

# 1 Summary

Stata/MP<sup>1</sup> is the version of Stata that is programmed to take full advantage of multicore and multiprocessor computers. It is exactly like Stata/SE in all ways except that it distributes many of Stata's most computationally demanding tasks across all the cores in your computer and thereby runs faster—much faster.

In a perfect world, software would run 2 times faster on 2 cores, 3 times faster on 3 cores, and so on. Stata/MP achieves about 78% efficiency. It runs 1.7 times faster on 2 cores, 2.6 times faster on 4 cores, and 3.3 times faster on 8 cores (see figure 1). Half the commands run faster than that. The other half run slower than the median speedup, and some of those commands are not sped up at all, either because they are inherently sequential (most time-series commands) or because they have not been parallelized (graphics, *mixed*).

In terms of evaluating average performance improvement, commands that take longer to run—such as estimation commands—are of greater importance. When estimation commands are taken as a group, Stata/MP achieves an even greater efficiency of approximately 85%. Taken at the median, estimation commands run 1.8 times faster on 2 cores, 3.1 times faster on 4 cores, and 4.2 times faster on 8 cores. Stata/MP supports up to 64 cores.

This paper provides a detailed report on the performance of Stata/MP. Command-by-command performance assessments are provided in section 8.

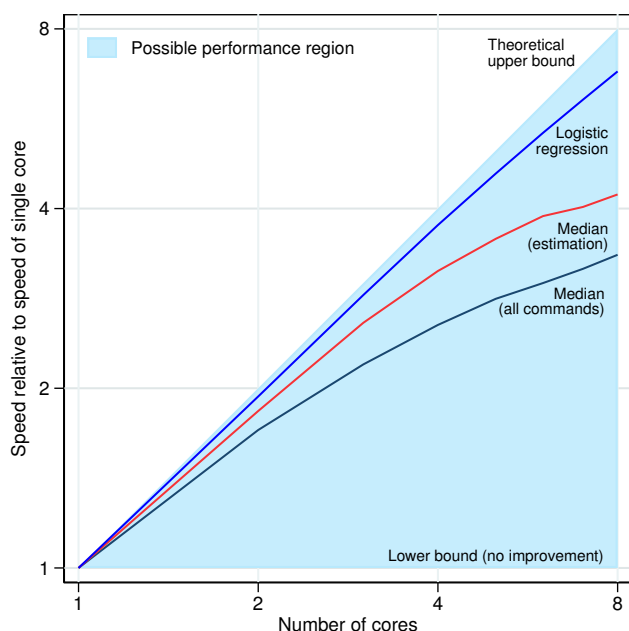


Figure 1. **Performance of Stata/MP.** Speed on multiple cores relative to speed on a single core.

1. Support for this effort was partially provided by the U.S. National Institutes of Health, National Institute on Aging grants 1R43AG019542-01A1, 2R44AG019542-02, and 5R44AG019542-03. We also thank Cornell Institute for Social and Economic Research (CISER) at Cornell University for graciously providing access to several highly parallel SMP platforms. CISER staff, in particular John Abowd, Kim Burlingame, Janet Heslop, and Lars Vilhuber, were exceptionally helpful in scheduling time and helping with configuration. The views expressed here do not necessarily reflect those of any of the parties thanked above.

## 2 Table of contents

|    |                                                         |     |
|----|---------------------------------------------------------|-----|
| 1  | Summary                                                 | 1   |
| 2  | Table of contents                                       | 2   |
| 3  | Introduction                                            | 3   |
| 4  | Parallel computing hardware                             | 4   |
| 5  | Constructing Stata/MP                                   | 5   |
| 6  | Measuring Stata/MP's performance                        | 5   |
| 7  | Performance summary                                     | 7   |
| 8  | Stata/MP performance, command by command                | 14  |
| 9  | Performance variability across computing platforms      | 42  |
| 10 | Hyperthreading—single- and multiple-processor platforms | 42  |
| A  | Performance assessment graphs for desktop computers     | 45  |
| B  | Performance assessment graphs for high-end servers      | 195 |
| C  | Command names and descriptions                          | 262 |
| D  | Problem sizes                                           | 290 |
| E  | Commands not assessed                                   | 319 |
| F  | Mata                                                    | 321 |
| G  | GLLAMM                                                  | 322 |

### 3 Introduction

Stata/MP was designed to take advantage of computers with multiple cores and multiple processors by partitioning the work among the multiple cores. From the outset, Stata/MP was required to be 100% compatible with all other flavors of Stata, including Stata/SE and Stata/IC. Stata/MP was also required to run all scripts, user-written programs, and analyses that run under existing Stata without any change or special action on the user's part.

Stata/MP runs on multicore and multiprocessor computers, including computers running MS Windows, Intel-based Mac OS computers, and Linux computers.

With multiple cores, one might expect to achieve the theoretical upper bound of doubling the speed by doubling the number of cores—2 cores run twice as fast as 1, 4 run twice as fast as 2, and so on. However, there are three reasons why such perfect scalability cannot be expected: 1) some calculations have parts that cannot be partitioned into parallel processes; 2) even when there are parts that can be partitioned, determining how to partition them takes computer time; and 3) multicore/multiprocessor systems only duplicate processors and cores, not all the other system resources.

Stata/MP achieved 78% efficiency overall and 85% efficiency among estimation commands.

Speed is more important for problems that are quantified as *large* in terms of the size of the dataset or some other aspect of the problem, such as the number of covariates. On large problems, Stata/MP with 2 cores runs half of Stata's commands at least 1.7 times faster than on a single core. With 4 cores, the same commands run at least 2.6 times faster than on a single core.

Figure 1, shown in the summary above, displays the theoretically possible performance as a shaded region. All Stata commands fall somewhere in the shaded region. Performance is measured as speed relative to speed on a single core: 1 indicates the speed on a single core; 2 means twice as fast as a single core; 4 means four times as fast as a single core; and so on. We could say the same thing in a different way: 2 means that a given problem runs in half the time required on a single core; 4 means that it runs in one-quarter the time; and so on.

The line in figure 1 for logistic regression reveals a speedup that is near the theoretical maximum. At the other end of the spectrum, some Stata commands experience no speedup at all. This is because their calculations are inherently sequential or because no effort was made to partition the work into parallel processes.

In typical use, Stata's estimation commands consume the bulk of the time required to perform analyses, so speeding them up was a priority for Stata programmers. Figure 1 also shows the median performance of Stata's estimation commands. The median estimation command runs 1.8 times faster on 2 cores and 3.1 times faster on 4 cores. Again, half the estimation commands speed up more than the median and half speed up less. Twenty-five percent of estimation commands speed up 2 times with 2 cores (the theoretical limit) and more than 3.9 times with 4 cores (this is not shown on the graph).

Figure 1 emphasizes dual-core, quad-core, and 8-core computers because those are the most common multicore platforms available. Stata/MP will work with up to 64 cores, however, and performance improvements continue to increase with more cores. For example, 25% of estimation commands run at least 6 times faster on 8-core computers, 11 times faster on 16-core computers, 16 times faster on

32-core computers, and 20 times faster on 64-core computers.

For assessments of performance gains of individual Stata commands, see section 8. See appendices A and B for results reported in graphical form.

## 4 Parallel computing hardware

Chip makers are increasing the number of cores on a computer processor, and computer makers are increasing the number of processors in a computer. Prior to 2005, chip makers essentially doubled the speed of computer processors every 18 months, a fact known informally as Moore's law ([Moore 1965](#)). The speed improvements were achieved by making components smaller—hence reducing electrical resistance—and by placing more transistors on a processor. Chip makers, however, are reaching the physical limits of what can be achieved through reduced size and increased complexity using existing technology. Although alternatives for further speeding up processors are on the horizon, these alternatives involve dramatic changes in technology and fabrication.

The other alternative to make computers run faster is simply to give you more processors or cores.

Modern computers run faster by having multiple processors in one box or multiple processors on one chip. When multiple processors are on one chip, the chip makers call such processors cores, and the chip they reside on is called a multicore processor. Each core is itself a processor that is bundled together with other cores onto a single chip.

Regardless, when they reside together in one box, all the processors and cores share the main memory, disk drives, and other devices on the computer. Most modern computers use multicore processors. Modern servers typically use multiple processors, each having multiple cores. Whether the cores are on one processor chip or on multiple processor chips does not much matter.

Following the lead of the chip makers, we are going to count cores and talk about cores on a computer. We are also going to use the term multicore to include both a single-processor computer whose processor has multiple cores and a multiprocessor computer whose processors also may have multiple cores.

Multicore designs work exceptionally well when running different programs simultaneously, especially when programs run independently. Hence, a 4-core computer can do almost as much work as four separate computers, and none of the programs needs to be modified to recognize that it is running in a multicore environment.

Single programs can take advantage of multicore environments, too, but the programs must be modified to do so. This modification is accomplished by allowing different parts of the program to run simultaneously in what are called separate execution threads. For example, a word processor might allow you to print and edit a document simultaneously. This type of threading is relatively easy to implement and is even allowed on single-core computers to make programs more convenient.

This type of threading adds convenience but does not address the issue of speeding the computations in a statistical package. To speed computations, a statistical package must be able to perform simultaneous computations on the same task. This ability is referred to as symmetric multiprocessing (SMP). Stata/MP is a modified version of Stata that uses SMP to speed up its computations.

Another type of parallel processing involves using multiple computers over a network. This type is known as cluster computing or distributed computing. Cluster computing requires problems that admit large-grain parallelization. Although cluster computing can be of interest in the computation of statistical results, Stata/MP does not address such parallel architectures.

For a thorough discussion of parallel processing, see [Culler, Singh, and Gupta \(1999\)](#) and [Grama et al. \(2003\)](#).

## 5 Constructing Stata/MP

For Stata to take advantage of multicore systems, sections of its code had to be rewritten to distribute their work across cores. Stata's internal design includes key algorithms that are used in many contexts. Once those key algorithms were rewritten, the benefits then spread themselves across Stata. Statistical computations lend themselves especially well to parallelization because observations are usually independent, and independent pieces can be calculated separately. One way parallelization happens is that many statistical computations can be partitioned over observations.

Parallelizing key algorithms resulted in a little more than half the observed performance gains. The remaining gains were achieved by modifying individual routines for important Stata commands and including custom code to parallelize them.

In all, approximately 250 sections of Stata's internal code were parallelized using the Open/MP API for developing SMP applications (see [Dagum and Menon \[1998\]](#)).

## 6 Measuring Stata/MP's performance

There is a theoretical limit to how much the performance of a program or command can be improved with multiple cores (or processors). With 2 cores, that limit is twice as fast (or half the run time); with 4 cores, the limit is 4 times as fast (or one-quarter the run time); and so on. This limit is called linear or perfect scaling.

Furthermore, not all algorithms or sections of code can be made to run in parallel. Some computations, or parts thereof, are inherently single threaded, for example, a formula that depends on prior values of itself, such as the autoregressive process:

$$y_t = \phi + \rho y_{t-1}$$

Statistical calculations are often more parallelizable than you might imagine. For instance, many inherently sequential computations can be parallelized when performed on longitudinal (panel) data because the dependencies that made the problem inherently sequential are broken at panel boundaries. Rather than partitioning on observations, Stata/MP partitions on panels. Whereas most time-series commands run only a little faster in the SMP environment, most panel-data commands run substantially faster.

Some sections of code are simply not worth the effort of parallelization because they take so little time to run or because parallelization would be technically difficult. Either way, the effort is just not worth the benefit.

Taken together, those sections of code are the nonparallelized region. Some authors refer to the parallelizable regions and the parallelized regions—the first referring to what could be parallelized and the second to what was actually parallelized—and even focus on the ratio between the two. We will focus on run times and their associated relative speeds, however, and draw no distinction between parallelizable and parallelized.

How much of a calculation has been parallelized is measurable, and measuring it is useful because it allows us to make extrapolations on how problems will run when the number of cores varies.

Figure 2 presents a stylized view of the component run times associated with a command that has been parallelized. Block *A* represents the time spent in parallelized regions of code; Block *B*, the unparallelized regions of code; and Block *C*, the additional overhead required for parallelization.

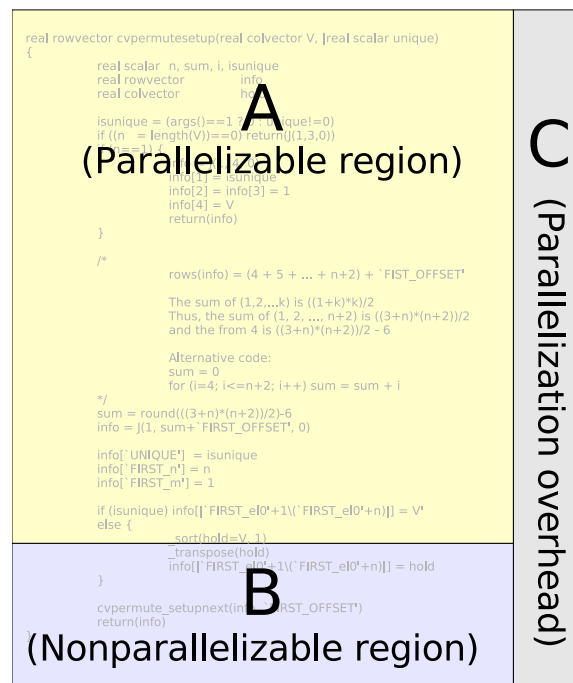


Figure 2. **Parallelization components.**

Let each letter represent an amount of time consumed in running a particular command on a particular dataset. Then  $A + B$  is the run time of the command when using a single core. If we parallelize the command, however, there is an additional time,  $C$ , associated with the overhead of partitioning the problem and coalescing the results from the cores.

If we know the percentage of time spent in  $A$ , then we have completely described the SMP performance of a command. Ignoring  $C$ , that is just  $100A/(A + B)$ .

We want to be more conservative, however, and account for the time required to parallelize the command. Considering  $C$  to be only the parts of the overhead that cannot be parallelized, we will refer to  $100A/(A + B + C)$  as the percentage parallelized:

$$\text{percentage parallelized} = \frac{100A}{A + B + C}$$

The percentage parallelized is a useful measure of how much performance will improve as cores (or processors) are added. All gains to parallelization occur because region  $A$  can be made to run on multiple cores at the same time. If we partition the region perfectly and each core runs uninterrupted, then when we double the number of cores, we halve the time used to perform  $A$ . As we add more cores, time spent in  $A$  continues to decrease. With 2 cores, it is  $A/2$ ; with 4 cores, it is  $A/4$ ; and with  $c$  cores, it is  $A/c$ . If we increase the number of cores without bound,  $A/c$  goes to zero. In contrast,  $B + C$  is a constant time for running the command; it cannot be reduced by adding more cores. As we add cores, the run time asymptotes to  $B + C$ .

We are ignoring another minor contribution to run time. Sometimes, overhead is associated with each core rather than, or in addition to, an overall parallelization overhead. Because of the methods used to build Stata/MP, this overhead is extremely small. In fact, it affects only four commands, and its effect on them is small.

The concept of percentage parallelizable helps clarify why some commands will have less-than-perfect scaling and allows results to be extrapolated to any number of cores. We also present performance results as simple relative speeds that can be read directly from tables or graphs to find the relative speed for multiple cores or processors compared with the speed for a single core or processor.

## 7 Performance summary

The performance of Stata/MP has been measured on 590 Stata commands. Excluding I/O commands, these 590 commands are most of the commands that take any appreciable time to run. Commands such as `display` (which writes output to the Results window) or `local` (which sets the value of a program macro) are not considered because they consume a negligible part of the time required to perform any analysis. Commands that run a target command repeatedly are not explicitly assessed, and some other commands are not timed for a variety of reasons; see appendix E. If you are searching this document for a specific command, know that we have tried to list every Stata command somewhere in the paper.

For each of the 590 commands, timings were recorded on a multicore computer where Stata/MP used 1, 2, ..., 40 cores to execute the same command. The computer contained four processors, each having 10 cores, for a total of 40 cores. All these timings were from the same installation of Stata/MP on the same computer. To reduce the impact of interruptions by the operating system, the timings were repeated three times and the shortest time was recorded.

Timings have also been performed on other dual-core, quad-core, 8-core, and 16-core computers.

Although speeds relative to a single core do vary among tested platforms, they are generally comparable, and the results presented are indicative of what can be expected across a spectrum of platforms. The results of the timings are presented in section 8, *Stata/MP performance, command by command*, and appendix A, *Performance assessment graphs for desktop computers*.

Appendix A, *Performance assessment graphs for desktop computers*, shows graphs for each of the 590 commands. Figure 3 shows the graph of Stata's logistic regression command, `logistic`:

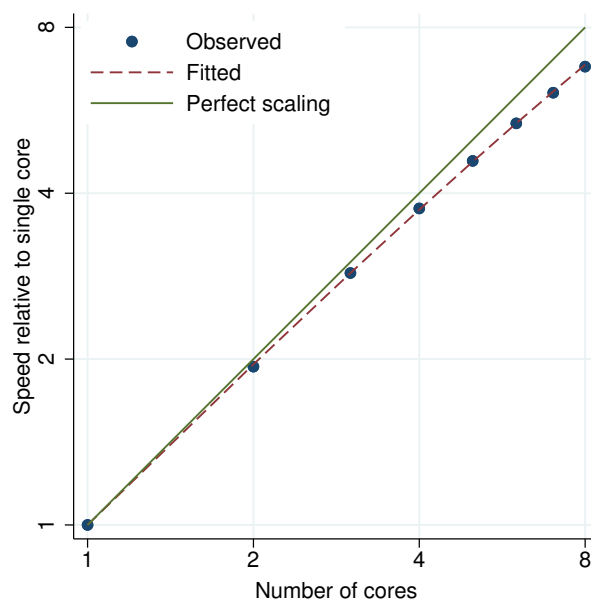


Figure 3. `logistic` performance plot.

The  $y$  axis shows speed relative to the speed on a single core. For `logistic`, the relative speeds are 1.9 (2 cores), 3.8 (4 cores), and 6.8 (8 cores). Also shown is a 45° reference line representing perfect scalability or, if you prefer, 100% parallelized: 2 times (2 cores), 4 times (4 cores), and 8 times (8 cores). `logistic` is 98% parallelized, but even so, you can see that its relative speeds are a bit below what is theoretically possible.

Stata's linear regression command, `regress`, very nearly achieves theoretical limits (see figure 4); its relative speeds increase in almost direct proportion to the number of cores.



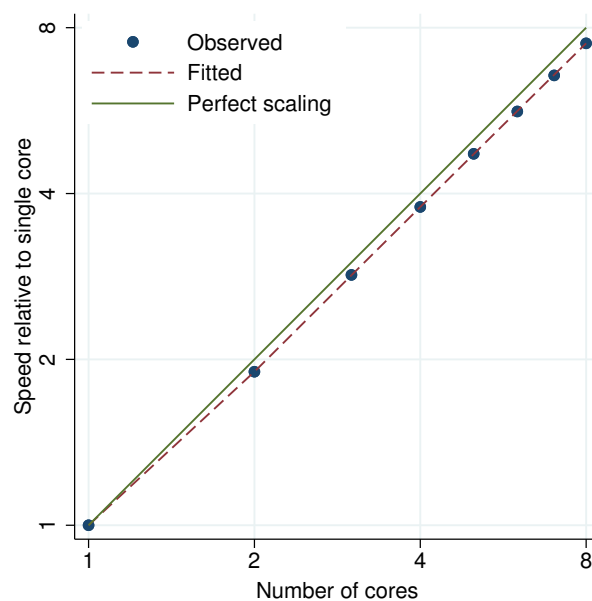


Figure 4. `regress` performance plot.

Figure 5 shows the graph for `arima`:

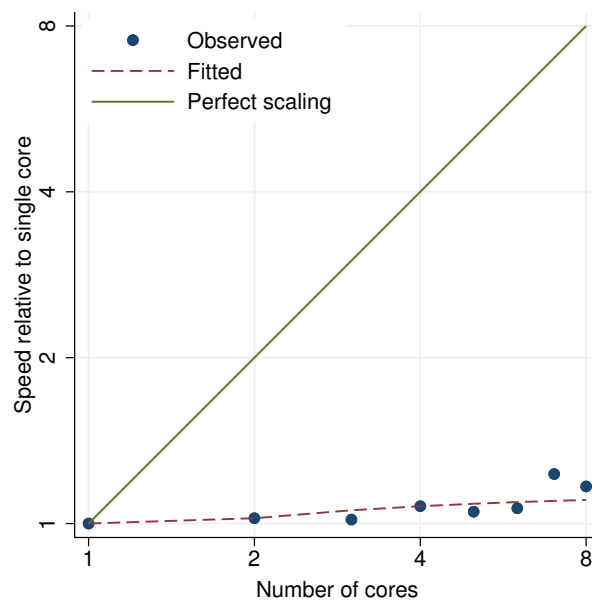


Figure 5. `arima` performance plot.

`arima`, a time-series command, hardly benefits from parallelization. Relative speeds are 1 (2 cores), 1.1 (4 cores), and 1.2 (8 cores).

Figure 6 shows the graph for Stata's command for Poisson regression with endogenous treatment effects, `etpoisson`:

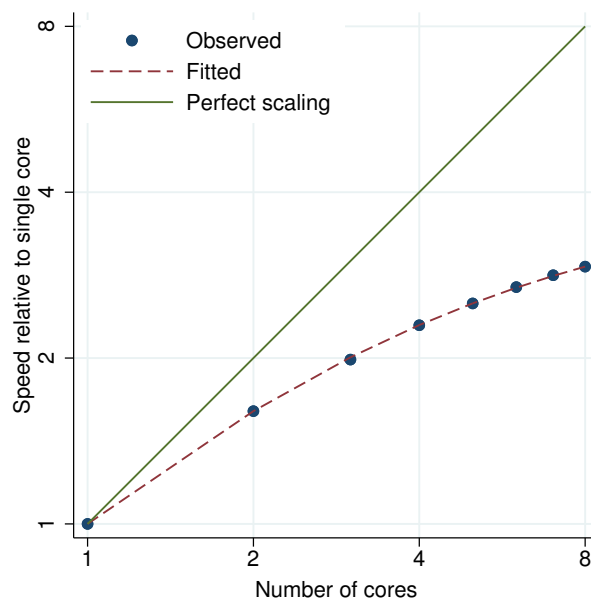


Figure 6. `etpoisson` performance plot.

Relative speeds are 1.6 (2 cores), 2.3 (4 cores), and 2.9 (8 cores). What is interesting about this graph is that the line flattens out as the number of cores increases. This is what happens when a command is not 100% parallelized: the relative run time approaches a horizontal asymptote that is related to the percent parallelized, which here is about 52%. Specifically, the asymptote is at  $1/\{1 - (\text{percentage parallelized})/100\}$ , which for `etpoisson` is about 2.1.

Finally, all 590 performance profiles can be combined into one figure, such as figure 7. The shaded area shows the region containing all possible performances. The diagonal top of the region represents perfect scaling (the maximum speed theoretically possible), while the horizontal lower boundary of the region represents no speed improvement.

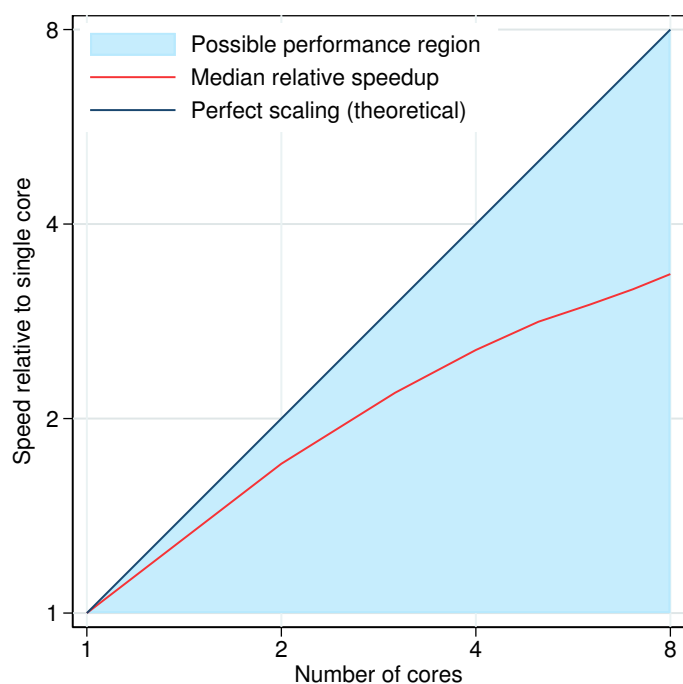


Figure 7. **Performance of Stata/MP.** Speed on multiple cores relative to speed on a single core.

Also included are the median results over all 590 commands; 295 commands have better performance gains (their curves lie above the median relative speedup line), and 295 exhibit lesser performance gains (their curves lie below the line).

Median performance for most Stata users will be better than median performance across commands as we calculated it. To be able to measure performance, we had to choose large problems even when, for a particular command, large problems are rarely run. For instance, few users would run analyses that spend as much time running *t* tests as did those analyses we had to run to record reliable results. Stata's command for *t* tests runs quickly on single or multiple cores. Meanwhile, Stata/MP development efforts focused on improving run times of commands that require substantial run times. Ergo, the median improvements are understated.

Figure 8 better illustrates the distribution of results by showing not just the median but also the quartiles. The most interesting thing about figure 8 is the first quartile (light-blue swath at the top). It shows that 25% of commands exhibit nearly perfect scaling. The worst commands among this group run about 2 times faster on 2 cores, 3.7 times faster on 4 cores, and 6.6 times faster on 8 cores.

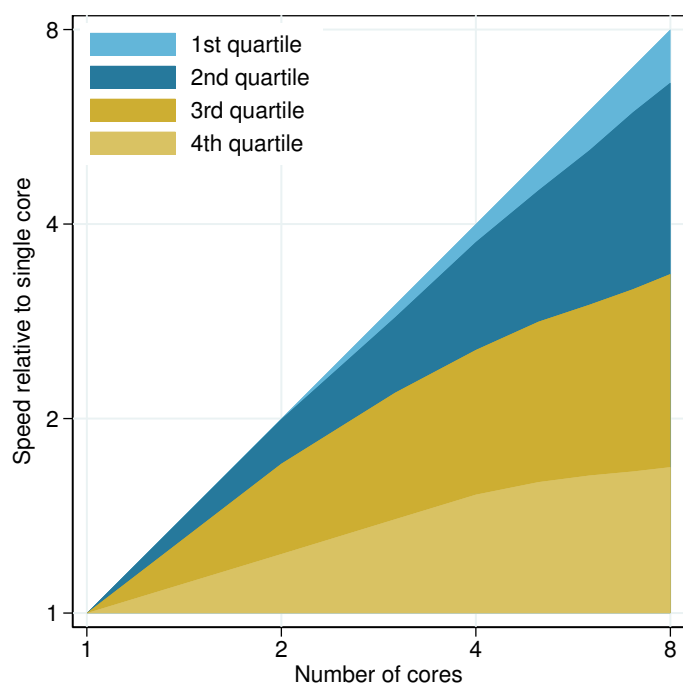


Figure 8. **Quartiles of Stata/MP performance.**  
Speed on multiple cores relative to speed on a single core.

Figures 7 and 8 present results for all commands, whereas the time required by most analyses is dominated by execution of estimation commands. Estimation commands tend to be the most computationally intensive, particularly those that require iterative solutions.

Figure 9 summarizes the observed performance and median performance for the 349 estimation commands. These include all the estimation commands in Stata, and some commands are included more than once to include critical options, such as `vce(robust)` and `vce(cluster)` for robust standard errors and correlation within groups. The options themselves are not important; what is important is that these options and a few others like them substantively affect how the calculation proceeds and thus affect speed.

Compared with figure 7, figure 9 shows that the median performance for estimation commands is better than the overall median. The median relative speed for estimation commands is 1.8 times faster on 2 cores, 3.1 times faster on 4 cores, and 4.2 times faster on 8 cores. Half of all estimation commands perform even better. Figure 10 reveals that only 25% of all estimation commands run less than 1.5 times faster on 2 cores, less than 2.1 times faster on 4 cores, and less than 2.5 times faster on 8 cores.

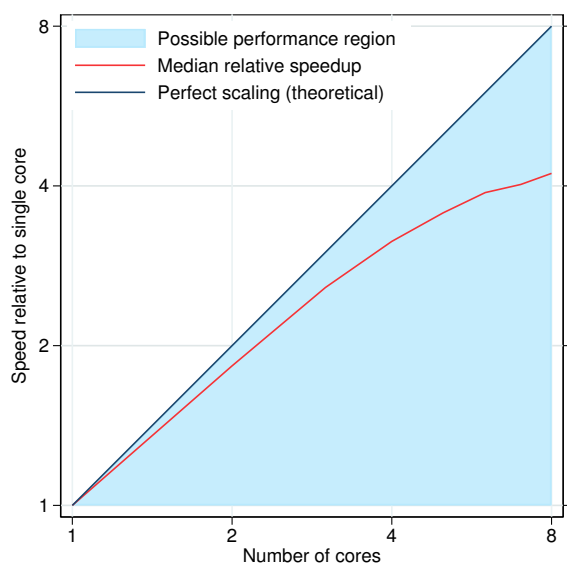


Figure 9. **Performance of Stata/MP on estimation commands.** Speed on multiple cores relative to speed on a single core.

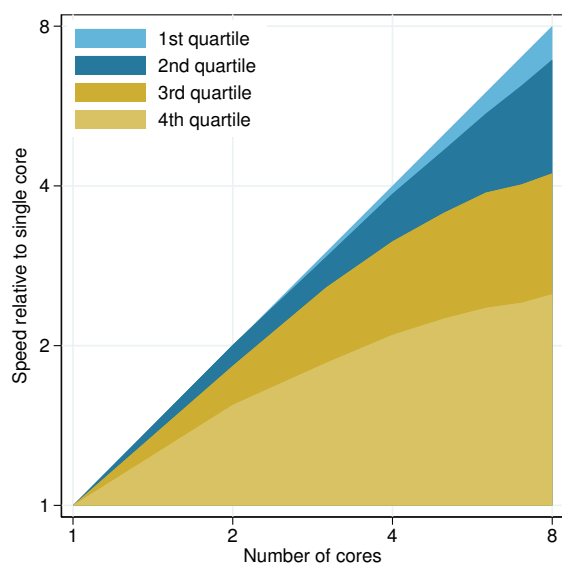


Figure 10. **Quartiles of Stata/MP performance on estimation commands.** Speed on multiple cores relative to speed on a single core.

We have emphasized results on 2, 4, and 8 cores because those are the most common desktop architectures currently available. Stata/MP supports up to 64 cores, and performance continues to improve as cores are added. Figure 11 shows the performance boundary and median for all 349 estimation commands on 1–40-core computers, and figure 12 shows their performance quartiles.

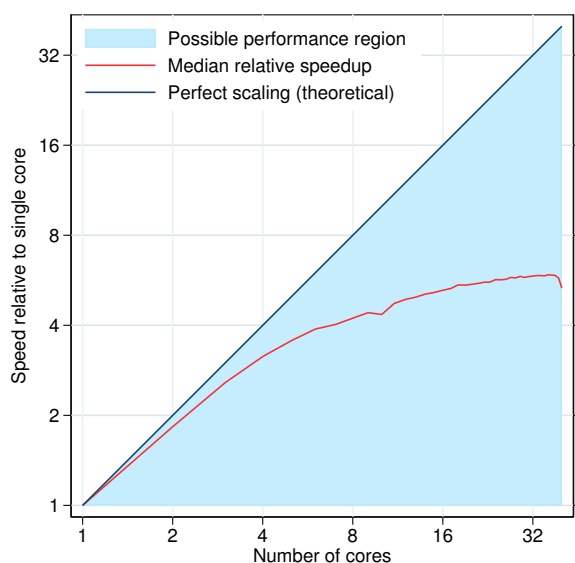


Figure 11. **Performance of Stata/MP on estimation commands (1 to 40 cores).** Speed of estimation commands on multiple cores relative to speed on a single core.

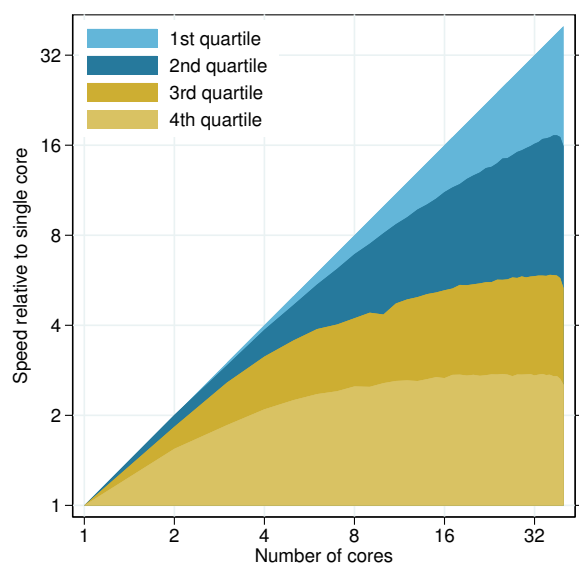


Figure 12. **Quartiles of Stata/MP performance on estimation commands (1 to 40 cores).** Speed of estimation commands on multiple cores relative to speed on a single core.

## 8 Stata/MP performance, command by command

The performance summaries from the prior section provide an overall sense of the performance of Stata/MP but will not reflect the experience of most users. Few users perform all the commands in Stata, and no users perform them with equal frequency. Most users will be interested in a subset of commands and often in only a few commands that they use regularly on large problems.

Table 1, toward the end of this section, provides relative speeds on individual commands, comparing the speed on 2, 4, 8, and 16 cores with the speed on a single core. The table also reports the degree to which each command is parallelized.

All commands were run on moderately-large-to-very-large problems. The goal was to measure performance on problems that required substantial time to solve and that were large enough to measure performance gains on 8, 16, 32, or even 64 cores. For commands that are parallelized, such problems have a larger parallelizable region (*A*) relative to the unparallelizable region (*B*) and are thus more amenable to parallelization, particularly when run on many cores. Longer timings also ameliorate variations in timings, such as interruptions caused by operating system processes or the memory status of the system when the command begins. Run-to-run variations are much greater for smaller problems that have shorter run times.

Timings were typically performed on commands that took 1–2 minutes to run on a single-core computer running at 2.2–3.4 GHz. For some commands, this meant that the problems used extremely large numbers of observations or covariates, because some commands are inherently fast. For others, the problems were smaller because the commands are inherently slow, because of, for example, iterative or even simulated solutions. For details on the sizes of the problems, see appendix D.

Stata/MP was designed mostly to improve performance on large problems, such as those reported in appendix D. Even so, the performance on small-to-moderate problems improves surprisingly well. Using the same commands as those in appendix D, but with problems 100 to 10,000 times smaller and run times of 0.4 seconds to just over 4 seconds on a machine running at 2.2–3.4 GHz, substantial speedups were still observed. Among commands that were at least 50% parallelized, more than half exhibited greater than 90% of the speedup exhibited on the larger problems. These are typical results. Run times for smaller problems vary more from computer to computer because small problems are more sensitive to the architecture of the computer, processor, and operating system.

All values were obtained from the minimum of three runs on a 40-core computer.

Stata/MP performance was tested on many computers under MS Windows, Mac, and Linux operating systems. Although performance varies somewhat across platforms, the results from the table below can reasonably be applied to any platform.

Most users should simply look at the column reporting results for the number of cores in which they are interested. This column estimates the speed on that number of cores relative to the speed on a single core. Given a computer with a known number of cores, this column of results is the most direct measure of performance improvement.

Relative speed is easy to understand. When relative speed is 2, you could run a given problem twice in the same amount of time that you could run it once on a single-core computer. When relative speed is 4, you could run a given problem four times, and so on. Equivalently, when relative speed is 2, you could run a given problem in half the time that you could run it on a single-core computer. When relative speed is 4, you could run a given problem in one-fourth the time, and so on.

Table 1 also presents the percentage parallelized discussed in section 6. Given a set of percentage run times (relative to the run times on a single core) for at least 3 different numbers of cores, we can estimate the percentage parallelized and parallelization overhead parameters. The form of the model is particularly simple:

$$\text{percentage run time} = \alpha + \widehat{PP} \frac{1}{c} + \widehat{O} \frac{\delta_1}{c} \quad (1)$$

where  $c$  is the number of cores,  $\delta_1$  is an indicator for  $c > 1$ , and  $\alpha$  is an intercept.

Our parameters of interest are directly estimated:

$$\text{percentage parallelized} = \widehat{PP} \quad (2)$$

and

$$\text{parallelization overhead} = \widehat{O} \quad (3)$$

Equation (1) is estimated by median regression (`qreg`) using Stata. Median regression is used in preference to ordinary least squares (OLS) because occasionally a timing will be far too large because of interruptions from the operating system. Such effects are ignored in median regression.

The estimated value for parallelization overhead is particularly sensitive to the computing platform, and so we do not report it here. Note from equation (1) that  $\widehat{O}$  captures any unexpected difference in the speed when using one core. Because different computer, processor, cache, and operating system architectures respond differently in moving from 1 to 2 cores,  $\widehat{O}$  captures not only the theoretical parallelization overhead, but also anything that causes the time from the first core to differ from the time from the second.

Percentage parallelized is the most concrete measure of how a command responds to more cores. For most commands, the run time in this percentage of the code falls by half for each doubling of the number of cores.

The estimated percentage parallelized is also the most comparable measure across computing platforms; it is very consistent from one platform to another. Most of the differences across computing platforms are captured in  $\widehat{O}$ . Because the relative speeds are compared with the run time on a single core, they necessarily include the parallelization overhead and are thus not quite as comparable across machines.

Each line in the table represents a command run on a particular problem. The command column shows the Stata command name and relevant options. For those unfamiliar with Stata syntax, appendix C provides short descriptions of what each command does. To learn more about any command, including worked examples, all of the Stata manuals can be access from <http://www.stata.com/features/documentation/>.

Appendix A contains performance graphs for each command using 1–8 cores. Appendix B contains graphs using 1–40 cores. The graphs plot the observed relative speed, the modeled performance using equation (1), and the perfect scalability reference line. If you are reading the PDF version of this document, you can click on the command name in table 1 to go to the page with the associated graph.



Table 1. Stata/MP performance, command by command

| Command               | Speed relative to a single core <sup>a</sup> |     |     |      | Percentage<br>parallelized <sup>b</sup> |
|-----------------------|----------------------------------------------|-----|-----|------|-----------------------------------------|
|                       | Number of cores                              |     |     |      |                                         |
|                       | 2                                            | 4   | 8   | 16   |                                         |
| alpha                 | 1.8                                          | 3.0 | 4.4 | 5.6  | 87                                      |
| ameans                | 1.8                                          | 2.7 | 3.7 | 4.6  | 83                                      |
| anova (one-way)       | 1.9                                          | 3.6 | 6.4 | 11.0 | 97                                      |
| anova (two-way)       | 2.6                                          | 4.7 | 7.7 | 11.2 | 95                                      |
| arch                  | 1.0                                          | 1.3 | 1.3 | 1.4  | 28                                      |
| areg                  | 2.0                                          | 3.6 | 5.4 | 7.0  | 90                                      |
| areg, vce(cluster)    | 2.0                                          | 3.4 | 5.4 | 7.4  | 92                                      |
| areg, vce(robust)     | 2.1                                          | 3.5 | 5.6 | 8.0  | 93                                      |
| arfima                | 1.1                                          | 1.1 | 1.1 | 1.1  | 4                                       |
| arima                 | 1.0                                          | 1.1 | 1.2 | 1.2  | 16                                      |
| bayes: logit          | 2.0                                          | 4.0 | 7.3 | 12.0 | 97                                      |
| bayes: poisson        | 1.7                                          | 2.7 | 3.9 | 5.0  | 86                                      |
| bayes: regress        | 1.9                                          | 3.4 | 5.4 | 7.4  | 91                                      |
| bayesmh logit         | 1.5                                          | 2.0 | 2.3 | 2.3  | 57                                      |
| bayesmh mvn           | 1.1                                          | 1.2 | 1.2 | 1.3  | 22                                      |
| bayesmh mylogit       | 1.4                                          | 1.8 | 2.1 | 2.4  | 62                                      |
| bayesmh nl            | 1.1                                          | 1.2 | 1.4 | 1.5  | 37                                      |
| bayesmh normal        | 1.6                                          | 2.1 | 2.6 | 2.7  | 66                                      |
| bayesmh normal gibbs  | 1.1                                          | 1.2 | 1.2 | 1.2  | 18                                      |
| bayesmh normal re     | 1.4                                          | 1.7 | 1.9 | 1.9  | 48                                      |
| betareg, link(logit)  | 2.0                                          | 3.8 | 7.0 | 11.7 | 97                                      |
| betareg, link(probit) | 2.0                                          | 3.8 | 6.9 | 11.7 | 97                                      |
| binreg                | 2.0                                          | 3.7 | 6.3 | 10.0 | 96                                      |
| biplot                | 1.0                                          | 1.0 | 1.0 | 1.0  | 0                                       |
| biprobit              | 1.9                                          | 3.6 | 6.9 | 12.5 | 98                                      |

All values are expressed as the speed relative to the speed of a single core.

*a.* Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

*b.* Bigger is better; 100 is perfect.

Table 1. Stata/MP performance, command by command

| Command                        | Speed relative to a single core <sup>a</sup> |      |      |      | Percentage<br>parallelized <sup>b</sup> |
|--------------------------------|----------------------------------------------|------|------|------|-----------------------------------------|
|                                | Number of cores                              |      |      |      |                                         |
|                                | 2                                            | 4    | 8    | 16   |                                         |
| biprobit (seemingly unrelated) | 1.9                                          | 4.0  | 7.4  | 12.6 | 98                                      |
| bitest                         | 1.7                                          | 2.7  | 3.8  | 4.8  | 85                                      |
| blogit                         | 2.0                                          | 3.7  | 6.7  | 11.1 | 97                                      |
| boxcox                         | 1.9                                          | 3.6  | 6.4  | 10.5 | 96                                      |
| bprobit                        | 1.9                                          | 3.6  | 6.4  | 10.8 | 97                                      |
| brier                          | 1.2                                          | 1.4  | 1.5  | 1.6  | 42                                      |
| bsample                        | 1.0                                          | 1.2  | 1.2  | 1.5  | 26                                      |
| bstat                          | 1.7                                          | 3.0  | 4.2  | 5.2  | 85                                      |
| by: generate                   | 2.7                                          | 5.8  | 11.6 | 23.1 | 100                                     |
| by: generate (small groups)    | 8.5                                          | 14.6 | 31.0 | 30.9 | 98                                      |
| by: replace                    | 2.9                                          | 5.8  | 11.7 | 23.2 | 100                                     |
| by: replace (small groups)     | 7.6                                          | 13.9 | 19.8 | 28.6 | 98                                      |
| ca                             | 1.5                                          | 1.9  | 2.3  | 2.6  | 64                                      |
| candisc                        | 2.2                                          | 4.1  | 7.3  | 12.0 | 97                                      |
| canon                          | 1.6                                          | 2.7  | 4.5  | 6.9  | 93                                      |
| cc                             | 1.2                                          | 1.3  | 1.3  | 1.4  | 44                                      |
| by: cc                         | 1.0                                          | 1.0  | 1.0  | 1.0  | 1                                       |
| centile                        | 1.0                                          | 1.0  | 1.0  | 1.0  | 1                                       |
| churdle linear                 | 1.8                                          | 3.2  | 5.0  | 6.7  | 90                                      |
| ci means                       | 1.5                                          | 2.7  | 3.2  | 3.5  | 74                                      |
| ci means, poisson              | 1.4                                          | 2.0  | 2.6  | 3.1  | 73                                      |
| ci proportions                 | 1.5                                          | 2.5  | 2.8  | 3.6  | 72                                      |
| clogit (k1 to k2 matching)     | 1.9                                          | 3.5  | 6.0  | 9.6  | 95                                      |
| clogit (1 to k matching)       | 1.5                                          | 2.1  | 2.3  | 2.7  | 66                                      |
| cloglog                        | 1.9                                          | 3.7  | 6.7  | 11.4 | 97                                      |

All values are expressed as the speed relative to the speed of a single core.

a. Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

b. Bigger is better; 100 is perfect.

Table 1. Stata/MP performance, command by command

| Command                 | Speed relative to a single core <sup>a</sup> |     |     |      | Percentage<br>parallelized <sup>b</sup> |
|-------------------------|----------------------------------------------|-----|-----|------|-----------------------------------------|
|                         | Number of cores                              |     |     |      |                                         |
|                         | 2                                            | 4   | 8   | 16   |                                         |
| cluster averagelinkage  | 2.2                                          | 4.3 | 8.1 | 15.2 | 99                                      |
| cluster centroidlinkage | 2.1                                          | 4.1 | 8.0 | 15.2 | 99                                      |
| cluster completelinkage | 2.2                                          | 4.3 | 8.1 | 14.9 | 99                                      |
| cluster generate        | 1.3                                          | 1.4 | 1.5 | 1.5  | 32                                      |
| cluster kmeans          | 2.2                                          | 4.2 | 8.2 | 15.3 | 99                                      |
| cluster kmedians        | 2.2                                          | 4.3 | 8.4 | 16.1 | 99                                      |
| cluster medianlinkage   | 2.1                                          | 4.2 | 8.0 | 14.7 | 99                                      |
| cluster singlelinkage   | 1.0                                          | 0.9 | 0.9 | 0.9  | 0                                       |
| cluster wardslinkage    | 2.0                                          | 3.8 | 7.3 | 13.3 | 99                                      |
| cluster waveragelinkage | 2.0                                          | 3.9 | 7.3 | 13.3 | 99                                      |
| cmclogit                | 2.0                                          | 3.5 | 5.5 | 7.4  | 92                                      |
| cmmprobit               | 1.6                                          | 2.3 | 2.9 | 3.2  | 70                                      |
| cmroprobit              | 1.7                                          | 2.4 | 3.0 | 3.3  | 71                                      |
| cnsreg                  | 2.0                                          | 4.0 | 7.9 | 15.3 | 100                                     |
| codebook                | 3.1                                          | 5.6 | 8.2 | 12.0 | 95                                      |
| collapse                | 1.3                                          | 1.8 | 1.9 | 2.0  | 57                                      |
| compare                 | 1.8                                          | 2.9 | 4.2 | 5.4  | 87                                      |
| compress                | 1.0                                          | 1.0 | 1.0 | 1.0  | 0                                       |
| contract                | 1.5                                          | 1.5 | 1.6 | 1.6  | 36                                      |
| corr2data               | 1.8                                          | 3.4 | 5.4 | 6.6  | 92                                      |
| correlate               | 1.9                                          | 3.8 | 7.6 | 15.2 | 100                                     |
| corrgram                | 1.1                                          | 1.3 | 1.2 | 1.5  | 19                                      |
| count                   | 2.0                                          | 4.0 | 8.0 | 16.0 | 100                                     |
| cpoisson                | 1.6                                          | 2.2 | 2.7 | 3.0  | 71                                      |
| cs                      | 1.2                                          | 1.3 | 1.3 | 1.4  | 28                                      |

All values are expressed as the speed relative to the speed of a single core.

*a.* Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

*b.* Bigger is better; 100 is perfect.

Table 1. Stata/MP performance, command by command

| Command              | Speed relative to a single core <sup>a</sup> |     |     |      | Percentage<br>parallelized <sup>b</sup> |
|----------------------|----------------------------------------------|-----|-----|------|-----------------------------------------|
|                      | Number of cores                              |     |     |      |                                         |
|                      | 2                                            | 4   | 8   | 16   |                                         |
| by: cs               | 1.0                                          | 1.0 | 1.0 | 1.0  | 0                                       |
| ctset                | 2.1                                          | 4.1 | 8.2 | 16.4 | 100                                     |
| cttost               | 1.0                                          | 1.0 | 1.0 | 1.0  | 5                                       |
| cumul                | 1.0                                          | 1.0 | 1.0 | 1.0  | 5                                       |
| cusum                | 1.1                                          | 1.1 | 1.2 | 1.3  | 18                                      |
| datasignature        | 1.0                                          | 1.0 | 1.0 | 1.0  | 0                                       |
| decode               | 1.0                                          | 1.0 | 1.0 | 1.0  | 2                                       |
| destring             | 1.0                                          | 1.1 | 1.1 | 1.2  | 24                                      |
| dfactor              | 1.5                                          | 1.9 | 2.3 | 2.3  | 56                                      |
| dfgls                | 1.1                                          | 1.2 | 1.3 | 1.3  | 27                                      |
| dfuller              | 2.8                                          | 3.7 | 4.4 | 4.8  | 80                                      |
| discrim knn          | 1.7                                          | 2.5 | 3.3 | 3.8  | 77                                      |
| discrim lda          | 2.1                                          | 3.6 | 5.7 | 8.1  | 93                                      |
| discrim logistic     | 2.1                                          | 4.1 | 7.6 | 13.6 | 98                                      |
| discrim qda          | 1.7                                          | 2.4 | 3.1 | 3.7  | 77                                      |
| dotplot              | 1.1                                          | 1.2 | 1.2 | 1.3  | 22                                      |
| drawnorm             | 2.0                                          | 3.7 | 6.3 | 9.8  | 95                                      |
| drop if <i>exp</i>   | 1.0                                          | 1.0 | 1.0 | 1.0  | 0                                       |
| drop in <i>range</i> | 1.0                                          | 1.0 | 1.0 | 1.0  | 0                                       |
| dsge                 | 1.0                                          | 1.0 | 1.0 | 1.0  | 0                                       |
| dsgenl               | 1.0                                          | 1.0 | 1.0 | 1.0  | 0                                       |
| dslogit              | 1.3                                          | 1.9 | 2.4 | 2.3  | 65                                      |
| dspoisson            | 1.5                                          | 2.0 | 2.4 | 2.7  | 66                                      |
| dsregress            | 1.4                                          | 1.6 | 2.2 | 2.0  | 62                                      |
| dstdize              | 1.0                                          | 1.0 | 1.0 | 1.0  | 0                                       |

All values are expressed as the speed relative to the speed of a single core.

a. Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

b. Bigger is better; 100 is perfect.

Table 1. Stata/MP performance, command by command

| Command                      | Speed relative to a single core <sup>a</sup> |     |     |      | Percentage<br>parallelized <sup>b</sup> |
|------------------------------|----------------------------------------------|-----|-----|------|-----------------------------------------|
|                              | Number of cores                              |     |     |      |                                         |
|                              | 2                                            | 4   | 8   | 16   |                                         |
| dvech                        | 1.1                                          | 1.0 | 1.1 | 1.1  | 6                                       |
| egen group()                 | 1.9                                          | 3.1 | 4.3 | 5.2  | 85                                      |
| by: egen mean                | 1.1                                          | 2.1 | 2.4 | 2.4  | 58                                      |
| eivreg                       | 1.9                                          | 3.9 | 7.6 | 14.2 | 99                                      |
| encode                       | 1.6                                          | 2.4 | 3.3 | 3.5  | 78                                      |
| eregress                     | 1.8                                          | 2.6 | 3.5 | 4.0  | 79                                      |
| esize twosample              | 1.4                                          | 1.8 | 2.1 | 2.3  | 60                                      |
| esize unpaired               | 1.6                                          | 2.4 | 3.1 | 3.7  | 78                                      |
| eteffects (exponential), ate | 1.1                                          | 2.9 | 4.0 | 2.4  | 84                                      |
| eteffects (linear), ate      | 2.0                                          | 3.5 | 5.6 | 7.9  | 93                                      |
| eteffects (linear), pomeans  | 2.0                                          | 3.5 | 5.7 | 8.0  | 93                                      |
| eteffects (probit), ate      | 2.0                                          | 3.5 | 5.6 | 7.9  | 92                                      |
| etpoisson                    | 1.6                                          | 2.3 | 2.9 | 2.5  | 52                                      |
| etregress, poutcomes         | 1.9                                          | 3.2 | 4.7 | 6.1  | 88                                      |
| etregress, twostep           | 1.9                                          | 3.7 | 6.8 | 12.1 | 98                                      |
| exlogistic                   | 1.0                                          | 1.0 | 1.0 | 1.0  | 2                                       |
| expand #                     | 1.0                                          | 1.0 | 1.0 | 1.0  | 0                                       |
| expand varname               | 1.0                                          | 1.0 | 1.0 | 1.0  | 0                                       |
| expandcl #                   | 1.6                                          | 2.0 | 2.5 | 2.9  | 76                                      |
| expandcl varname             | 1.6                                          | 2.4 | 3.1 | 3.6  | 76                                      |
| expoisson                    | 1.1                                          | 1.1 | 1.1 | 1.1  | 11                                      |
| factor                       | 1.2                                          | 1.1 | 1.2 | 1.8  | 36                                      |
| fcast compute                | 1.0                                          | 1.0 | 1.0 | 1.0  | 1                                       |
| fillin                       | 1.6                                          | 2.4 | 3.2 | 3.9  | 79                                      |
| fmm 2: poisson               | 1.7                                          | 2.4 | 2.9 | 3.2  | 70                                      |

All values are expressed as the speed relative to the speed of a single core.

*a.* Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

*b.* Bigger is better; 100 is perfect.

Table 1. Stata/MP performance, command by command

| Command                      | Speed relative to a single core <sup>a</sup> |     |      |      | Percentage<br>parallelized <sup>b</sup> |
|------------------------------|----------------------------------------------|-----|------|------|-----------------------------------------|
|                              | Number of cores                              |     |      |      |                                         |
|                              | 2                                            | 4   | 8    | 16   |                                         |
| fmm 2: regress               | 1.6                                          | 2.0 | 2.3  | 2.4  | 61                                      |
| fmm 3: poisson               | 1.5                                          | 2.0 | 2.3  | 2.5  | 61                                      |
| fmm 3: regress               | 1.4                                          | 1.8 | 1.7  | 1.9  | 39                                      |
| fracreg probit               | 2.2                                          | 4.3 | 8.2  | 14.8 | 99                                      |
| frontier                     | 2.0                                          | 3.9 | 7.5  | 13.6 | 99                                      |
| fvrevar (factors)            | 2.5                                          | 3.3 | 6.3  | 11.0 | 97                                      |
| fvrevar (interaction)        | 1.9                                          | 3.3 | 5.3  | 7.2  | 90                                      |
| generate (small expressions) | 3.2                                          | 6.2 | 11.3 | 19.3 | 99                                      |
| generate                     | 2.0                                          | 4.1 | 8.1  | 16.1 | 100                                     |
| glm, family(gamma)           | 2.0                                          | 3.8 | 6.6  | 10.5 | 97                                      |
| glm, family(gaussian)        | 1.9                                          | 3.7 | 6.6  | 11.2 | 96                                      |
| glm, family(igaussian)       | 2.0                                          | 3.9 | 7.0  | 11.8 | 97                                      |
| glm, family(nbinomial)       | 2.0                                          | 4.0 | 7.3  | 12.4 | 98                                      |
| glm, family(poisson)         | 2.1                                          | 4.1 | 7.4  | 12.7 | 98                                      |
| glogit                       | 2.1                                          | 4.0 | 7.7  | 14.7 | 99                                      |
| gmm                          | 1.3                                          | 1.6 | 1.7  | 1.7  | 40                                      |
| gmm (with derivatives)       | 2.0                                          | 3.5 | 5.4  | 6.7  | 90                                      |
| gprobit                      | 2.0                                          | 3.9 | 7.5  | 14.1 | 99                                      |
| graph bar                    | 1.2                                          | 2.5 | 2.6  | 2.6  | 64                                      |
| graph box                    | 1.1                                          | 1.1 | 1.1  | 1.1  | 10                                      |
| graph pie                    | 1.2                                          | 1.3 | 1.4  | 1.4  | 31                                      |
| grmeanby                     | 1.3                                          | 1.3 | 2.2  | 2.8  | 70                                      |
| gsem, oprobit (CFA, 2-level) | 2.0                                          | 3.1 | 4.2  | 4.4  | 77                                      |
| gsem, oprobit (CFA)          | 2.1                                          | 3.0 | 3.5  | 3.6  | 70                                      |
| gsem, logit group()          | 1.7                                          | 2.7 | 3.5  | 4.1  | 80                                      |

All values are expressed as the speed relative to the speed of a single core.

a. Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

b. Bigger is better; 100 is perfect.

Table 1. Stata/MP performance, command by command

| Command               | Speed relative to a single core <sup>a</sup> |     |     |      | Percentage<br>parallelized <sup>b</sup> |
|-----------------------|----------------------------------------------|-----|-----|------|-----------------------------------------|
|                       | Number of cores                              |     |     |      |                                         |
|                       | 2                                            | 4   | 8   | 16   |                                         |
| gsem, group()         | 1.8                                          | 2.8 | 3.8 | 4.7  | 83                                      |
| gsem, ologit group()  | 1.7                                          | 2.6 | 3.4 | 3.9  | 79                                      |
| gsem, poisson group() | 1.8                                          | 2.7 | 3.5 | 4.2  | 81                                      |
| gsort                 | 1.1                                          | 1.2 | 1.3 | 1.3  | 25                                      |
| hausman               | 1.2                                          | 1.3 | 1.3 | 1.3  | 23                                      |
| heckman               | 1.8                                          | 3.4 | 6.4 | 11.0 | 97                                      |
| heckman, twostep      | 1.9                                          | 3.6 | 6.7 | 11.5 | 98                                      |
| heckoprobit           | 2.0                                          | 3.8 | 7.0 | 12.1 | 98                                      |
| heckpoisson           | 1.8                                          | 3.0 | 4.4 | 5.5  | 86                                      |
| heckprob              | 1.9                                          | 3.8 | 7.0 | 11.8 | 97                                      |
| hetoprobit            | 2.0                                          | 3.5 | 5.6 | 7.8  | 92                                      |
| hetprob               | 1.8                                          | 3.9 | 6.4 | 9.8  | 95                                      |
| hetregress            | 1.9                                          | 3.7 | 6.6 | 10.5 | 96                                      |
| hetregress, twostep   | 2.0                                          | 3.6 | 6.7 | 11.5 | 97                                      |
| histogram             | 1.4                                          | 1.7 | 1.9 | 2.0  | 54                                      |
| hotelling             | 1.7                                          | 3.6 | 7.1 | 13.5 | 99                                      |
| icc, mixed            | 1.2                                          | 1.8 | 2.2 | 2.4  | 65                                      |
| icc (one-way)         | 1.3                                          | 2.0 | 2.5 | 3.0  | 72                                      |
| icc (two-way)         | 1.1                                          | 1.7 | 2.1 | 2.3  | 63                                      |
| intreg                | 2.0                                          | 4.1 | 7.7 | 14.1 | 99                                      |
| ir                    | 1.3                                          | 1.6 | 1.7 | 1.9  | 50                                      |
| by: ir                | 1.0                                          | 1.0 | 1.0 | 1.0  | 0                                       |
| irf create            | 1.3                                          | 1.7 | 1.9 | 2.1  | 57                                      |
| irt 1pl               | 1.6                                          | 2.3 | 2.8 | 2.8  | 62                                      |
| irt 2pl               | 1.6                                          | 2.3 | 2.9 | 2.9  | 63                                      |

All values are expressed as the speed relative to the speed of a single core.

*a.* Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

*b.* Bigger is better; 100 is perfect.

Table 1. Stata/MP performance, command by command

| Command                       | Speed relative to a single core <sup>a</sup> |     |     |      | Percentage<br>parallelized <sup>b</sup> |
|-------------------------------|----------------------------------------------|-----|-----|------|-----------------------------------------|
|                               | Number of cores                              |     |     |      |                                         |
|                               | 2                                            | 4   | 8   | 16   |                                         |
| irt 3pl                       | 1.7                                          | 2.0 | 2.1 | 2.0  | 44                                      |
| irt grm                       | 1.6                                          | 2.3 | 2.8 | 2.7  | 63                                      |
| irt nrm                       | 1.5                                          | 2.1 | 2.5 | 2.4  | 57                                      |
| irt pcm                       | 1.5                                          | 2.1 | 2.4 | 2.4  | 56                                      |
| irt rsm                       | 1.5                                          | 2.2 | 2.5 | 2.4  | 53                                      |
| istdize                       | 1.0                                          | 1.0 | 1.0 | 1.0  | 3                                       |
| ivpoisson cfunction           | 2.6                                          | 4.5 | 6.3 | 8.3  | 91                                      |
| ivpoisson gmm, additive       | 2.5                                          | 5.1 | 7.0 | 8.7  | 91                                      |
| ivpoisson gmm, multiplicative | 1.8                                          | 2.9 | 4.1 | 4.8  | 84                                      |
| ivprobit                      | 1.7                                          | 2.5 | 3.3 | 4.1  | 79                                      |
| ivregress 2sls                | 2.0                                          | 3.8 | 6.7 | 11.3 | 97                                      |
| ivregress gmm                 | 1.9                                          | 3.3 | 5.1 | 6.9  | 91                                      |
| ivregress liml                | 1.9                                          | 3.5 | 5.9 | 9.0  | 95                                      |
| ivtobit                       | 1.7                                          | 2.3 | 2.9 | 3.2  | 73                                      |
| kap                           | 1.1                                          | 1.3 | 1.3 | 1.3  | 25                                      |
| kappa                         | 1.9                                          | 3.6 | 6.1 | 9.3  | 94                                      |
| kdensity                      | 1.9                                          | 3.3 | 3.9 | 4.1  | 79                                      |
| keep if <i>exp</i>            | 1.0                                          | 1.1 | 1.2 | 1.1  | 6                                       |
| keep in <i>range</i>          | 0.9                                          | 0.9 | 0.9 | 0.9  | 0                                       |
| keep <i>varlist</i>           | 0.9                                          | 1.0 | 1.0 | 0.9  | 0                                       |
| ksmirnov                      | 2.1                                          | 2.4 | 2.8 | 3.0  | 70                                      |
| ksmirnov, by()                | 2.4                                          | 2.4 | 2.6 | 2.9  | 67                                      |
| ktau                          | 1.1                                          | 1.1 | 1.1 | 1.1  | 5                                       |
| kwallis                       | 2.0                                          | 2.0 | 2.1 | 2.1  | 54                                      |
| ladder                        | 1.6                                          | 2.7 | 3.6 | 4.5  | 83                                      |

All values are expressed as the speed relative to the speed of a single core.

a. Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

b. Bigger is better; 100 is perfect.



Table 1. Stata/MP performance, command by command

| Command                               | Speed relative to a single core <sup>a</sup> |     |     |      | Percentage<br>parallelized <sup>b</sup> |
|---------------------------------------|----------------------------------------------|-----|-----|------|-----------------------------------------|
|                                       | Number of cores                              |     |     |      |                                         |
|                                       | 2                                            | 4   | 8   | 16   |                                         |
| <a href="#">lasso_linear</a>          | 1.4                                          | 1.8 | 2.0 | 2.1  | 54                                      |
| <a href="#">lasso_logit</a>           | 1.3                                          | 1.5 | 1.7 | 1.7  | 39                                      |
| <a href="#">lasso_poisson</a>         | 1.3                                          | 1.5 | 1.6 | 1.7  | 39                                      |
| <a href="#">gsem, lclass(C 2)</a>     | 1.6                                          | 2.5 | 3.2 | 3.7  | 77                                      |
| <a href="#">gsem, lclass(C 3)</a>     | 1.6                                          | 2.3 | 3.0 | 3.3  | 74                                      |
| <a href="#">levelsof</a>              | 1.1                                          | 1.1 | 1.1 | 1.1  | 12                                      |
| <a href="#">loadingplot</a>           | 1.0                                          | 1.0 | 1.0 | 0.9  | 0                                       |
| <a href="#">logistic</a>              | 1.9                                          | 3.8 | 6.8 | 11.6 | 98                                      |
| <a href="#">logit</a>                 | 2.0                                          | 3.8 | 7.0 | 12.1 | 98                                      |
| <a href="#">loneway</a>               | 1.0                                          | 1.1 | 1.3 | 1.4  | 31                                      |
| <a href="#">lowess</a>                | 1.9                                          | 3.4 | 6.8 | 13.6 | 100                                     |
| <a href="#">lpoly</a>                 | 1.6                                          | 2.6 | 3.5 | 4.3  | 82                                      |
| <a href="#">ltable</a>                | 0.8                                          | 0.8 | 0.9 | 0.8  | 0                                       |
| <a href="#">manova (one-way)</a>      | 1.8                                          | 2.8 | 3.8 | 4.7  | 83                                      |
| <a href="#">manova (two-way)</a>      | 1.4                                          | 2.1 | 2.9 | 3.9  | 83                                      |
| <a href="#">margins</a>               | 1.9                                          | 3.4 | 6.5 | 11.1 | 98                                      |
| <a href="#">margins, dydx() exp()</a> | 1.8                                          | 3.3 | 5.6 | 8.7  | 95                                      |
| <a href="#">margins, dydx()</a>       | 1.7                                          | 3.2 | 5.3 | 8.1  | 94                                      |
| <a href="#">margins, exp()</a>        | 1.8                                          | 3.3 | 5.7 | 9.2  | 96                                      |
| <a href="#">markout</a>               | 2.1                                          | 4.1 | 8.3 | 16.5 | 100                                     |
| <a href="#">marksample</a>            | 2.0                                          | 4.0 | 8.1 | 16.1 | 100                                     |
| <a href="#">marksample if exp</a>     | 2.0                                          | 4.0 | 8.0 | 16.0 | 100                                     |
| <a href="#">matrix accum</a>          | 1.9                                          | 3.8 | 7.6 | 15.2 | 100                                     |
| <a href="#">matrix eigenvalues</a>    | 1.0                                          | 1.0 | 1.0 | 1.0  | 0                                       |
| <a href="#">matrix score</a>          | 2.1                                          | 4.2 | 8.3 | 16.4 | 100                                     |

All values are expressed as the speed relative to the speed of a single core.

*a.* Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

*b.* Bigger is better; 100 is perfect.

Table 1. Stata/MP performance, command by command

| Command                        | Speed relative to a single core <sup>a</sup> |     |     |      | Percentage<br>parallelized <sup>b</sup> |
|--------------------------------|----------------------------------------------|-----|-----|------|-----------------------------------------|
|                                | Number of cores                              |     |     |      |                                         |
|                                | 2                                            | 4   | 8   | 16   |                                         |
| matrix svd                     | 1.0                                          | 1.0 | 1.0 | 1.0  | 0                                       |
| matrix syeigen                 | 1.0                                          | 1.0 | 1.0 | 1.0  | 0                                       |
| matrix syminv                  | 1.2                                          | 2.0 | 3.5 | 6.2  | 96                                      |
| mca                            | 1.0                                          | 1.1 | 1.1 | 1.1  | 8                                       |
| mcc                            | 1.1                                          | 1.2 | 1.2 | 1.3  | 22                                      |
| mds                            | 2.3                                          | 2.7 | 3.0 | 3.1  | 70                                      |
| mdslong                        | 2.3                                          | 2.7 | 3.0 | 3.2  | 70                                      |
| mean                           | 2.0                                          | 3.9 | 7.0 | 11.7 | 97                                      |
| mecloglog                      | 1.6                                          | 2.5 | 3.2 | 3.6  | 74                                      |
| median                         | 1.8                                          | 2.8 | 3.8 | 4.5  | 81                                      |
| meintreg                       | 1.8                                          | 3.0 | 4.5 | 5.9  | 88                                      |
| melogit                        | 1.7                                          | 2.6 | 3.5 | 4.0  | 78                                      |
| menbreg, dispersion(constant)  | 1.4                                          | 1.6 | 1.9 | 1.9  | 49                                      |
| menbreg, dispersion(mean)      | 1.5                                          | 1.8 | 2.2 | 2.4  | 63                                      |
| menl                           | 1.0                                          | 1.0 | 1.0 | 0.9  | 0                                       |
| meologit                       | 1.7                                          | 2.6 | 3.5 | 4.1  | 79                                      |
| meoprobit                      | 1.8                                          | 2.8 | 4.0 | 4.9  | 84                                      |
| mepoisson                      | 1.6                                          | 2.3 | 2.8 | 3.1  | 69                                      |
| meprobit                       | 1.8                                          | 2.8 | 3.9 | 4.7  | 83                                      |
| mestreg, distribution(exp)     | 1.7                                          | 2.5 | 3.1 | 3.5  | 74                                      |
| mestreg, distribution(weibull) | 1.7                                          | 2.7 | 3.7 | 4.4  | 81                                      |
| metobit                        | 1.8                                          | 2.7 | 3.8 | 4.5  | 82                                      |
| mgarch                         | 1.1                                          | 1.1 | 1.1 | 1.0  | 0                                       |
| mhodds                         | 1.1                                          | 1.4 | 1.5 | 1.6  | 40                                      |
| mhodds (adjusted)              | 2.1                                          | 3.0 | 3.8 | 4.3  | 80                                      |

All values are expressed as the speed relative to the speed of a single core.

*a.* Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

*b.* Bigger is better; 100 is perfect.

Table 1. Stata/MP performance, command by command

| Command                         | Speed relative to a single core <sup>a</sup> |     |     |      | Percentage<br>parallelized <sup>b</sup> |
|---------------------------------|----------------------------------------------|-----|-----|------|-----------------------------------------|
|                                 | Number of cores                              |     |     |      |                                         |
|                                 | 2                                            | 4   | 8   | 16   |                                         |
| by: mhodds                      | 1.0                                          | 1.0 | 1.1 | 1.1  | 6                                       |
| mhodds (trend)                  | 1.4                                          | 1.7 | 1.9 | 2.0  | 52                                      |
| mi estimate: logit (flong)      | 1.5                                          | 2.0 | 2.5 | 2.6  | 68                                      |
| mi estimate: logit (flongsep)   | 2.0                                          | 3.5 | 5.4 | 7.5  | 92                                      |
| mi estimate: logit (mlong)      | 1.7                                          | 2.7 | 3.7 | 4.5  | 83                                      |
| mi estimate: logit (wide)       | 1.9                                          | 3.4 | 5.6 | 8.2  | 93                                      |
| mi estimate: mlogit             | 1.9                                          | 3.8 | 6.7 | 10.8 | 97                                      |
| mi estimate: ologit             | 1.9                                          | 3.5 | 6.1 | 9.8  | 96                                      |
| mi estimate: regress (flong)    | 1.1                                          | 1.2 | 1.4 | 1.4  | 34                                      |
| mi estimate: regress (flongsep) | 1.7                                          | 2.7 | 3.6 | 4.6  | 83                                      |
| mi estimate: regress (mlong)    | 1.4                                          | 1.8 | 2.0 | 2.0  | 55                                      |
| mi estimate: regress (wide)     | 1.7                                          | 2.7 | 4.1 | 5.2  | 86                                      |
| mi impute chained (flong)       | 1.2                                          | 1.5 | 1.6 | 1.7  | 46                                      |
| mi impute chained (flongsep)    | 1.2                                          | 1.4 | 1.7 | 1.8  | 50                                      |
| mi impute chained (mlong)       | 1.2                                          | 1.4 | 1.6 | 1.8  | 49                                      |
| mi impute chained (wide)        | 1.3                                          | 1.6 | 1.9 | 2.0  | 53                                      |
| mi impute logit (flong)         | 1.2                                          | 1.2 | 1.3 | 1.4  | 28                                      |
| mi impute logit (flongsep)      | 1.6                                          | 2.0 | 2.4 | 2.7  | 66                                      |
| mi impute logit (mlong)         | 1.4                                          | 1.6 | 1.7 | 1.9  | 45                                      |
| mi impute logit (wide)          | 1.8                                          | 2.8 | 4.0 | 4.9  | 83                                      |
| mi impute mlogit                | 1.7                                          | 2.4 | 2.9 | 3.2  | 73                                      |
| mi impute mono pmm              | 1.5                                          | 2.3 | 2.8 | 2.8  | 65                                      |
| mi impute mono regress          | 1.5                                          | 2.2 | 2.8 | 3.3  | 75                                      |
| mi impute mvn                   | 1.4                                          | 1.5 | 1.6 | 1.6  | 36                                      |
| mi impute ologit                | 1.6                                          | 2.0 | 2.3 | 2.4  | 62                                      |

All values are expressed as the speed relative to the speed of a single core.

a. Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

b. Bigger is better; 100 is perfect.

Table 1. Stata/MP performance, command by command

| Command                     | Speed relative to a single core <sup>a</sup> |     |      |      | Percentage<br>parallelized <sup>b</sup> |
|-----------------------------|----------------------------------------------|-----|------|------|-----------------------------------------|
|                             | Number of cores                              |     |      |      |                                         |
|                             | 2                                            | 4   | 8    | 16   |                                         |
| mi impute pmm               | 1.6                                          | 2.2 | 2.9  | 3.3  | 73                                      |
| mi impute regress           | 1.4                                          | 1.7 | 1.8  | 1.9  | 49                                      |
| misstable nested            | 1.3                                          | 1.8 | 2.3  | 2.5  | 66                                      |
| misstable patterns          | 1.3                                          | 1.7 | 1.9  | 2.1  | 57                                      |
| misstable summarize         | 1.0                                          | 1.0 | 1.0  | 1.0  | 0                                       |
| misstable tree              | 1.4                                          | 1.9 | 2.3  | 2.5  | 65                                      |
| mixed                       | 1.1                                          | 1.1 | 1.1  | 1.1  | 5                                       |
| mixed (crossed effects)     | 1.1                                          | 1.1 | 1.1  | 1.1  | 10                                      |
| mkspline                    | 2.0                                          | 2.9 | 3.8  | 4.3  | 80                                      |
| mlevel                      | 2.1                                          | 4.2 | 8.3  | 16.4 | 100                                     |
| mlevel, nocons              | 2.1                                          | 4.2 | 8.3  | 16.4 | 100                                     |
| mlmatbysum                  | 2.0                                          | 3.7 | 6.9  | 12.4 | 98                                      |
| mlmatsum                    | 2.0                                          | 4.0 | 8.0  | 15.8 | 100                                     |
| mlogit                      | 2.0                                          | 4.0 | 7.8  | 15.1 | 99                                      |
| mlsum                       | 1.8                                          | 3.3 | 5.4  | 8.3  | 94                                      |
| mlvecsum                    | 2.0                                          | 3.9 | 7.6  | 14.4 | 99                                      |
| mprobit                     | 1.6                                          | 1.6 | 1.6  | 1.6  | 37                                      |
| mswitch ar                  | 1.2                                          | 1.3 | 1.4  | 1.4  | 31                                      |
| mswitch dr                  | 1.0                                          | 1.0 | 1.0  | 1.0  | 0                                       |
| mvdecode                    | 4.5                                          | 9.5 | 18.9 | 37.7 | 100                                     |
| mvencode                    | 2.0                                          | 3.9 | 7.8  | 15.5 | 100                                     |
| mvreg                       | 1.8                                          | 3.5 | 6.2  | 10.2 | 97                                      |
| mvtest correlations         | 1.7                                          | 2.7 | 3.9  | 4.7  | 84                                      |
| mvtest covariances          | 1.7                                          | 2.7 | 3.7  | 5.0  | 85                                      |
| mvtest means, heterogeneous | 1.6                                          | 2.2 | 2.6  | 2.9  | 68                                      |

All values are expressed as the speed relative to the speed of a single core.

*a.* Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

*b.* Bigger is better; 100 is perfect.

Table 1. Stata/MP performance, command by command

| Command                   | Speed relative to a single core <sup>a</sup> |     |     |      | Percentage<br>parallelized <sup>b</sup> |
|---------------------------|----------------------------------------------|-----|-----|------|-----------------------------------------|
|                           | Number of cores                              |     |     |      |                                         |
|                           | 2                                            | 4   | 8   | 16   |                                         |
| mvtest means, homogeneous | 1.7                                          | 2.7 | 3.3 | 3.8  | 78                                      |
| mvtest means, lr          | 1.6                                          | 2.2 | 2.6 | 2.9  | 69                                      |
| mvtest normality          | 1.0                                          | 1.0 | 1.0 | 1.0  | 0                                       |
| nbreg                     | 1.9                                          | 3.7 | 6.5 | 10.6 | 96                                      |
| newey                     | 1.2                                          | 1.4 | 1.5 | 1.5  | 36                                      |
| nl                        | 2.2                                          | 4.3 | 6.8 | 6.8  | 93                                      |
| nlogit                    | 2.0                                          | 2.5 | 3.1 | 3.3  | 72                                      |
| nlstur                    | 1.3                                          | 1.6 | 1.7 | 1.8  | 46                                      |
| npregress kernel          | 1.1                                          | 1.2 | 1.3 | 1.2  | 21                                      |
| nptrend                   | 2.3                                          | 2.5 | 2.7 | 2.8  | 65                                      |
| ologit                    | 2.0                                          | 3.9 | 7.4 | 13.5 | 99                                      |
| ologit, vce(cluster)      | 1.9                                          | 3.7 | 6.7 | 10.6 | 97                                      |
| ologit, vce(robust)       | 2.0                                          | 3.9 | 7.5 | 13.7 | 99                                      |
| oneway                    | 1.0                                          | 1.0 | 1.0 | 1.0  | 0                                       |
| oprobit                   | 2.0                                          | 3.9 | 7.4 | 13.5 | 99                                      |
| oprobit, vce(cluster)     | 2.0                                          | 3.9 | 7.1 | 12.3 | 98                                      |
| oprobit, vce(robust)      | 2.0                                          | 4.0 | 7.7 | 14.4 | 99                                      |
| orthog                    | 1.7                                          | 4.1 | 7.0 | 10.1 | 93                                      |
| pca                       | 1.5                                          | 2.1 | 2.3 | 2.6  | 68                                      |
| pcorr                     | 2.0                                          | 3.9 | 7.7 | 14.9 | 100                                     |
| pctile                    | 1.8                                          | 3.2 | 4.5 | 5.8  | 87                                      |
| pergram                   | 1.0                                          | 1.0 | 1.0 | 1.0  | 0                                       |
| pkcollapse                | 0.9                                          | 0.9 | 1.0 | 1.0  | 0                                       |
| pkexamine                 | 1.1                                          | 1.2 | 1.3 | 1.3  | 28                                      |
| pksumm                    | 0.9                                          | 1.0 | 1.0 | 1.0  | 0                                       |

All values are expressed as the speed relative to the speed of a single core.

a. Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

b. Bigger is better; 100 is perfect.

Table 1. Stata/MP performance, command by command

| Command               | Speed relative to a single core <sup>a</sup> |     |     |      | Percentage<br>parallelized <sup>b</sup> |
|-----------------------|----------------------------------------------|-----|-----|------|-----------------------------------------|
|                       | Number of cores                              |     |     |      |                                         |
|                       | 2                                            | 4   | 8   | 16   |                                         |
| poisson               | 2.0                                          | 3.9 | 7.4 | 13.1 | 98                                      |
| poisson, vce(cluster) | 2.0                                          | 3.9 | 7.0 | 11.9 | 97                                      |
| poisson, exposure()   | 2.0                                          | 3.9 | 7.3 | 12.9 | 98                                      |
| poisson, vce(robust)  | 2.0                                          | 3.9 | 7.3 | 13.2 | 98                                      |
| pologit               | 1.6                                          | 2.1 | 2.5 | 2.7  | 67                                      |
| popoisson             | 1.5                                          | 2.5 | 2.8 | 3.0  | 71                                      |
| poregress             | 1.5                                          | 2.0 | 1.9 | 2.3  | 61                                      |
| pperron               | 1.2                                          | 1.3 | 1.3 | 1.9  | 50                                      |
| prais                 | 1.1                                          | 1.2 | 1.2 | 1.2  | 21                                      |
| predict, cooks        | 2.0                                          | 4.0 | 8.0 | 16.0 | 100                                     |
| predict, covratio     | 2.0                                          | 4.0 | 7.9 | 15.6 | 100                                     |
| predict, dfbeta       | 2.0                                          | 4.0 | 7.7 | 14.5 | 99                                      |
| predict, dfits        | 2.0                                          | 4.0 | 7.9 | 15.6 | 100                                     |
| predict, e            | 2.0                                          | 3.7 | 6.6 | 11.4 | 97                                      |
| predict, leverage     | 2.0                                          | 4.0 | 8.0 | 15.7 | 100                                     |
| predict, pr           | 2.0                                          | 3.7 | 6.6 | 10.9 | 97                                      |
| predict, residuals    | 2.0                                          | 4.1 | 8.2 | 16.3 | 100                                     |
| predict, rstandard    | 2.0                                          | 4.0 | 8.1 | 16.1 | 100                                     |
| predict, rstudent     | 2.0                                          | 4.0 | 8.0 | 16.0 | 100                                     |
| predict, stdf         | 2.0                                          | 4.0 | 8.0 | 16.0 | 100                                     |
| predict, stdp         | 2.0                                          | 4.0 | 8.0 | 16.1 | 100                                     |
| predict, stdr         | 2.0                                          | 4.0 | 8.0 | 16.0 | 100                                     |
| predict, welsch       | 2.0                                          | 4.0 | 8.0 | 15.9 | 100                                     |
| predict, ystar        | 1.9                                          | 3.5 | 5.7 | 9.0  | 95                                      |
| predictnl             | 2.0                                          | 3.7 | 6.0 | 9.1  | 94                                      |

All values are expressed as the speed relative to the speed of a single core.

*a.* Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

*b.* Bigger is better; 100 is perfect.

Table 1. Stata/MP performance, command by command

| Command                     | Speed relative to a single core <sup>a</sup> |     |      |      | Percentage<br>parallelized <sup>b</sup> |
|-----------------------------|----------------------------------------------|-----|------|------|-----------------------------------------|
|                             | Number of cores                              |     |      |      |                                         |
|                             | 2                                            | 4   | 8    | 16   |                                         |
| probit                      | 2.3                                          | 4.3 | 7.9  | 13.4 | 98                                      |
| procrustes                  | 1.8                                          | 3.6 | 4.9  | 6.8  | 89                                      |
| proportion                  | 1.5                                          | 2.0 | 2.5  | 2.5  | 82                                      |
| prtest1                     | 1.8                                          | 3.0 | 4.6  | 6.1  | 89                                      |
| prtest2                     | 1.8                                          | 2.8 | 4.2  | 5.5  | 87                                      |
| prtest, by()                | 1.1                                          | 1.2 | 1.3  | 1.3  | 25                                      |
| pwcorr                      | 1.9                                          | 3.3 | 6.1  | 9.7  | 95                                      |
| qreg                        | 2.3                                          | 4.0 | 6.0  | 7.8  | 91                                      |
| ranksum                     | 2.2                                          | 3.1 | 3.7  | 4.0  | 79                                      |
| ratio                       | 1.5                                          | 1.9 | 2.2  | 2.4  | 62                                      |
| ratio (exp1) (exp2)         | 1.5                                          | 2.0 | 2.4  | 2.7  | 66                                      |
| recode                      | 1.4                                          | 1.8 | 2.1  | 2.3  | 60                                      |
| reg3                        | 1.8                                          | 3.4 | 5.7  | 8.4  | 94                                      |
| regress                     | 1.9                                          | 3.8 | 7.5  | 14.7 | 100                                     |
| regress, vce(cluster)       | 1.8                                          | 3.1 | 4.8  | 6.7  | 91                                      |
| regress, vce(robust)        | 1.9                                          | 3.7 | 7.0  | 12.8 | 99                                      |
| replace                     | 2.0                                          | 4.0 | 7.9  | 15.9 | 100                                     |
| replace (small expressions) | 4.1                                          | 8.3 | 16.5 | 31.8 | 100                                     |
| reshape long                | 1.1                                          | 2.3 | 2.5  | 2.5  | 62                                      |
| reshape wide                | 1.1                                          | 1.1 | 1.2  | 1.2  | 16                                      |
| robvar                      | 1.2                                          | 1.2 | 1.2  | 2.2  | 58                                      |
| rocfit                      | 1.8                                          | 2.4 | 2.9  | 3.3  | 73                                      |
| roctab                      | 1.0                                          | 1.3 | 1.4  | 1.5  | 36                                      |
| rotate                      | 1.0                                          | 1.0 | 1.0  | 1.0  | 2                                       |
| rotatemat                   | 1.0                                          | 1.0 | 1.0  | 1.0  | 2                                       |

All values are expressed as the speed relative to the speed of a single core.

*a.* Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

*b.* Bigger is better; 100 is perfect.

Table 1. Stata/MP performance, command by command

| Command                 | Speed relative to a single core <sup>a</sup> |     |     |      | Percentage<br>parallelized <sup>b</sup> |
|-------------------------|----------------------------------------------|-----|-----|------|-----------------------------------------|
|                         | Number of cores                              |     |     |      |                                         |
|                         | 2                                            | 4   | 8   | 16   |                                         |
| rreg                    | 2.0                                          | 3.8 | 6.9 | 11.8 | 97                                      |
| runtest                 | 1.8                                          | 2.9 | 3.9 | 4.8  | 84                                      |
| scobit                  | 2.0                                          | 3.8 | 7.1 | 12.4 | 98                                      |
| scoreplot               | 1.6                                          | 2.2 | 2.7 | 3.0  | 70                                      |
| screepplot              | 1.1                                          | 1.3 | 1.3 | 1.3  | 21                                      |
| sdtest1                 | 1.4                                          | 1.8 | 2.1 | 2.3  | 60                                      |
| sdtest2                 | 1.4                                          | 1.7 | 2.0 | 2.2  | 58                                      |
| sdtest, by()            | 1.4                                          | 1.8 | 2.1 | 2.2  | 59                                      |
| sem, method(adf) (CFA)  | 2.0                                          | 4.0 | 7.8 | 15.3 | 100                                     |
| sem, method(ml) (CFA)   | 1.6                                          | 2.1 | 2.8 | 3.0  | 72                                      |
| sem, method(mlmv) (CFA) | 1.2                                          | 1.2 | 1.2 | 1.2  | 17                                      |
| sem (SEM latent)        | 1.6                                          | 2.3 | 2.9 | 3.3  | 74                                      |
| sem (SEM observed)      | 1.6                                          | 2.2 | 2.8 | 3.2  | 71                                      |
| separate                | 1.2                                          | 1.5 | 1.5 | 1.7  | 41                                      |
| sfrancia                | 1.9                                          | 2.2 | 3.9 | 4.2  | 80                                      |
| signrank                | 2.3                                          | 3.0 | 3.6 | 4.0  | 78                                      |
| signtest                | 1.8                                          | 3.9 | 7.4 | 13.7 | 99                                      |
| sktest                  | 2.0                                          | 3.5 | 4.7 | 5.8  | 87                                      |
| slogit                  | 1.3                                          | 1.6 | 1.9 | 2.0  | 54                                      |
| sort                    | 1.0                                          | 1.5 | 2.4 | 3.0  | 75                                      |
| spearman                | 3.0                                          | 3.9 | 4.8 | 5.5  | 84                                      |
| sspace                  | 1.7                                          | 2.3 | 3.0 | 3.2  | 68                                      |
| stack                   | 1.6                                          | 2.3 | 2.7 | 3.0  | 70                                      |
| stci                    | 1.1                                          | 1.2 | 2.4 | 2.4  | 62                                      |
| stcox                   | 1.0                                          | 1.1 | 1.0 | 1.1  | 8                                       |

All values are expressed as the speed relative to the speed of a single core.

*a.* Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

*b.* Bigger is better; 100 is perfect.



Table 1. Stata/MP performance, command by command

| Command                              | Speed relative to a single core <sup>a</sup> |     |     |      | Percentage<br>parallelized <sup>b</sup> |
|--------------------------------------|----------------------------------------------|-----|-----|------|-----------------------------------------|
|                                      | Number of cores                              |     |     |      |                                         |
|                                      | 2                                            | 4   | 8   | 16   |                                         |
| stcrreg                              | 1.1                                          | 1.1 | 1.1 | 1.1  | 12                                      |
| stgen                                | 1.7                                          | 2.0 | 3.4 | 4.0  | 80                                      |
| stintreg, d(exponential)             | 1.8                                          | 2.8 | 3.9 | 4.7  | 84                                      |
| stintreg, d(weibull)                 | 2.6                                          | 4.4 | 6.1 | 8.2  | 91                                      |
| stir                                 | 1.6                                          | 2.1 | 2.4 | 2.6  | 65                                      |
| stmcc                                | 2.7                                          | 2.9 | 3.1 | 3.2  | 70                                      |
| by: stmcc                            | 2.5                                          | 2.8 | 3.1 | 3.2  | 70                                      |
| stmh                                 | 1.3                                          | 1.7 | 1.9 | 2.1  | 53                                      |
| by: stmh                             | 1.2                                          | 1.5 | 1.6 | 1.7  | 47                                      |
| stptime                              | 1.1                                          | 1.3 | 1.3 | 1.4  | 28                                      |
| strate                               | 1.2                                          | 1.3 | 1.5 | 1.6  | 38                                      |
| streg, distribution(exponential)     | 2.0                                          | 3.8 | 6.7 | 11.4 | 97                                      |
| streg, dist(exp) vce(cluster)        | 2.0                                          | 3.9 | 7.3 | 12.9 | 98                                      |
| streg, dist(exp) frailty()           | 2.0                                          | 3.9 | 7.0 | 11.8 | 97                                      |
| streg, dist(exp) frailty() shared()  | 2.1                                          | 3.9 | 6.7 | 10.4 | 96                                      |
| streg, dist(exp) vce(robust)         | 2.0                                          | 3.9 | 7.5 | 13.8 | 99                                      |
| streg, distribution(ggamma)          | 2.2                                          | 4.4 | 7.9 | 13.0 | 97                                      |
| streg, dist(ggamma) vce(cluster)     | 2.0                                          | 4.4 | 8.3 | 13.8 | 98                                      |
| streg, dist(ggamma) vce(robust)      | 2.0                                          | 4.5 | 8.3 | 14.3 | 98                                      |
| streg, distribution(gompertz)        | 2.2                                          | 5.0 | 8.8 | 14.0 | 97                                      |
| streg, dist(gompertz) vce(cluster)   | 2.0                                          | 4.3 | 7.6 | 11.8 | 96                                      |
| streg, dist(gompertz) frailty()      | 2.1                                          | 5.0 | 8.6 | 13.6 | 97                                      |
| streg, dist(gomp) frailty() shared() | 2.4                                          | 4.9 | 7.8 | 10.9 | 95                                      |
| streg, dist(gompertz) vce(robust)    | 2.2                                          | 4.3 | 7.6 | 12.1 | 97                                      |
| streg, distribution(llogistic)       | 2.0                                          | 4.2 | 7.4 | 13.0 | 98                                      |

All values are expressed as the speed relative to the speed of a single core.

*a.* Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

*b.* Bigger is better; 100 is perfect.

Table 1. Stata/MP performance, command by command

| Command                               | Speed relative to a single core <sup>a</sup> |     |     |      | Percentage<br>parallelized <sup>b</sup> |
|---------------------------------------|----------------------------------------------|-----|-----|------|-----------------------------------------|
|                                       | Number of cores                              |     |     |      |                                         |
|                                       | 2                                            | 4   | 8   | 16   |                                         |
| streg, dist(llogistic) vce(cluster)   | 2.1                                          | 4.0 | 7.5 | 12.6 | 98                                      |
| streg, dist(llogistic) frailty()      | 2.0                                          | 3.9 | 7.1 | 12.2 | 97                                      |
| streg, dist(llog) frailty() shared()  | 2.2                                          | 4.5 | 8.0 | 12.3 | 97                                      |
| streg, dist(llogistic) vce(robust)    | 2.1                                          | 4.2 | 7.9 | 14.2 | 99                                      |
| streg, distribution(lnormal)          | 2.1                                          | 4.3 | 7.6 | 12.0 | 97                                      |
| streg, dist(lnormal) vce(cluster)     | 2.0                                          | 4.0 | 7.3 | 11.7 | 97                                      |
| streg, dist(lnormal) frailty()        | 2.0                                          | 3.9 | 7.0 | 11.4 | 97                                      |
| streg, dist(lnorm) frailty() shared() | 2.1                                          | 4.1 | 6.7 | 9.6  | 94                                      |
| streg, dist(lnormal) vce(robust)      | 2.1                                          | 4.2 | 7.6 | 13.4 | 98                                      |
| streg, distribution(weibull)          | 2.0                                          | 3.9 | 7.5 | 13.7 | 99                                      |
| streg, dist(weibull) vce(cluster)     | 2.0                                          | 3.8 | 7.3 | 12.8 | 98                                      |
| streg, dist(weibull) frailty()        | 2.1                                          | 4.8 | 8.4 | 13.5 | 97                                      |
| streg, dist(weib) frailty() shared()  | 2.2                                          | 4.1 | 7.1 | 10.9 | 96                                      |
| streg, dist(weibull) vce(robust)      | 2.0                                          | 3.9 | 7.5 | 13.9 | 99                                      |
| sts generate                          | 1.1                                          | 1.1 | 1.1 | 1.1  | 13                                      |
| sts graph                             | 1.1                                          | 1.2 | 1.2 | 1.2  | 16                                      |
| sts list                              | 1.2                                          | 1.2 | 1.2 | 1.2  | 19                                      |
| sts test                              | 1.1                                          | 1.1 | 1.2 | 1.2  | 15                                      |
| stset                                 | 1.5                                          | 2.0 | 2.3 | 2.5  | 63                                      |
| stsplit                               | 1.1                                          | 1.1 | 1.2 | 1.2  | 19                                      |
| stsum                                 | 1.0                                          | 1.1 | 1.1 | 1.6  | 42                                      |
| stteffects ipw (weibull)              | 2.0                                          | 3.5 | 5.3 | 6.9  | 91                                      |
| stteffects ipwra (weibull)            | 1.8                                          | 2.8 | 3.7 | 4.7  | 83                                      |
| stteffects ra (weibull)               | 1.7                                          | 2.5 | 3.3 | 3.9  | 79                                      |
| stteffects wra (weibull)              | 1.8                                          | 2.7 | 3.7 | 4.5  | 82                                      |

All values are expressed as the speed relative to the speed of a single core.

a. Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

b. Bigger is better; 100 is perfect.

Table 1. Stata/MP performance, command by command

| Command                 | Speed relative to a single core <sup>a</sup> |     |     |      | Percentage<br>parallelized <sup>b</sup> |
|-------------------------|----------------------------------------------|-----|-----|------|-----------------------------------------|
|                         | Number of cores                              |     |     |      |                                         |
|                         | 2                                            | 4   | 8   | 16   |                                         |
| stvary                  | 1.0                                          | 1.5 | 2.1 | 2.4  | 64                                      |
| suest                   | 2.0                                          | 3.8 | 7.2 | 12.5 | 98                                      |
| summarize               | 1.7                                          | 3.5 | 7.0 | 13.8 | 100                                     |
| sunflower               | 1.4                                          | 2.0 | 5.4 | 6.6  | 90                                      |
| sureg                   | 1.8                                          | 3.3 | 5.7 | 8.8  | 95                                      |
| svar                    | 1.7                                          | 2.0 | 2.2 | 2.3  | 60                                      |
| svmat                   | 1.1                                          | 1.1 | 1.1 | 1.1  | 5                                       |
| svy brr: logit          | 1.5                                          | 2.1 | 2.6 | 3.0  | 70                                      |
| svy brr: poisson        | 1.6                                          | 2.4 | 3.0 | 3.5  | 76                                      |
| svy brr: regress        | 1.9                                          | 3.6 | 6.0 | 9.4  | 95                                      |
| svy jackknife: logit    | 1.6                                          | 2.6 | 3.5 | 4.3  | 82                                      |
| svy jackknife: poisson  | 1.5                                          | 2.3 | 3.1 | 3.8  | 79                                      |
| svy jackknife: regress  | 1.9                                          | 3.2 | 5.0 | 6.8  | 90                                      |
| svy linearized: logit   | 2.0                                          | 3.6 | 5.9 | 9.1  | 95                                      |
| svy linearized: poisson | 2.0                                          | 3.6 | 5.9 | 8.9  | 95                                      |
| svy linearized: regress | 1.8                                          | 3.1 | 4.7 | 6.4  | 90                                      |
| swilk                   | 2.2                                          | 2.5 | 3.3 | 3.6  | 74                                      |
| symmetry                | 1.1                                          | 1.2 | 1.2 | 1.3  | 22                                      |
| table (one-way)         | 1.0                                          | 1.3 | 1.5 | 1.6  | 47                                      |
| table (two-way)         | 1.1                                          | 1.4 | 1.6 | 1.8  | 49                                      |
| tabodds                 | 1.0                                          | 1.1 | 1.1 | 1.1  | 8                                       |
| tabodds (adjusted)      | 0.9                                          | 1.0 | 1.1 | 1.1  | 12                                      |
| tabstat                 | 1.4                                          | 1.7 | 1.9 | 2.1  | 55                                      |
| tabstat, by()           | 1.2                                          | 1.1 | 3.0 | 3.2  | 72                                      |
| tabulate (one-way)      | 1.0                                          | 1.0 | 1.0 | 1.0  | 0                                       |

All values are expressed as the speed relative to the speed of a single core.

a. Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

b. Bigger is better; 100 is perfect.

Table 1. Stata/MP performance, command by command

| Command                             | Speed relative to a single core <sup>a</sup> |     |     |      | Percentage<br>parallelized <sup>b</sup> |
|-------------------------------------|----------------------------------------------|-----|-----|------|-----------------------------------------|
|                                     | Number of cores                              |     |     |      |                                         |
|                                     | 2                                            | 4   | 8   | 16   |                                         |
| tabulate (two-way)                  | 1.0                                          | 1.0 | 1.0 | 1.0  | 0                                       |
| teffects aipw (linear)              | 2.0                                          | 3.5 | 5.8 | 8.8  | 94                                      |
| teffects aipw (probit)              | 1.9                                          | 3.4 | 5.6 | 8.2  | 93                                      |
| teffects ipw (logit)                | 2.1                                          | 4.1 | 6.8 | 10.4 | 95                                      |
| teffects ipwra (linear)             | 1.9                                          | 3.4 | 5.8 | 8.4  | 94                                      |
| teffects ipwra (probit)             | 1.9                                          | 3.4 | 5.5 | 7.9  | 93                                      |
| teffects nnmatch                    | 2.0                                          | 4.0 | 7.8 | 15.3 | 100                                     |
| teffects psmatch, logit             | 1.1                                          | 1.1 | 1.2 | 1.1  | 13                                      |
| teffects ra (linear)                | 2.1                                          | 3.8 | 6.6 | 10.5 | 96                                      |
| teffects ra (probit)                | 2.0                                          | 3.6 | 6.0 | 9.1  | 95                                      |
| tetrachoric                         | 1.2                                          | 1.4 | 1.5 | 1.5  | 37                                      |
| threshold, threshvar()              | 1.0                                          | 1.0 | 1.0 | 1.0  | 0                                       |
| threshold, threshvar() regionvars() | 1.0                                          | 1.1 | 1.1 | 1.1  | 5                                       |
| tnbreg                              | 1.5                                          | 2.6 | 3.4 | 3.7  | 79                                      |
| tobit                               | 2.2                                          | 3.8 | 6.1 | 8.7  | 93                                      |
| tostring                            | 1.0                                          | 1.0 | 1.0 | 1.1  | 10                                      |
| total                               | 2.0                                          | 3.9 | 7.1 | 12.0 | 97                                      |
| tpoisson                            | 1.9                                          | 3.5 | 6.2 | 8.8  | 95                                      |
| truncreg                            | 2.0                                          | 3.5 | 5.8 | 8.4  | 94                                      |
| tsfilter bk                         | 1.1                                          | 1.2 | 1.3 | 1.3  | 25                                      |
| tsfilter bw                         | 1.0                                          | 1.0 | 1.0 | 1.0  | 5                                       |
| tsfilter cf                         | 1.1                                          | 1.2 | 1.3 | 1.3  | 26                                      |
| tsfilter hp                         | 1.0                                          | 1.0 | 1.0 | 1.0  | 4                                       |
| tsrevar                             | 2.5                                          | 4.9 | 9.4 | 15.0 | 95                                      |
| tsset                               | 1.0                                          | 1.5 | 1.9 | 2.3  | 62                                      |

All values are expressed as the speed relative to the speed of a single core.

a. Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

b. Bigger is better; 100 is perfect.

Table 1. Stata/MP performance, command by command

| Command             | Speed relative to a single core <sup>a</sup> |     |     |      | Percentage<br>parallelized <sup>b</sup> |
|---------------------|----------------------------------------------|-----|-----|------|-----------------------------------------|
|                     | Number of cores                              |     |     |      |                                         |
|                     | 2                                            | 4   | 8   | 16   |                                         |
| tssmooth exp        | 1.3                                          | 1.6 | 1.8 | 1.9  | 50                                      |
| tssmooth ma         | 1.2                                          | 1.3 | 1.3 | 1.4  | 28                                      |
| ttest1              | 1.4                                          | 1.8 | 2.1 | 2.3  | 61                                      |
| ttest2              | 1.7                                          | 2.4 | 3.0 | 3.4  | 75                                      |
| ttest, by()         | 1.4                                          | 1.8 | 2.1 | 2.2  | 59                                      |
| twoway fpfit        | 1.5                                          | 2.3 | 4.4 | 5.1  | 86                                      |
| twoway lfitci       | 1.2                                          | 1.5 | 1.6 | 1.7  | 45                                      |
| twoway mband        | 2.3                                          | 3.3 | 4.1 | 4.5  | 79                                      |
| twoway mspline      | 2.4                                          | 3.0 | 3.8 | 4.3  | 80                                      |
| ucm, model(rwdrift) | 1.3                                          | 1.3 | 1.3 | 1.3  | 20                                      |
| var                 | 1.4                                          | 3.1 | 3.6 | 3.6  | 74                                      |
| vargranger          | 1.0                                          | 1.0 | 1.0 | 1.0  | 1                                       |
| varlmar             | 1.3                                          | 1.7 | 2.1 | 2.3  | 59                                      |
| varnorm             | 1.3                                          | 1.9 | 2.3 | 2.6  | 67                                      |
| varsoc              | 1.4                                          | 1.7 | 2.1 | 2.2  | 58                                      |
| varstable           | 2.4                                          | 2.4 | 2.4 | 2.4  | 58                                      |
| vec                 | 1.2                                          | 1.4 | 1.5 | 1.6  | 39                                      |
| veclmar             | 1.1                                          | 1.3 | 1.5 | 1.5  | 38                                      |
| vecnorm             | 1.3                                          | 1.6 | 1.7 | 2.8  | 69                                      |
| vecrank             | 1.1                                          | 1.4 | 1.5 | 1.6  | 39                                      |
| vecstable           | 1.0                                          | 1.0 | 1.0 | 1.0  | 0                                       |
| vwls                | 2.0                                          | 3.9 | 7.3 | 12.7 | 98                                      |
| wntestb             | 1.0                                          | 1.0 | 1.0 | 1.0  | 1                                       |
| wntestq             | 1.1                                          | 1.1 | 1.1 | 1.2  | 14                                      |
| xcorr               | 1.1                                          | 1.1 | 1.1 | 1.1  | 12                                      |

All values are expressed as the speed relative to the speed of a single core.

*a.* Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

*b.* Bigger is better; 100 is perfect.

Table 1. Stata/MP performance, command by command

| Command                           | Speed relative to a single core <sup>a</sup> |     |     |      | Percentage<br>parallelized <sup>b</sup> |
|-----------------------------------|----------------------------------------------|-----|-----|------|-----------------------------------------|
|                                   | Number of cores                              |     |     |      |                                         |
|                                   | 2                                            | 4   | 8   | 16   |                                         |
| xpologit                          | 1.3                                          | 1.5 | 1.6 | 1.7  | 42                                      |
| xpopoisson                        | 1.2                                          | 1.5 | 1.6 | 1.7  | 44                                      |
| xporegress                        | 1.3                                          | 1.5 | 1.6 | 1.7  | 41                                      |
| xtabond                           | 1.2                                          | 1.5 | 1.6 | 1.8  | 46                                      |
| xtabond, twostep                  | 1.2                                          | 1.5 | 1.7 | 1.8  | 49                                      |
| xtcloglog, re                     | 2.1                                          | 4.0 | 6.9 | 10.9 | 96                                      |
| xtdata, be                        | 1.9                                          | 1.9 | 2.0 | 2.3  | 59                                      |
| xtdata, fe                        | 3.2                                          | 3.7 | 3.9 | 4.2  | 77                                      |
| xtdata, re                        | 3.3                                          | 3.9 | 4.1 | 4.3  | 78                                      |
| xtdpd                             | 1.5                                          | 1.6 | 1.7 | 1.8  | 46                                      |
| xtdpdsys                          | 1.4                                          | 1.6 | 1.7 | 1.9  | 49                                      |
| xteregress                        | 1.4                                          | 1.8 | 2.0 | 1.9  | 46                                      |
| xtfrontier                        | 2.6                                          | 4.4 | 6.9 | 9.5  | 93                                      |
| xtgee, family(gaussian) corr(ar2) | 1.8                                          | 2.0 | 2.1 | 2.2  | 54                                      |
| xtgee, fam(gauss) corr(unstruct)  | 1.7                                          | 1.9 | 2.1 | 2.1  | 54                                      |
| xtcloglog, pa                     | 1.8                                          | 2.6 | 3.3 | 3.8  | 78                                      |
| xtlogit, pa                       | 1.6                                          | 1.9 | 2.1 | 2.2  | 58                                      |
| xtnbreg, pa                       | 1.7                                          | 2.3 | 2.7 | 2.9  | 70                                      |
| xtpoisson, pa                     | 1.6                                          | 2.0 | 2.3 | 2.5  | 62                                      |
| xtprobit, pa                      | 1.4                                          | 1.7 | 1.9 | 2.0  | 53                                      |
| xtreg, pa                         | 1.4                                          | 1.7 | 1.8 | 1.9  | 48                                      |
| xtgls                             | 1.6                                          | 2.2 | 2.7 | 3.0  | 70                                      |
| xthtaylor                         | 1.6                                          | 2.8 | 3.9 | 4.9  | 84                                      |
| xtile                             | 1.0                                          | 1.0 | 1.0 | 1.0  | 0                                       |
| xtintreg                          | 2.0                                          | 3.9 | 7.0 | 11.7 | 97                                      |

All values are expressed as the speed relative to the speed of a single core.

a. Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

b. Bigger is better; 100 is perfect.

Table 1. Stata/MP performance, command by command

| Command                            | Speed relative to a single core <sup>a</sup> |     |      |      | Percentage<br>parallelized <sup>b</sup> |
|------------------------------------|----------------------------------------------|-----|------|------|-----------------------------------------|
|                                    | Number of cores                              |     |      |      |                                         |
|                                    | 2                                            | 4   | 8    | 16   |                                         |
| xtivreg, be                        | 2.8                                          | 4.2 | 5.0  | 5.6  | 84                                      |
| xtivreg, fd                        | 2.4                                          | 3.5 | 4.0  | 4.3  | 79                                      |
| xtivreg, fe                        | 2.5                                          | 3.6 | 4.1  | 4.5  | 80                                      |
| xtivreg, re                        | 3.0                                          | 4.2 | 5.1  | 5.6  | 84                                      |
| xtlogit, fe                        | 1.6                                          | 2.4 | 2.8  | 3.1  | 71                                      |
| xtlogit, re                        | 2.2                                          | 3.7 | 6.0  | 7.6  | 92                                      |
| xtnbreg, fe                        | 3.4                                          | 5.9 | 8.3  | 10.3 | 93                                      |
| xtnbreg, re                        | 3.0                                          | 5.1 | 7.4  | 9.6  | 93                                      |
| xtologit                           | 1.8                                          | 2.7 | 3.7  | 4.4  | 81                                      |
| xtoprobit                          | 1.8                                          | 2.9 | 4.2  | 5.2  | 86                                      |
| xtpcse                             | 1.4                                          | 1.8 | 2.0  | 2.1  | 56                                      |
| xtpoisson, fe                      | 3.0                                          | 5.2 | 7.8  | 9.9  | 93                                      |
| xtpoisson, re                      | 3.0                                          | 5.8 | 10.2 | 15.4 | 97                                      |
| xtprobit, re                       | 2.0                                          | 3.7 | 6.3  | 9.8  | 95                                      |
| xtrc                               | 1.6                                          | 2.3 | 2.8  | 3.1  | 72                                      |
| xtreg, be                          | 2.2                                          | 2.9 | 3.4  | 3.6  | 75                                      |
| xtreg, fe                          | 1.9                                          | 3.4 | 5.5  | 7.9  | 93                                      |
| xtreg, fe vce(robust)              | 1.8                                          | 3.4 | 5.8  | 9.1  | 95                                      |
| xtreg, mle                         | 1.3                                          | 1.8 | 3.1  | 3.3  | 72                                      |
| xtreg, re                          | 2.8                                          | 4.0 | 5.0  | 5.3  | 84                                      |
| xtregar, fe                        | 2.2                                          | 4.9 | 5.4  | 5.9  | 85                                      |
| xtregar, re                        | 2.3                                          | 3.2 | 3.6  | 4.0  | 77                                      |
| xtset                              | 1.2                                          | 1.6 | 2.0  | 2.2  | 60                                      |
| xtstreg, distribution(exponential) | 1.9                                          | 2.8 | 3.6  | 4.0  | 78                                      |
| xtstreg, distribution(weibull)     | 1.9                                          | 2.9 | 4.0  | 4.7  | 83                                      |

All values are expressed as the speed relative to the speed of a single core.

*a.* Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

*b.* Bigger is better; 100 is perfect.

Table 1. Stata/MP performance, command by command

| Command             | Speed relative to a single core <sup>a</sup> |     |     |      | Percentage<br>parallelized <sup>b</sup> |
|---------------------|----------------------------------------------|-----|-----|------|-----------------------------------------|
|                     | Number of cores                              |     |     |      |                                         |
|                     | 2                                            | 4   | 8   | 16   |                                         |
| xtsum               | 2.7                                          | 3.8 | 4.5 | 4.9  | 82                                      |
| xttab               | 1.1                                          | 1.2 | 1.3 | 1.4  | 32                                      |
| xttobit             | 2.0                                          | 3.9 | 6.5 | 9.5  | 94                                      |
| xtunitroot breitung | 1.2                                          | 1.6 | 1.8 | 2.0  | 55                                      |
| xtunitroot fisher   | 1.0                                          | 1.1 | 1.1 | 1.1  | 10                                      |
| xtunitroot hadri    | 1.2                                          | 1.3 | 1.3 | 1.3  | 26                                      |
| xtunitroot ht       | 1.3                                          | 1.6 | 1.8 | 1.9  | 52                                      |
| xtunitroot ips      | 1.0                                          | 1.0 | 1.0 | 1.0  | 1                                       |
| xtunitroot llc      | 1.0                                          | 1.0 | 1.0 | 1.0  | 2                                       |
| zinb                | 2.0                                          | 4.3 | 7.7 | 12.5 | 97                                      |
| zioprobit           | 1.9                                          | 3.4 | 5.6 | 7.8  | 92                                      |
| zip                 | 2.0                                          | 4.1 | 7.5 | 13.5 | 98                                      |
| _predict, xb        | 2.0                                          | 4.1 | 8.2 | 16.3 | 100                                     |
| _rmcoll             | 1.8                                          | 3.7 | 7.4 | 14.1 | 100                                     |
| _robust             | 2.0                                          | 4.0 | 7.7 | 15.0 | 99                                      |

All values are expressed as the speed relative to the speed of a single core.

*a.* Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

*b.* Bigger is better; 100 is perfect.



Nine of the lines in table 1 represent estimation commands run on survey data. Each of these commands begins with **svy**. These are only a few of the estimation commands that support estimation on survey data, but we can make some generalizations about how the three primary methods of estimation with survey data will perform with other estimation commands. With the linearization method, prefix **svy linearized**, estimation commands will be parallelized just as well and sometimes better than they were parallelized on non-survey data. This is true because the linearization computation is itself almost 100% parallelized. When using the balanced repeated replications (BRR) method, **svy brr**, or the jackknife method, **svy jackknife**, almost all estimation commands are slightly less parallelized. The BRR and jackknife VCE computations are not themselves parallelized, but the overall estimation time is dominated by standard estimation.

More than a full page of table 1 is dedicated to performance when using multiple imputation (MI). These commands begin with the **mi** prefix. All the results in the table are from problems with five imputations. The number of imputations does not affect parallelization performance much. As with all commands, problems with more observations and covariates are better parallelized; see appendix D for the sizes of problems used to assess performance.

There are two particularly computationally intensive aspects to using MI data—creating the MI datasets (imputation) and estimation. The table reports the results for all the primary methods of imputation; these lines are prefixed with **mi impute**. It also reports the results for four representative estimators—linear regression (**regress**), logistic regression (**logit**), multinomial logistic regression (**mlogit**), and ordered logistic regression (**ologit**).

Performance on MI data is affected by the style in which the MI data are stored. Stata allows four styles for storing MI data: wide (**wide**), marginal long (**mlong**), full long (**flong**), and full long and separate (**flongsep**). Each style has advantages with regard to storage required and ease of use in some analyses.

With regard to imputation performance under Stata/MP, imputation is fastest and most parallelized when using style **flongsep**. **flongsep** is the native style in which imputations are performed. Table 1 reports performance across all MI storage styles for only the **logit** imputer; the relative performance of the styles is similar for other imputers.

Estimation is fastest and most parallelized when using storage style **wide**, although style **flongsep** is also well parallelized. Style **wide** is fastest because the overhead for managing **wide** data mostly involves simply changing the names of variables. The table reports estimation results in all four MI storage styles for only **regress** and **logit**. The relative performance is similar for other estimators.

As with estimation using survey data, MI can be applied to many more estimation commands than those listed in the table. Only some of the MI computations are themselves parallelized, so most commands are less parallelized when used with MI data, regardless of the style in which the data are stored. Computationally intensive estimation commands that involve iterative solutions, such as logistic regression, are less affected than are commands with closed-form solutions, such as linear regression.

For maximum performance using Stata/MP, set the MI storage style to **flongsep** when performing imputations and to **wide** when performing estimations. The short time you invest to convert between styles will be more than repaid in faster imputation and estimation. If you have insufficient memory to store an MI dataset in style **wide**, then continue to use **flongsep** during estimation.

When using many imputations on moderate-to-small problems, the overhead of the MI computations can dominate the time required. Such problems are less parallelized than reported in the table. Conversely, very large problems with few imputations are parallelized even more than reported in the table.

## 9 Performance variability across computing platforms

As discussed in sections 3 and 4, multicore/multiprocessor performance will vary across computing platforms for many reasons. Those reasons include differences in how operating systems partition tasks, how processors pipeline and partition instructions, how memory is accessed, and how onboard processor cache is handled.

Stata/MP performance has been tested on dozens of different platforms, including different processors (both Intel and AMD), different cache architectures, different operating systems (including Microsoft Windows, Mac OS X for Intel, Linux, and Oracle Solaris), and different architectures for accessing memory. Despite the possibility for varying performance, the results from all these tests support the results presented in section 8 and appendixes A and B.

It is not helpful to break these results down by platform. There were no conclusive patterns among operating systems, CPUs, or other platform characteristics.

## 10 Hyperthreading—single- and multiple-processor platforms

Hyperthreading is an Intel technology for allowing each core of a processor to masquerade as two cores. The operating system and other applications see each physical core as two virtual cores and treat each just as they would any physical core. Intel achieves performance improvements primarily because main computer memory is slow compared with the processor and its onboard cache memory. When the execution thread of one process must wait for something from main memory, the thread for another process can execute. The effect is clearly not the same as having two cores, but for many applications, performance can be improved by treating a computer with a hyperthreaded processor as having twice as many cores as it actually has.

Stata/MP runs on hyperthreaded processors.

Most Stata commands are computationally intense and Stata/MP has been optimized to access main memory efficiently. For these reasons, we would not expect hyperthreading to substantially improve the performance of most commands. Our timings indicate that this is true for most Stata commands, but a few performance gains were surprisingly good.

Figure 13 presents the now familiar boundary region and median performance of Stata/MP running on a quad-core computer with hyperthreading – making for 8 virtual cores. Through the first 4 cores, performance is almost identical to what we saw in Figure 7 for a non-hyperthreaded processor. That is to say, so long as we do not exceed the number of physical cores on a system, hyperthreaded computers behave just like non-hyperthreaded computers.

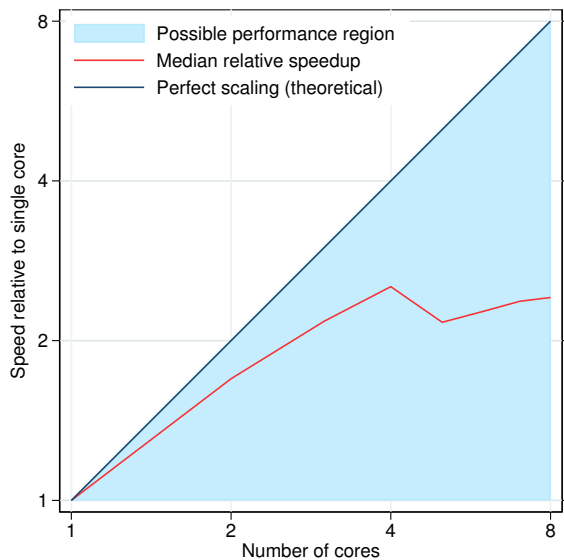


Figure 13. **Performance of Stata/MP on hyperthreaded CPUs.** Speed on multiple cores relative to speed on a single core.

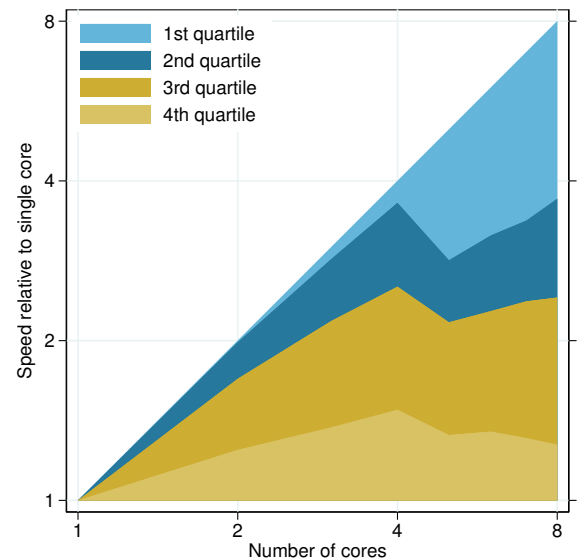


Figure 14. **Quartiles of Stata/MP Performance on hyperthreaded CPUs.** Speed on multiple cores relative speed on a single core.

Above 4 cores, median performance drops for 5 cores, one of them virtual, but improves to approaching the performance of 4 physical cores. The most interesting point beyond 4 cores is 8 cores – all of the virtual cores on the computer. The median relative speed with 8 cores is 2.4, which is slightly less than the median speed of 2.5 for the 4 physical cores.

Figure 14 presents the quartiles of command performance. The diagonal top of the light-blue region indicates that at least one command has perfect parallelization over all 8 virtual cores. Moreover, for the 25% of commands that perform best with hyperthreading their relative speed is at least 3.7 with all 8 virtual cores as compared to 3.6 with 4 physical cores — a 2% improvement. At the other end of the spectrum, for the 25% take the least advantage of hyperthreading, the performance on 8 virtual cores is worse than that of 4 physical cores.

By way of caution, Stata/MP has not been evaluated on a wide range of single-processor hyper-threaded computers, and these results should therefore be considered provisional.

On multiprocessor computers where each CPU is hyperthreaded, the current recommendation is to set Stata/MP to use the number of real CPUs, not the number of virtual processors. Under such architectures, each CPU appears to Stata/MP and the operating system as two processors, and Stata/MP would by default try to use all the (virtual) processors. On these computers, users should type

```
. set processors #
```

where # is the number of CPUs on the computer. Here we use “CPU” to mean a physical core on the computer and not a virtual core created by hyperthreading. So, we could equivalently say, where # is the number of physical cores on the computer.

This can be done either interactively or placed in Stata's `profile.do` startup script.

Current experience indicates that setting the number of processors to be used above the number of real CPUs on the computer leads to contention for the floating-point unit (FPU), which can make commands run slower when trying to use virtual processors.

Figures 15 and 16 show the results of two commands run on an 4-processor computer, each hyper-threaded, giving the appearance of 8 virtual processors.

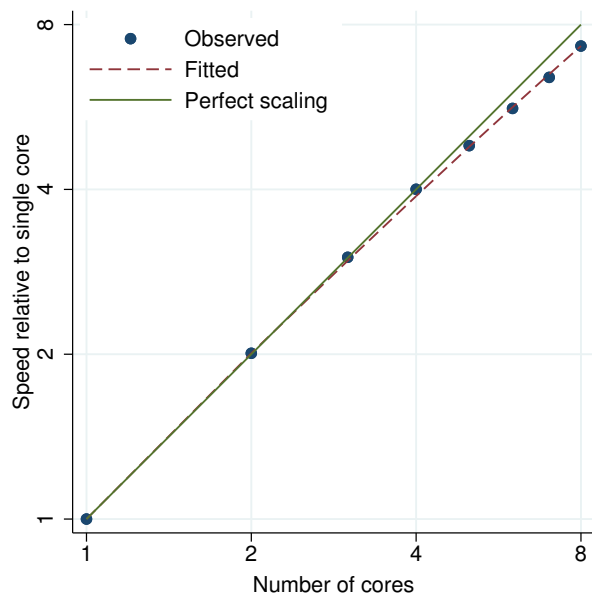


Figure 15. `predict, leverage` performance plot on computers with hyperthreaded CPUs.

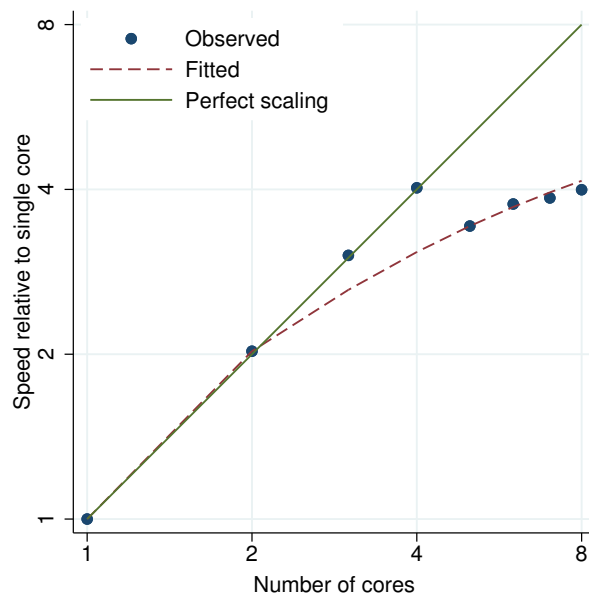


Figure 16. `regress` performance plot on computers with hyperthreaded CPUs.

The `predict, leverage` command, however, is an exception to this recommendation. This command remains nearly perfectly parallelized through all 8 processors (half of which are virtual).

Most commands do not exhibit results like this, and `regress` is an example. Beyond the number of real CPUs, performance actually degrades. This occurs because each CPU has only one FPU, and `regress`, along with most Stata commands, requires many floating-point computations. The computations are dominated by access to the FPU, and the virtual processors must contend for access to this single FPU.

## A Performance assessment graphs for desktop computers

The performance of Stata/MP as reported in columns 2, 3, and 4 of table 1 is presented graphically below, along with the modeled performance from equation 1 and a line representing perfectly scalable performance.

Figures 17 and 18 show two typical graphs. As with table 1, performance is measured as the speed of executing the command on multiple cores relative to the speed on a single core.

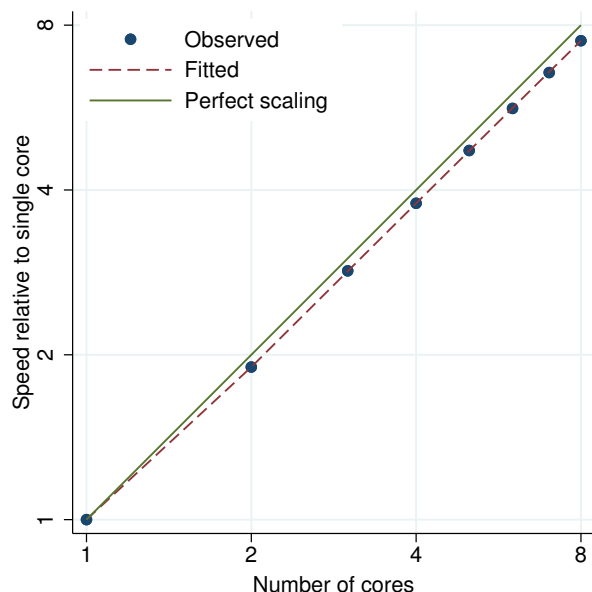


Figure 17. `regress` performance plot.

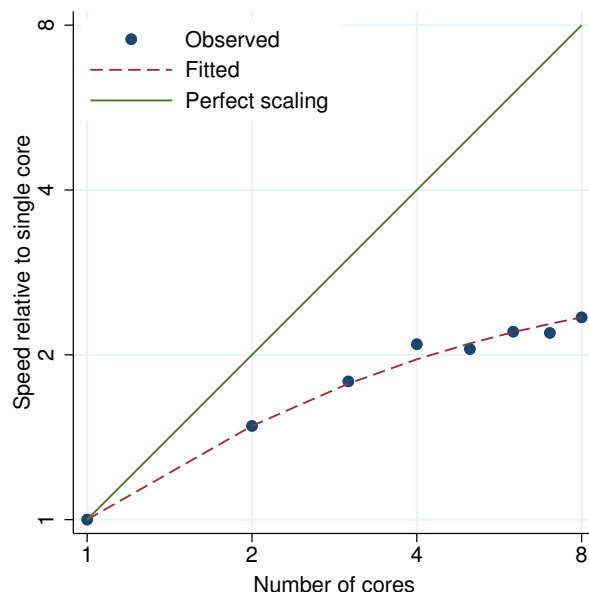


Figure 18. `clogit` (1 to k matching) performance plot.

For a perfectly scalable command, the speed doubles each time the number of cores is doubled. This type of scalability is linear when the number of cores and the relative speed are graphed on a logarithmic scale, which is the scale used in these graphs. Perfect scaling is shown on each graph as a green line that diagonally bisects the graph.

Linear regression, shown in figure 17, is nearly perfectly scalable. Both the observed values and the modeled performance are nearly on the perfect-scalability reference line. The speed is doubled each time the number of cores doubles.

As shown in figure 18, conditional logistic regression clearly performs better as the number of cores increases, but not as much better as linear regression. From table 1, we can see that `clogit` (1 to k matching) is 66% parallelized as compared with 100% for `regress`. From figure 18, we see that `clogit` run with 2 cores on a large dataset is 1.5 times faster than when run with one core; with 4 cores, this relative speed climbs to 2.1; and with 8 cores, it climbs further to 2.3.

Figure 8, from section 7, summarizes the information from all the graphs in this section by placing the observed performance for each command into one of the performance quartiles on the graph.

In a few of the graphs that follow, the observed performance exceeds the theoretical limit of perfect scaling—some of the relative speeds are above the diagonal perfect-scaling line. An example of this can be seen when the `replace` command is evaluating small expressions, as shown in figure 19.

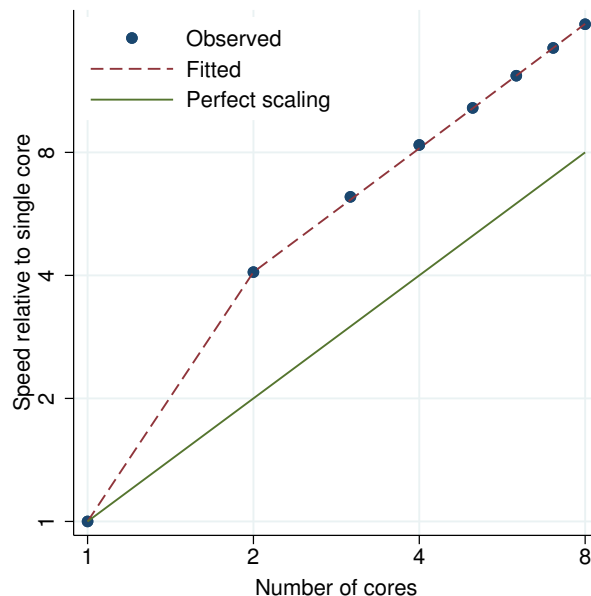
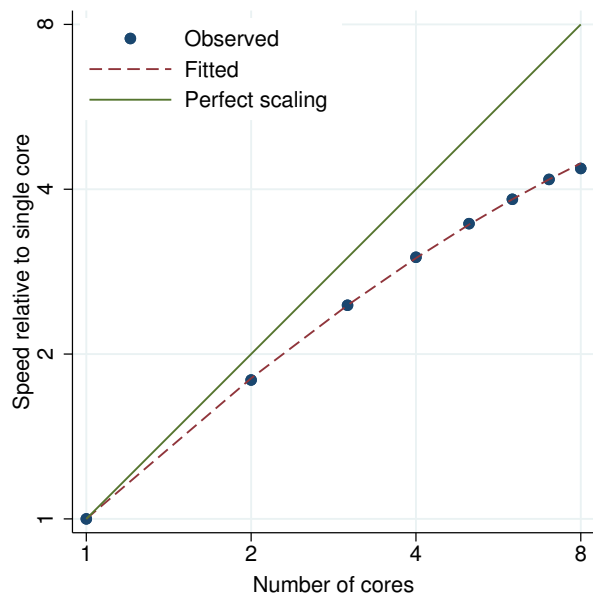
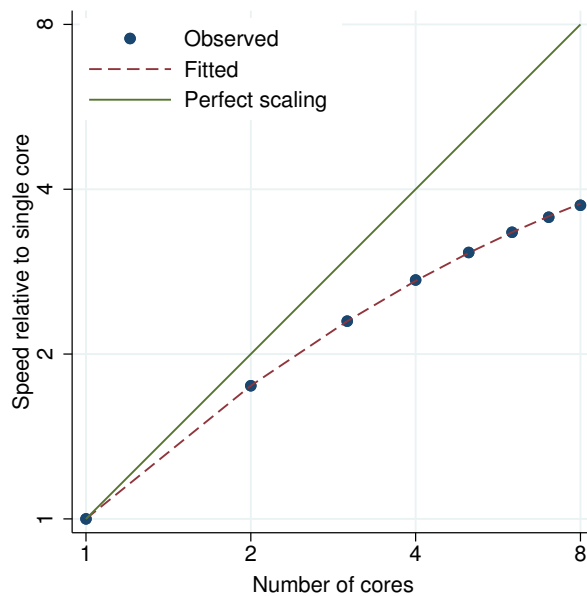
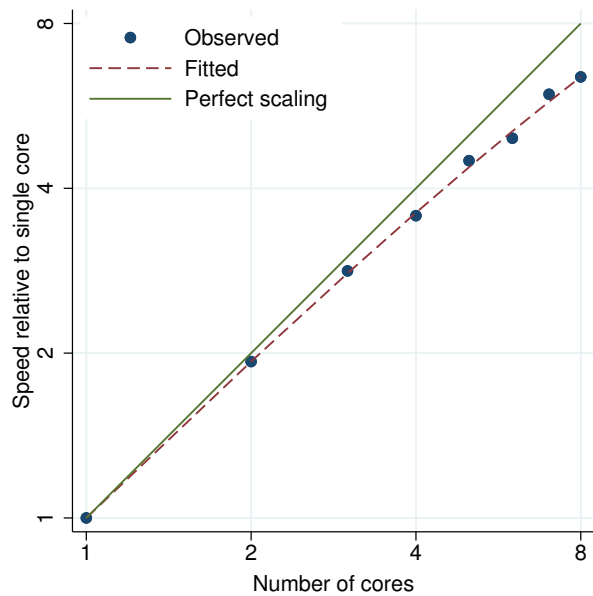
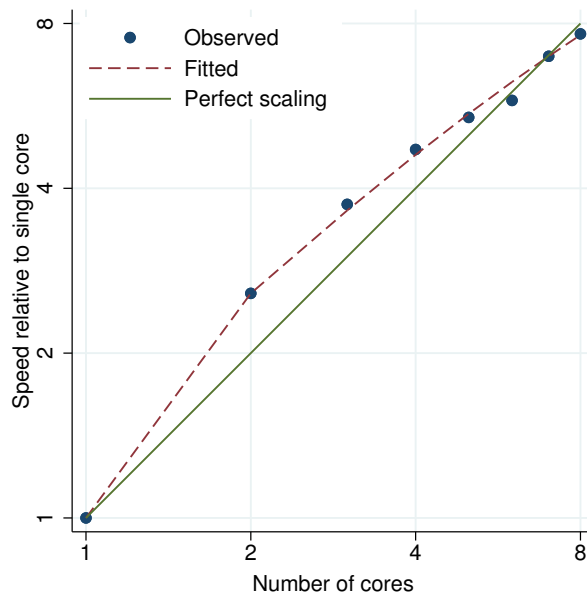


Figure 19. `replace` performance plot.

This phenomenon is nothing more than a cache effect. Cache is very high speed memory that processors use to store data and code that they use often or expect to use often. Cores run much faster when the data they need can be found in cache rather than in standard memory. The `replace` command above was able to find far more of the data it needed in cache when running on 2 or more cores than it could find when running on a single core. The model that we used to determine percentage parallelized ignored that cache effect and correctly determined that the `replace` command was just under 100% parallelized, not greater than 100%.

Observant readers will have noted that the `regress` command in figure 17 exhibited some mild cache effects. Its observed performance is slightly above perfect scaling.

Figure 20. `alpha` performance plot.Figure 21. `ameans` performance plot.Figure 22. `anova` (one-way) performance plot.Figure 23. `anova` (two-way) performance plot.

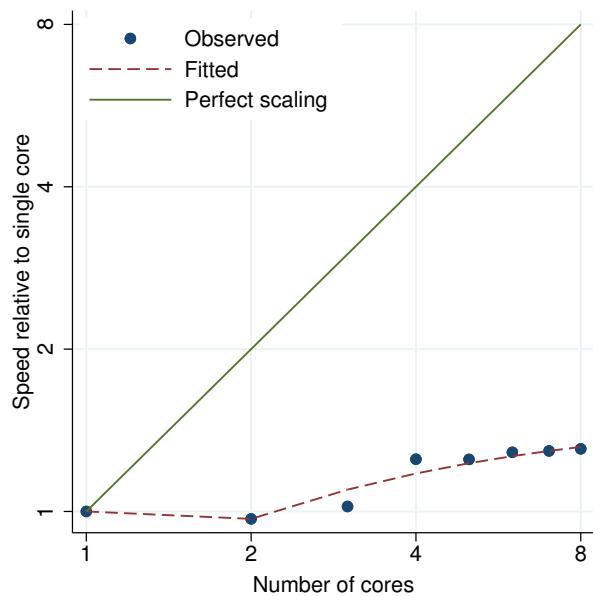


Figure 24. arch performance plot.

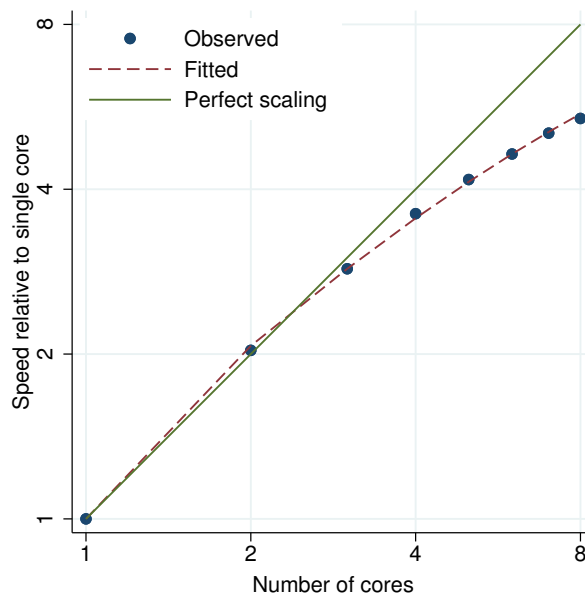


Figure 25. areg performance plot.

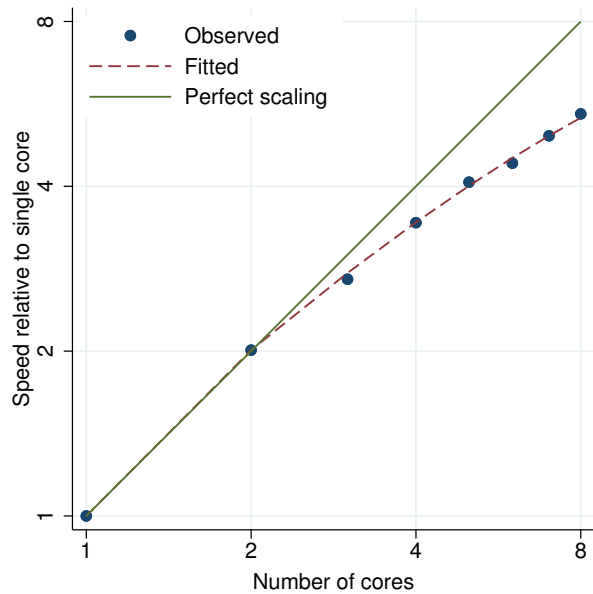


Figure 26. areg, vce(cluster) performance plot.

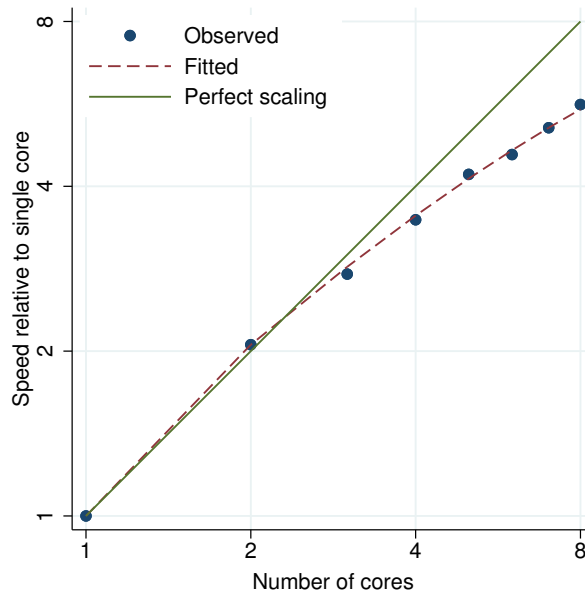
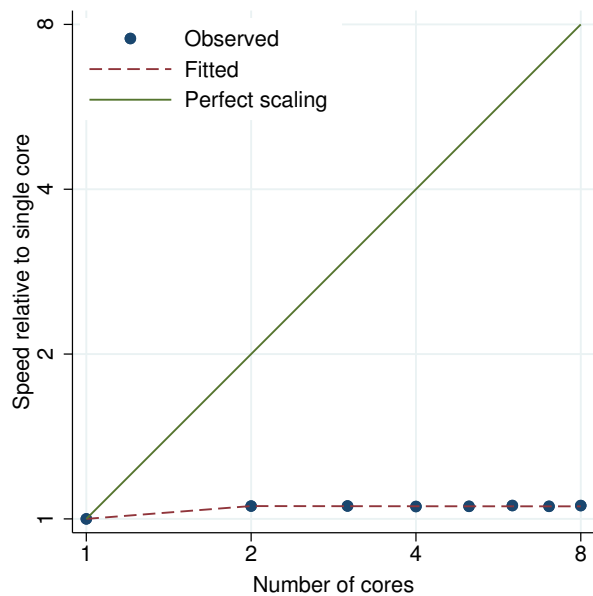
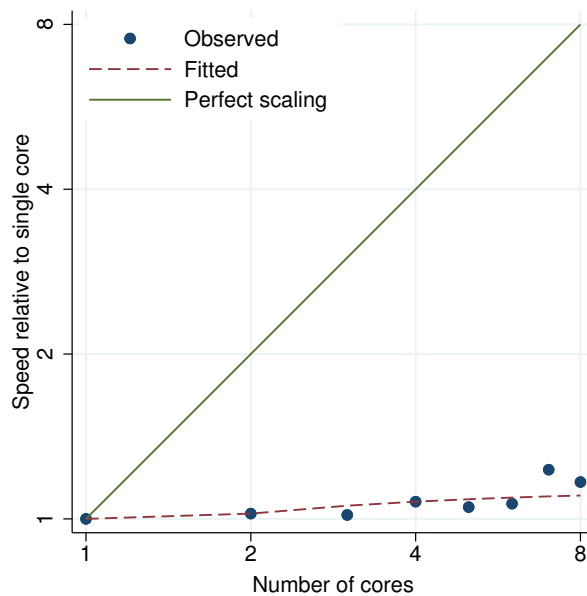
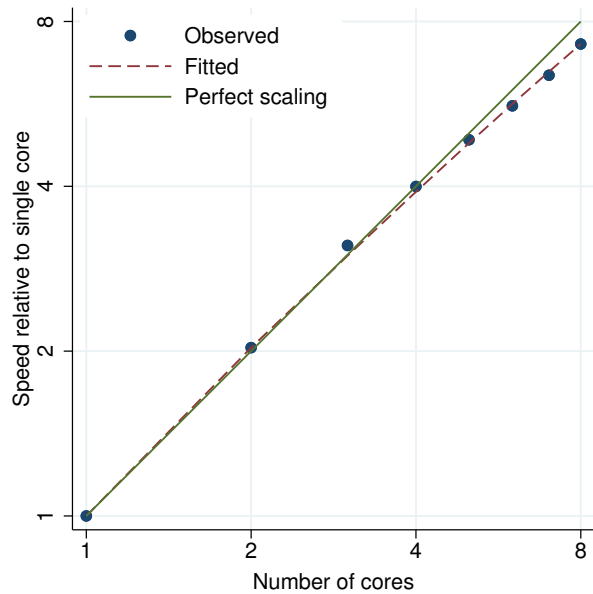
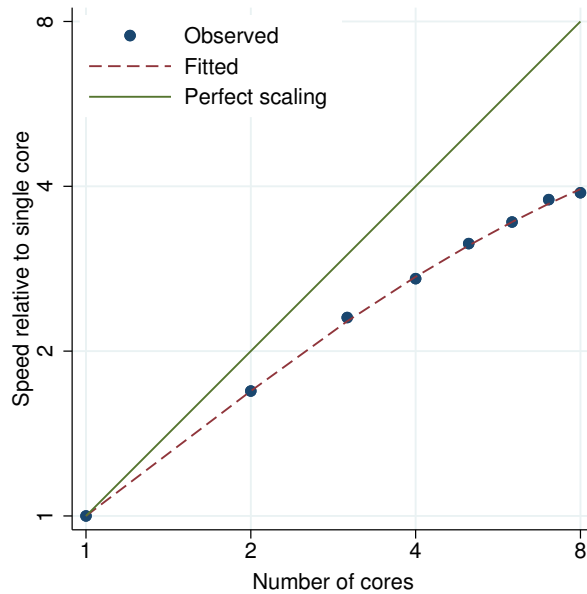


Figure 27. areg, vce(robust) performance plot.



Figure 28. `arfima` performance plot.Figure 29. `arima` performance plot.Figure 30. `bayes: logit` performance plot.Figure 31. `bayes: poisson` performance plot.

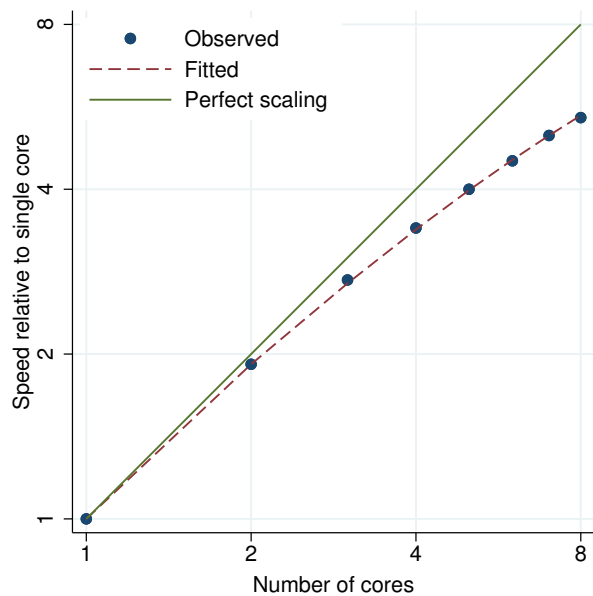


Figure 32. bayes: regress performance plot.

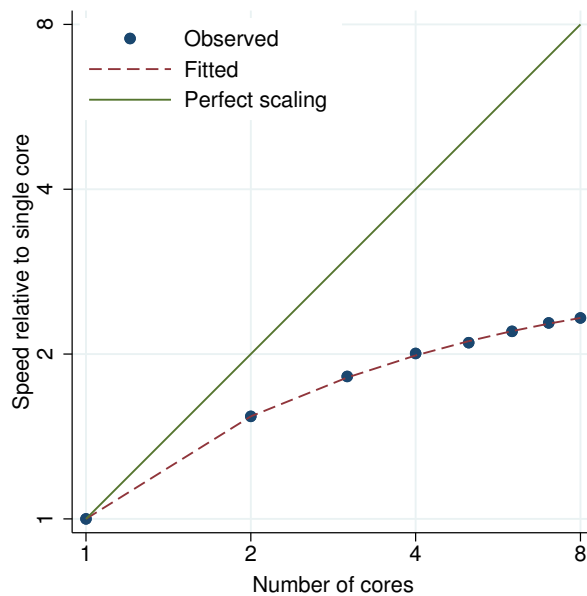


Figure 33. bayesmh logit performance plot.

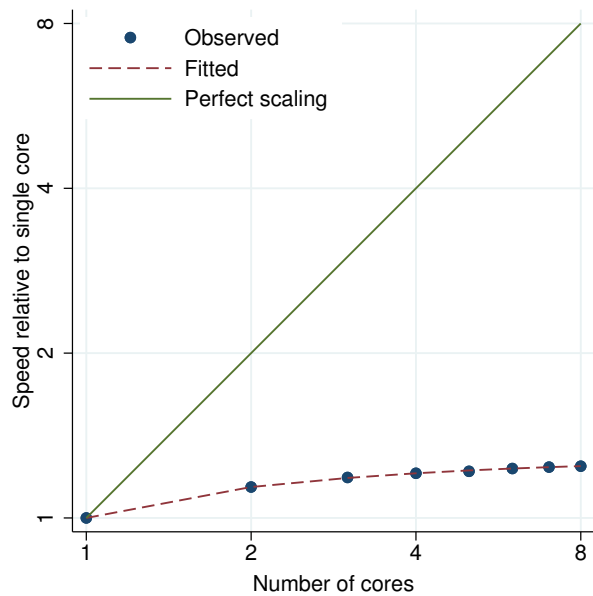


Figure 34. bayesmh mvn performance plot.

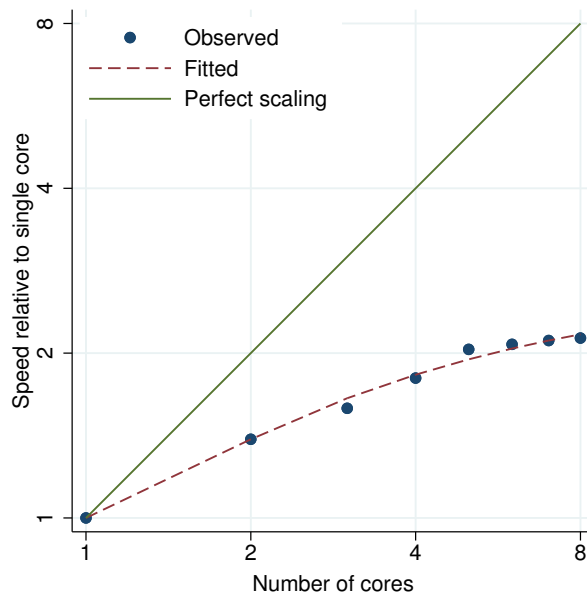


Figure 35. bayesmh mylogit performance plot.

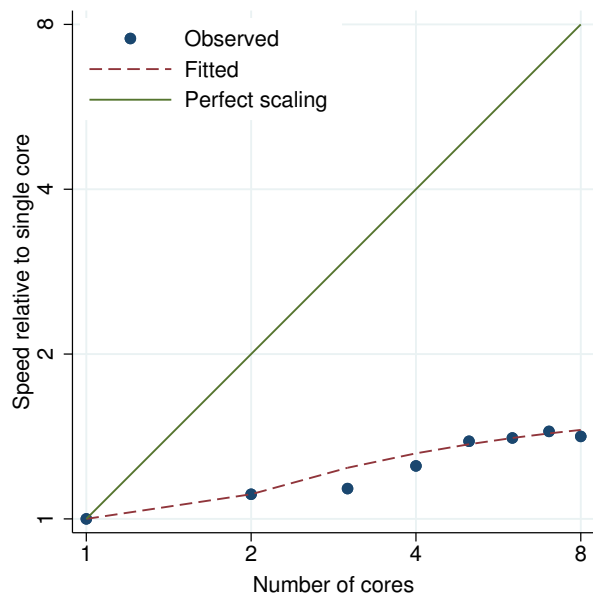


Figure 36. bayesmh nl performance plot.

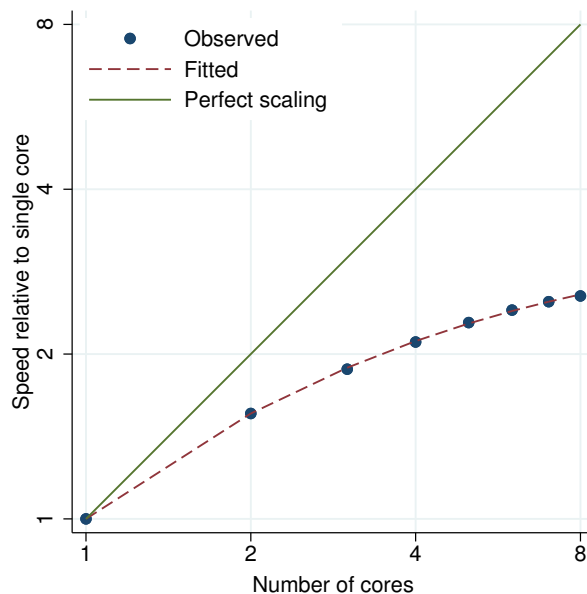


Figure 37. bayesmh normal performance plot.

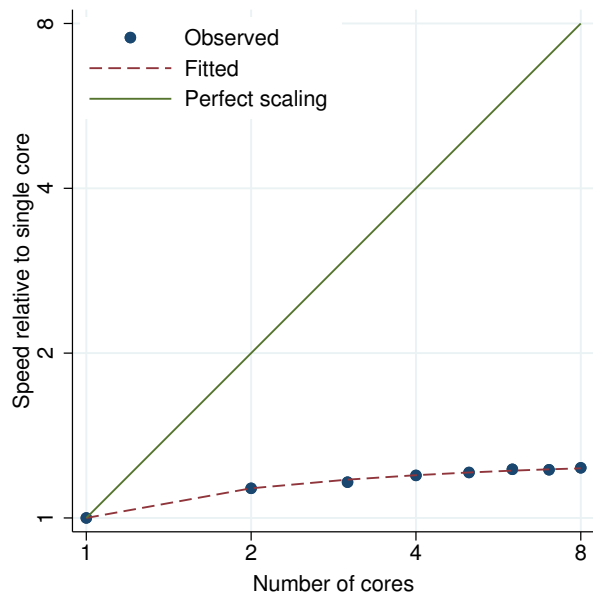


Figure 38. bayesmh normal gibbs performance plot.

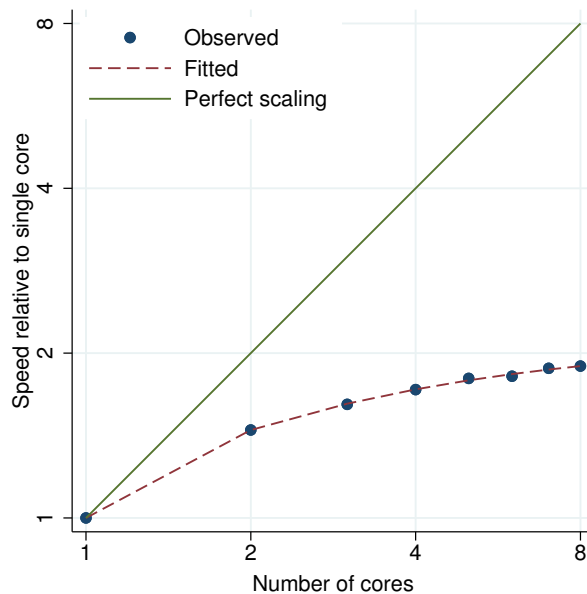


Figure 39. bayesmh normal re performance plot.

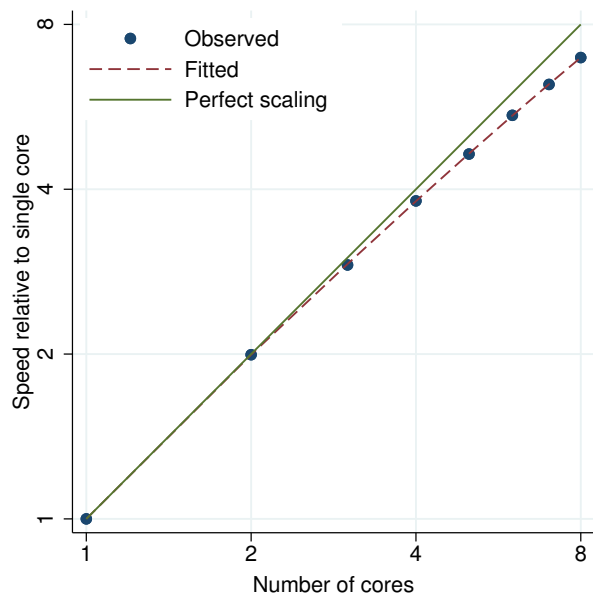


Figure 40. `betareg, link(logit)` performance plot.

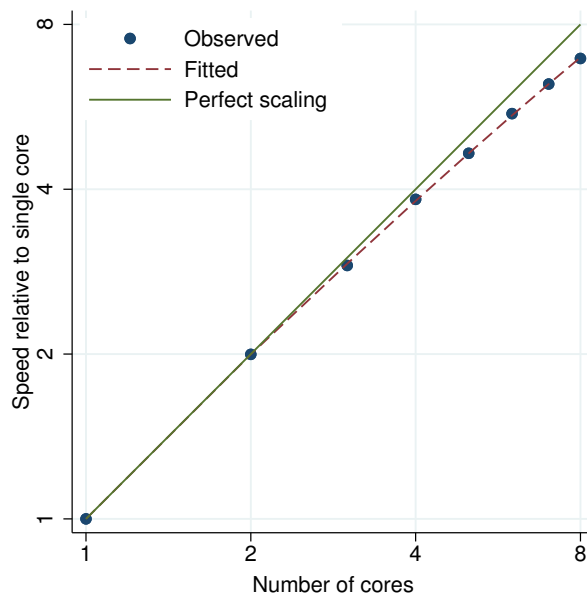


Figure 41. `betareg, link(probit)` performance plot.

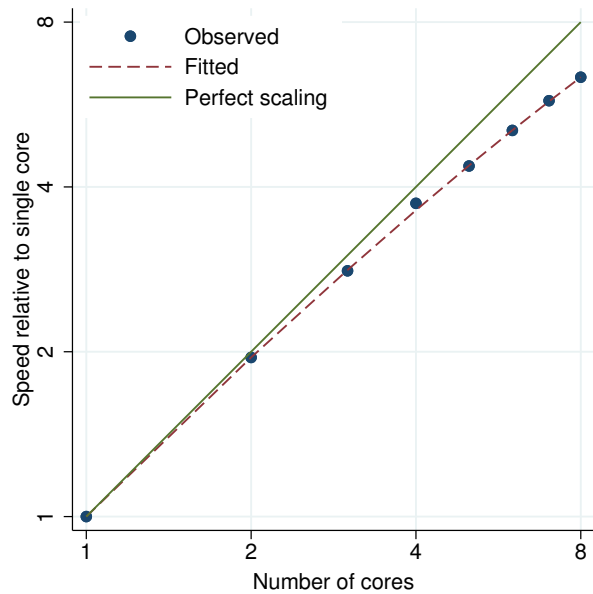


Figure 42. `binreg` performance plot.

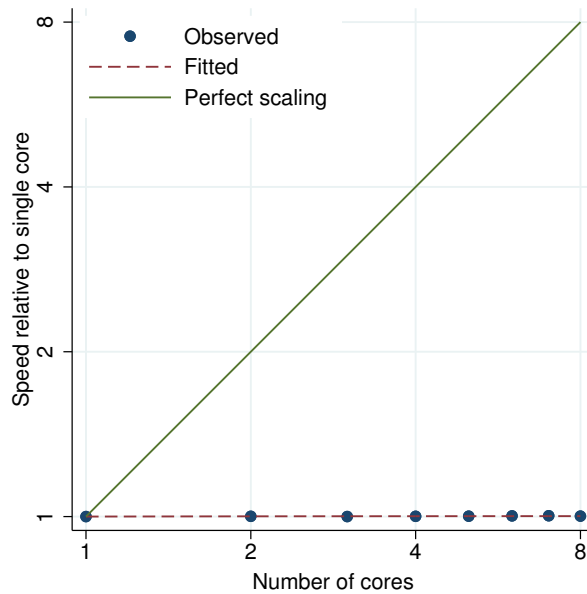
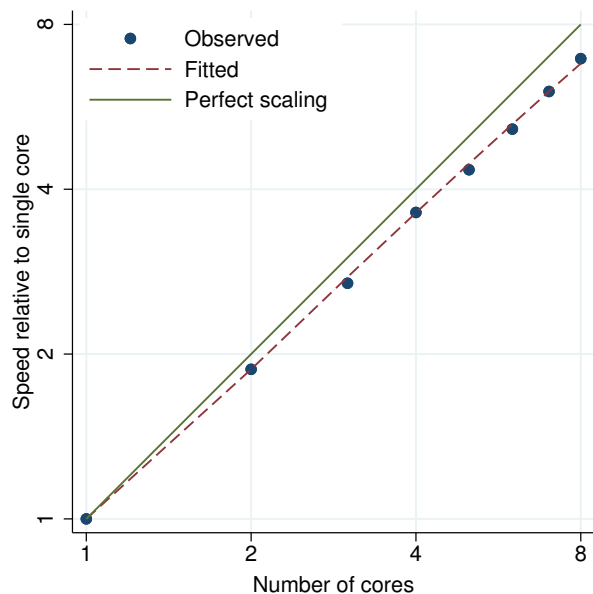
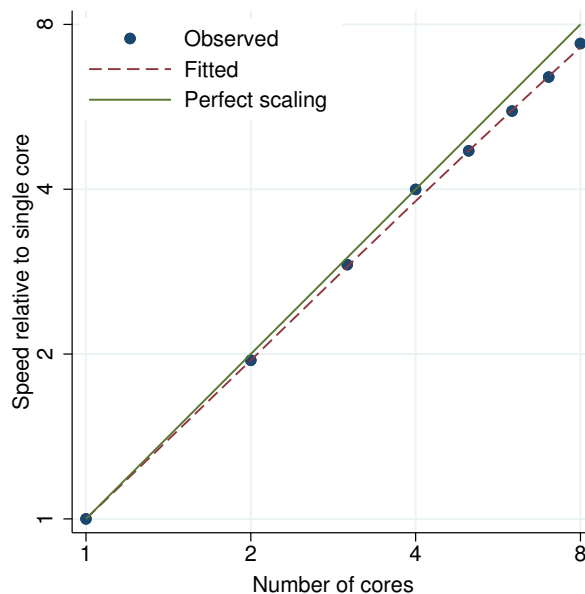
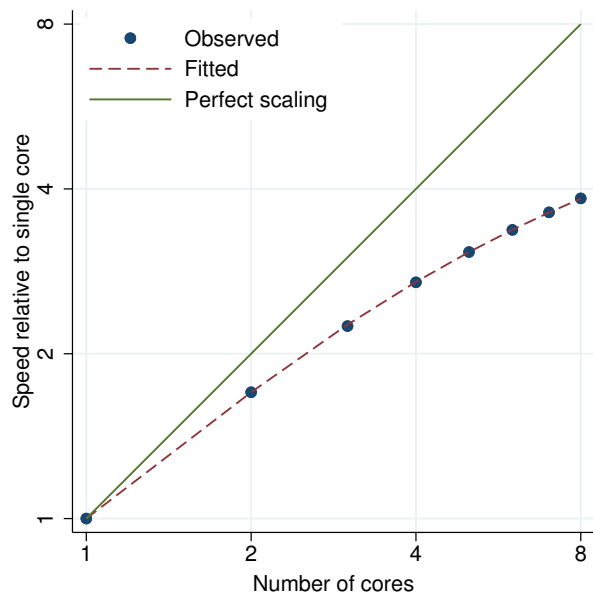
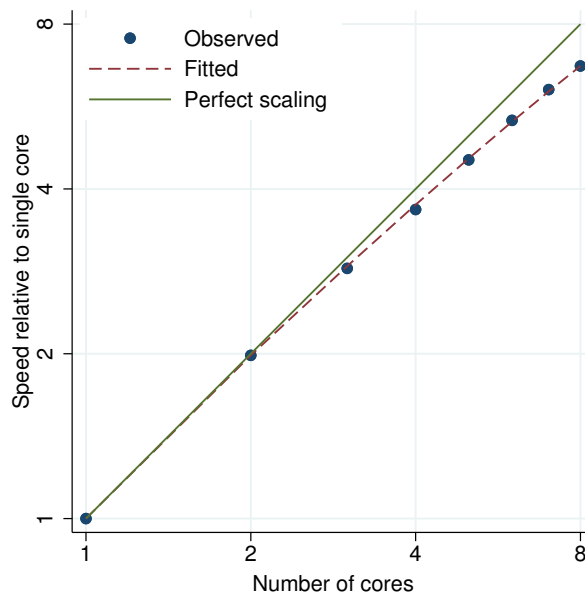


Figure 43. `biplot` performance plot.

Figure 44. `biprobit` performance plot.Figure 45. `biprobit` (seemingly unrelated) performance plot.Figure 46. `bitest` performance plot.Figure 47. `blogit` performance plot.

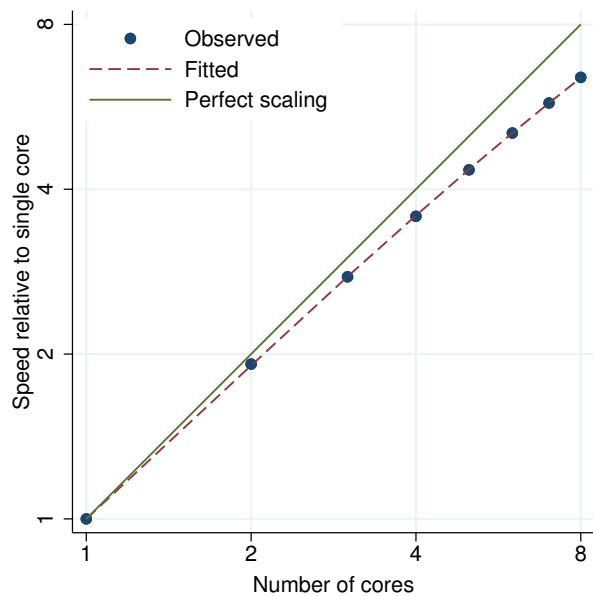


Figure 48. boxcox performance plot.

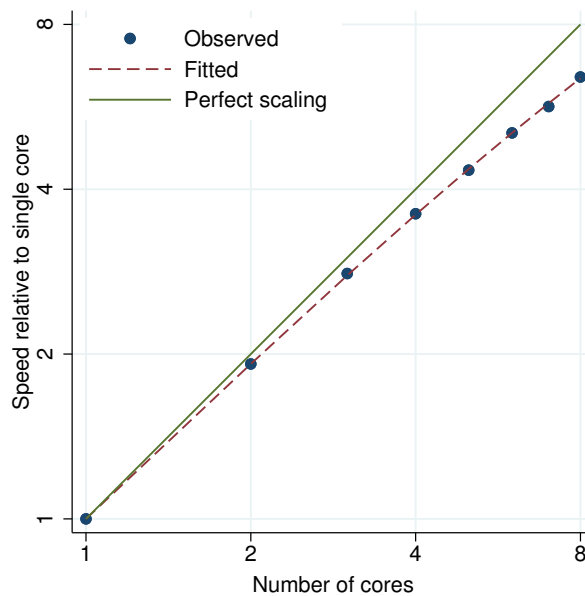


Figure 49. bprobit performance plot.

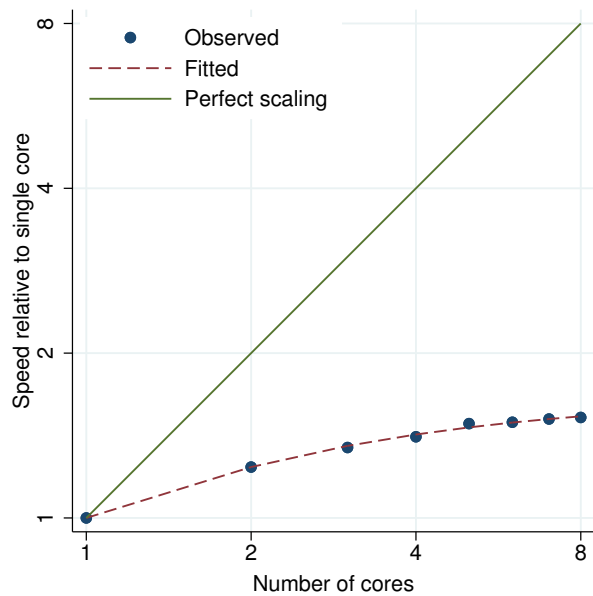


Figure 50. brier performance plot.

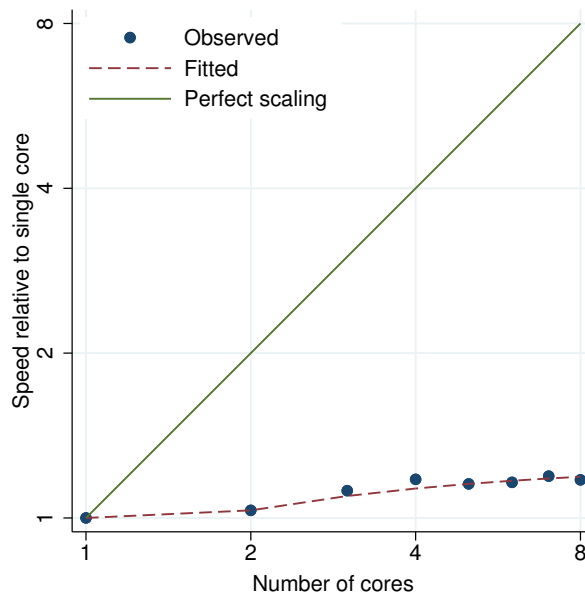
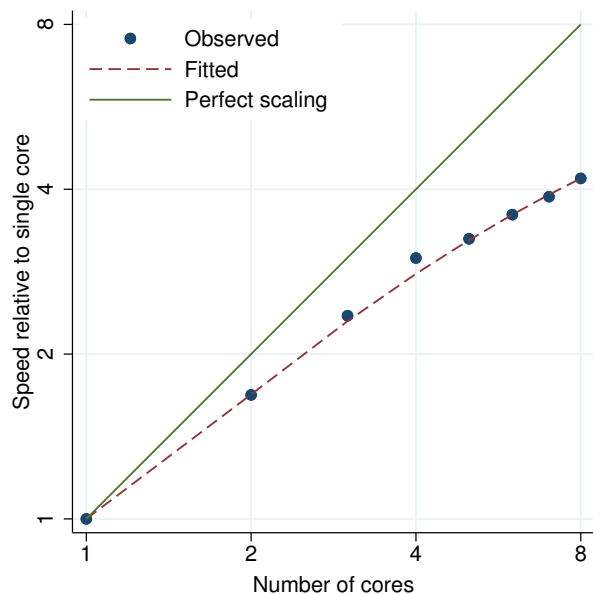
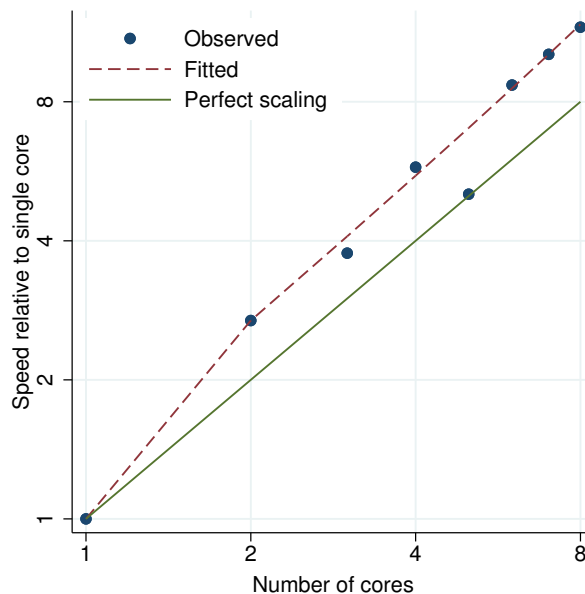
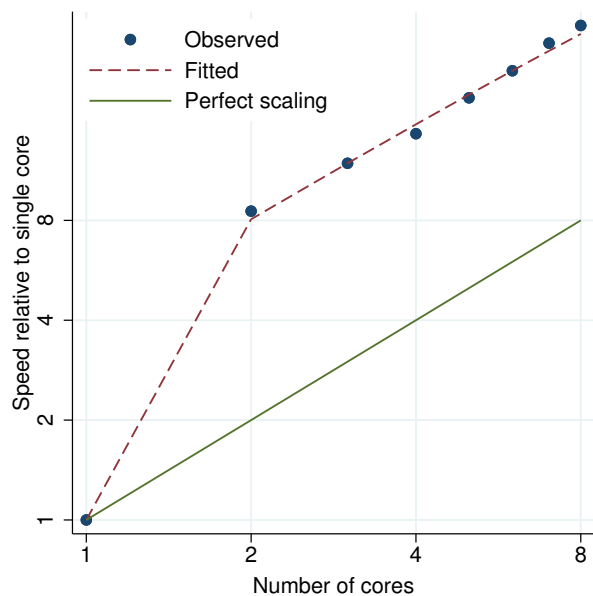
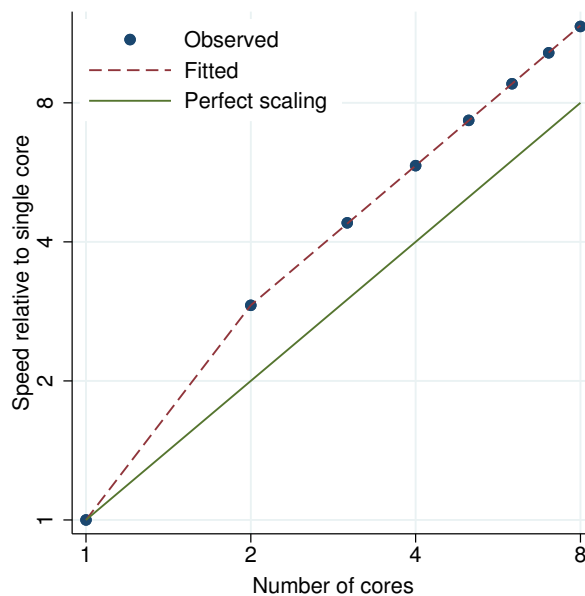


Figure 51. bsample performance plot.

Figure 52. `bstat` performance plot.Figure 53. `by: generate` performance plot.Figure 54. `by: generate` (small groups) performance plot.Figure 55. `by: replace` performance plot.

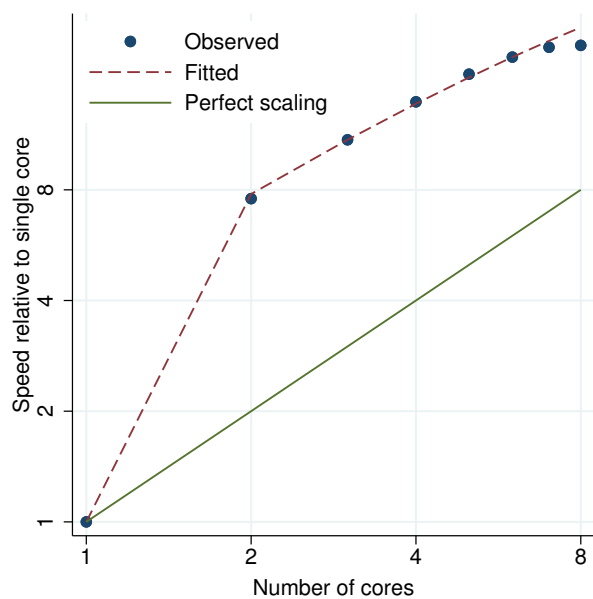


Figure 56. `by: replace` (small groups) performance plot.

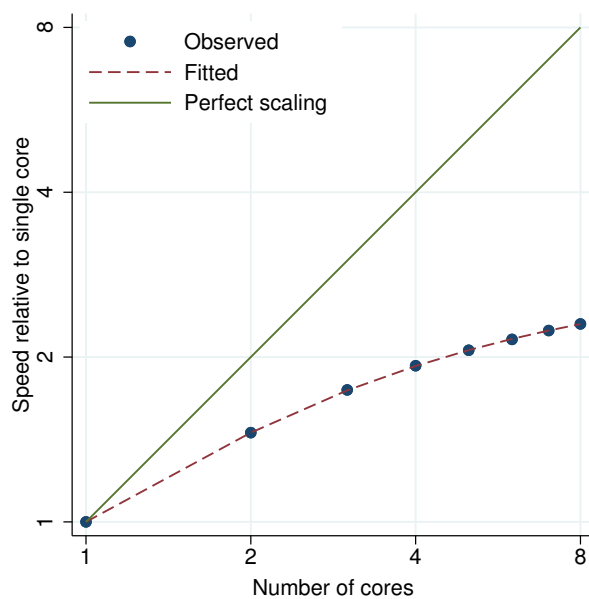


Figure 57. `ca` performance plot.

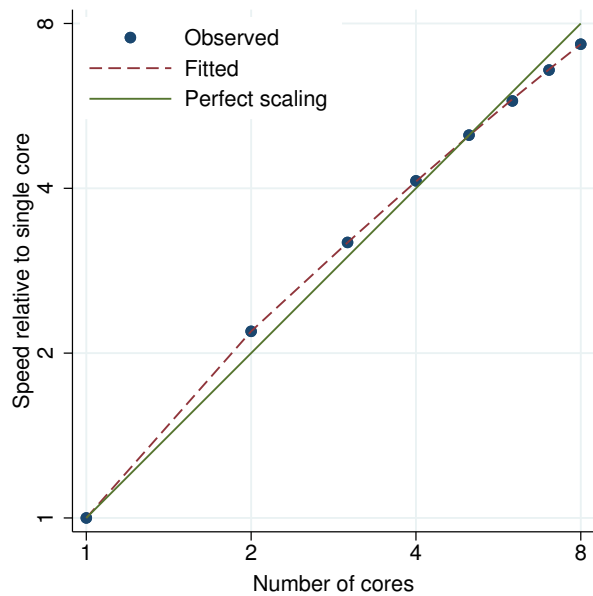


Figure 58. `candisc` performance plot.

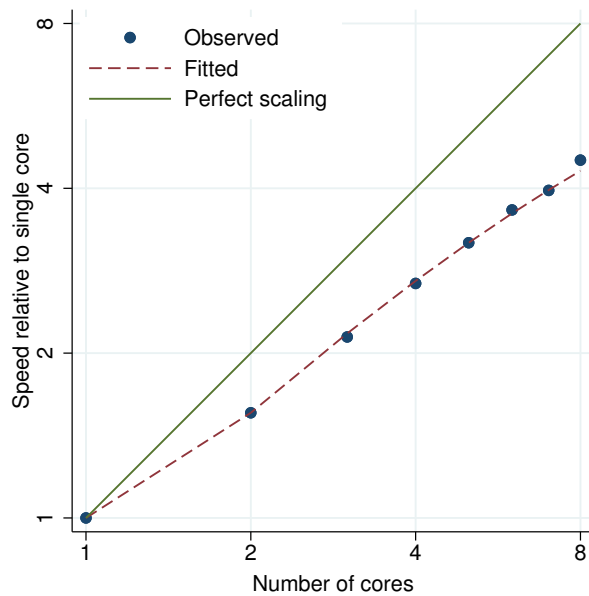


Figure 59. `canon` performance plot.



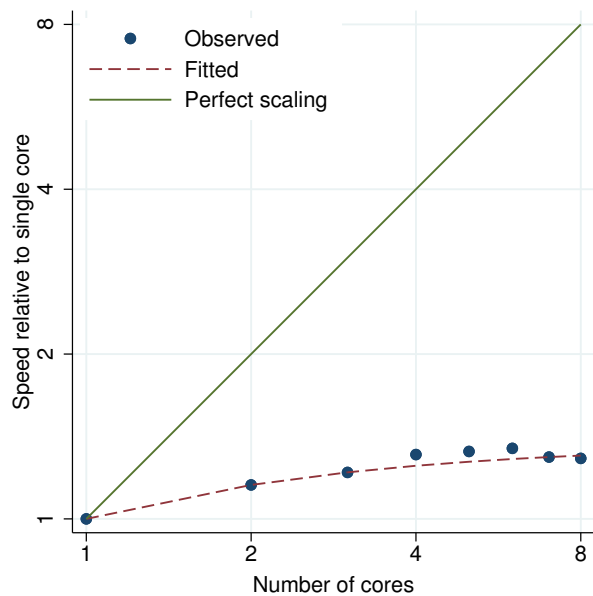


Figure 60. cc performance plot.

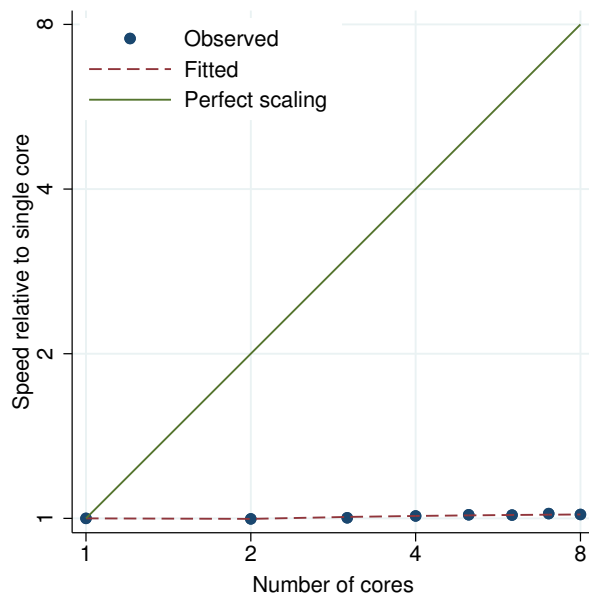


Figure 61. by: cc performance plot.

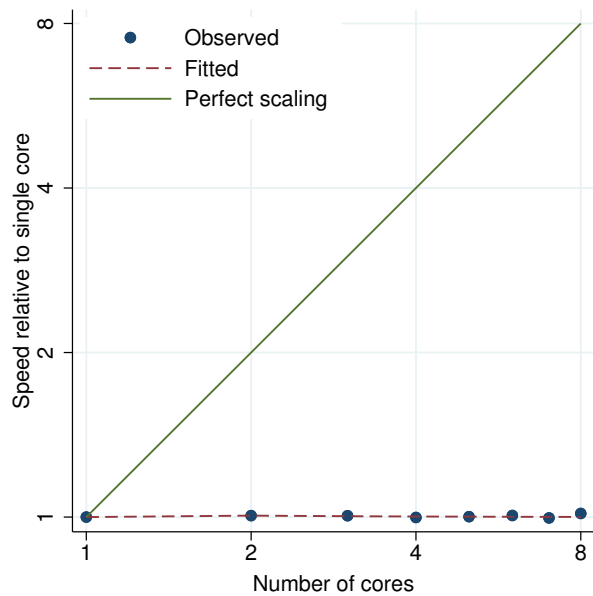


Figure 62. centile performance plot.

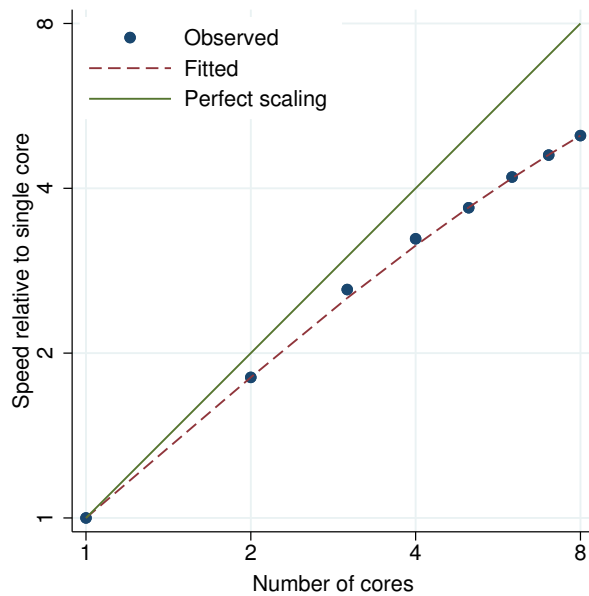
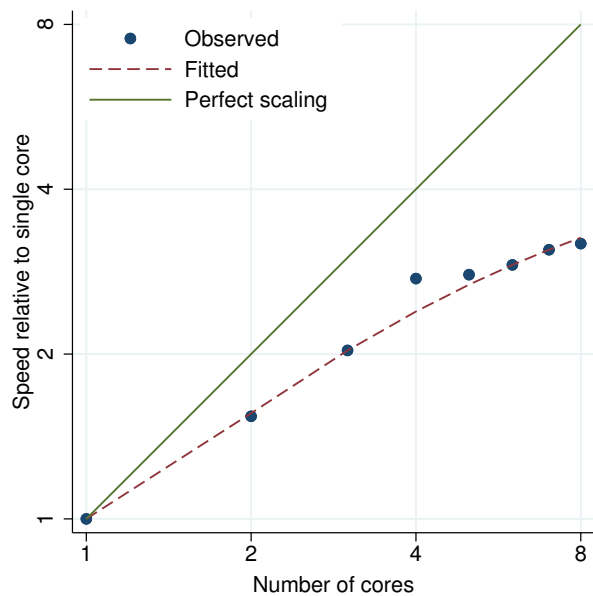
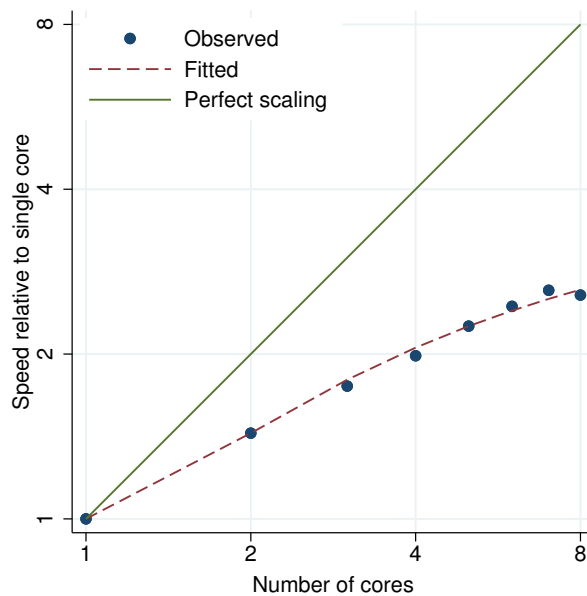
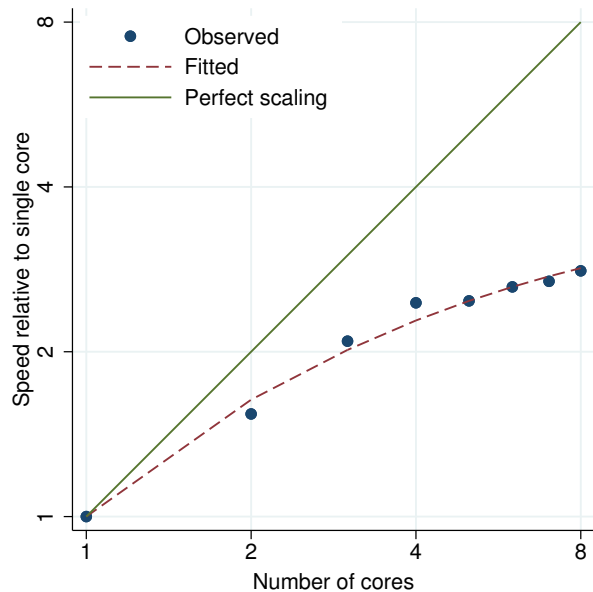
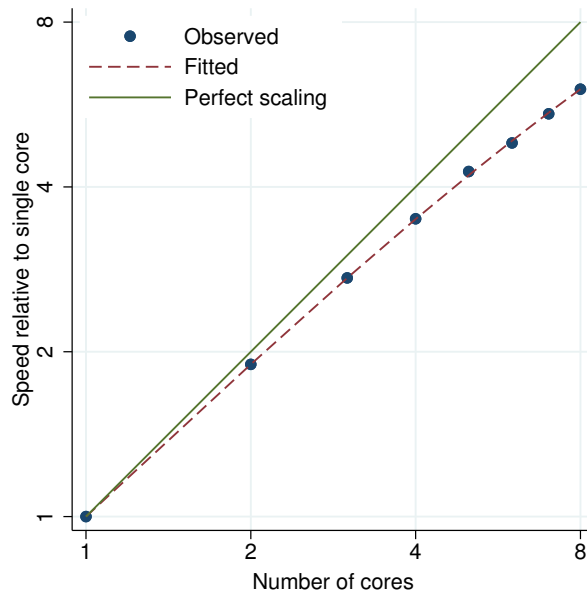


Figure 63. churdle linear performance plot.

Figure 64. `ci means` performance plot.Figure 65. `ci means, poisson` performance plot.Figure 66. `ci proportions` performance plot.Figure 67. `clogit (k1 to k2 matching)` performance plot.

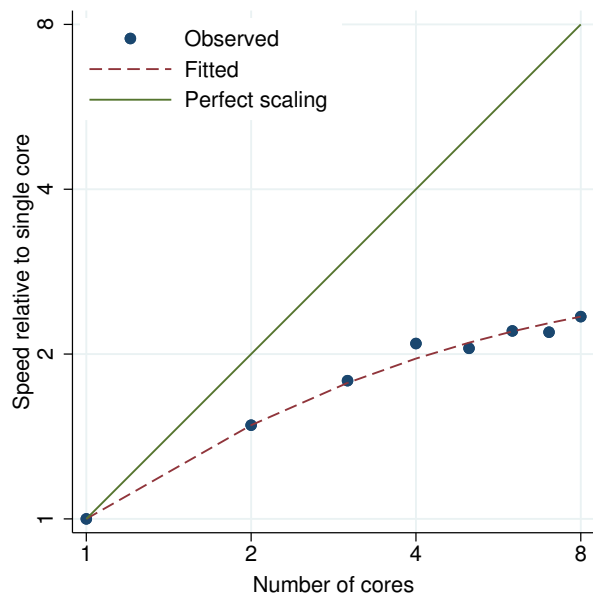


Figure 68. `clogit` (1 to k matching) performance plot.

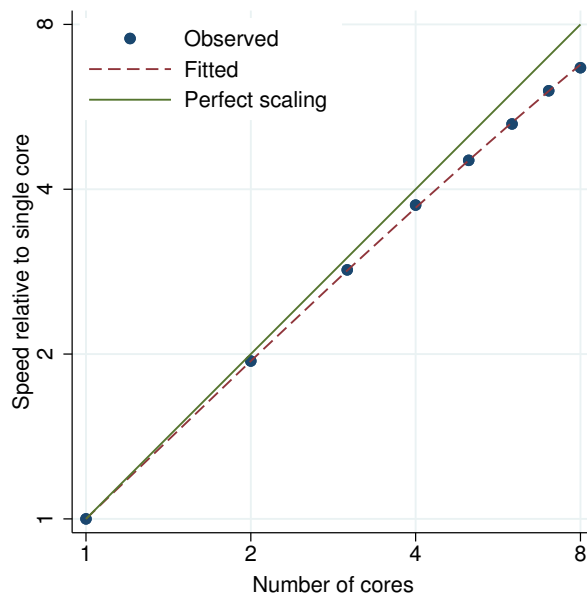


Figure 69. `cloglog` performance plot.

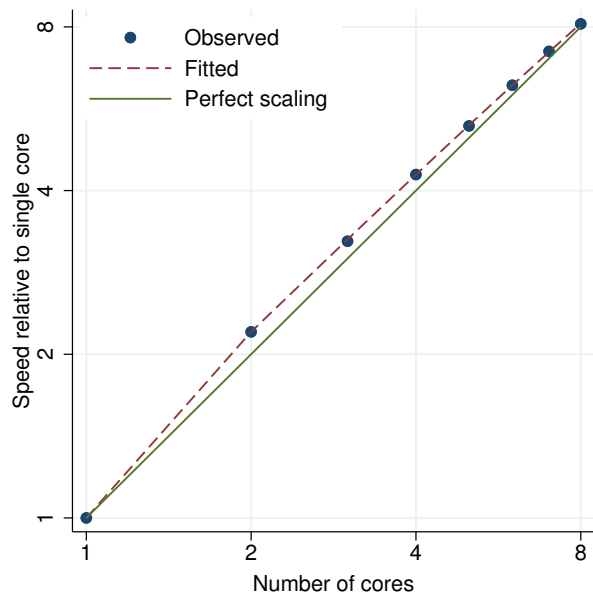


Figure 70. `cluster averagelinkage` performance plot.

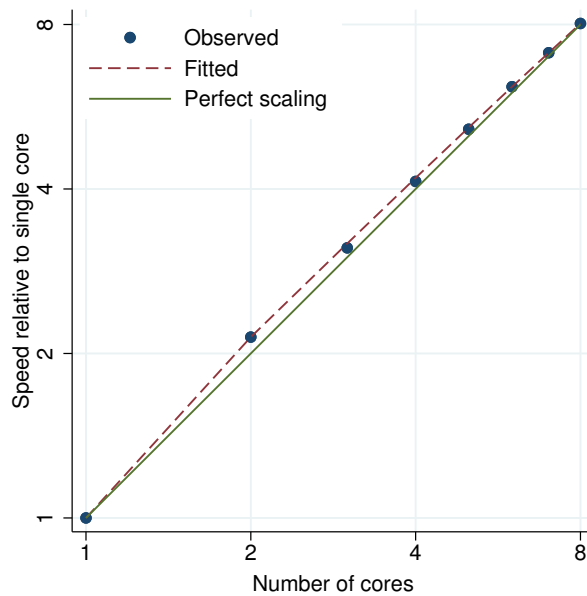
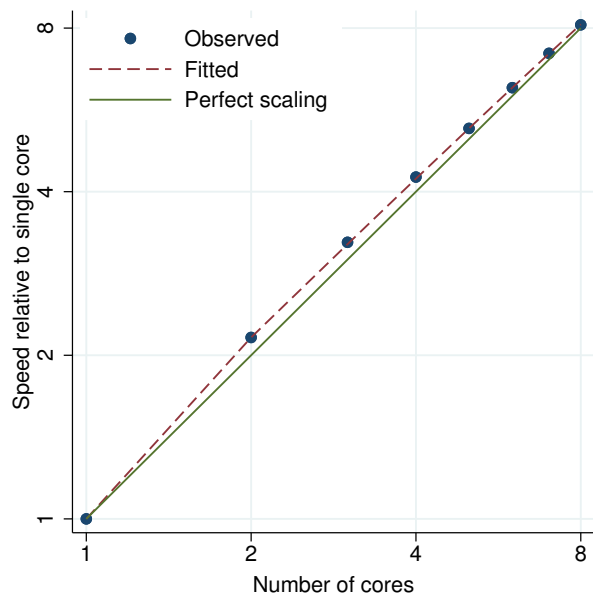
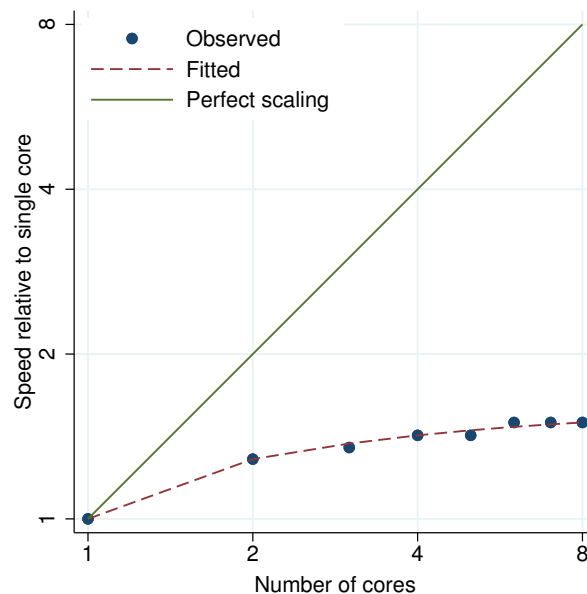
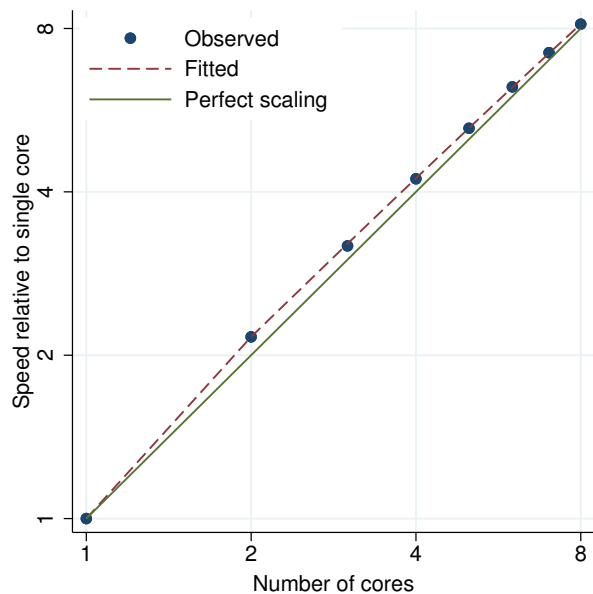
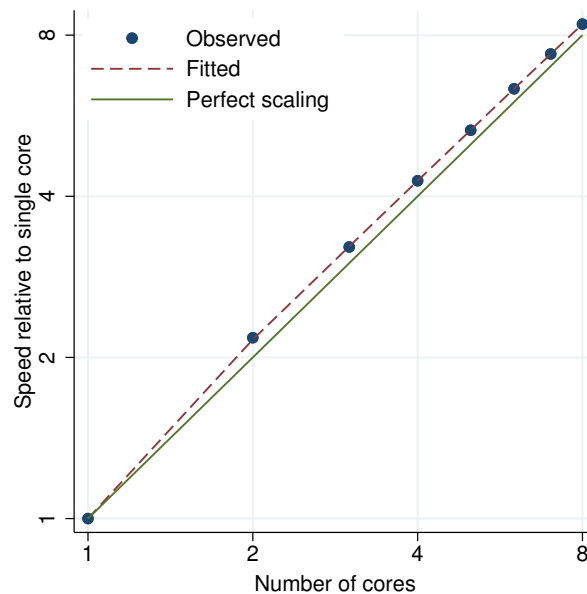


Figure 71. `cluster centroidlinkage` performance plot.

Figure 72. `cluster completelinkage` performance plot.Figure 73. `cluster generate` performance plot.Figure 74. `cluster kmeans` performance plot.Figure 75. `cluster kmedians` performance plot.

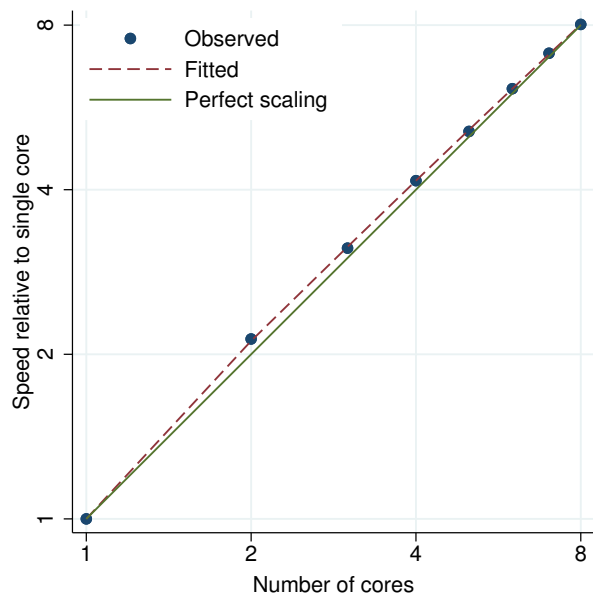


Figure 76. cluster medianlinkage performance plot.

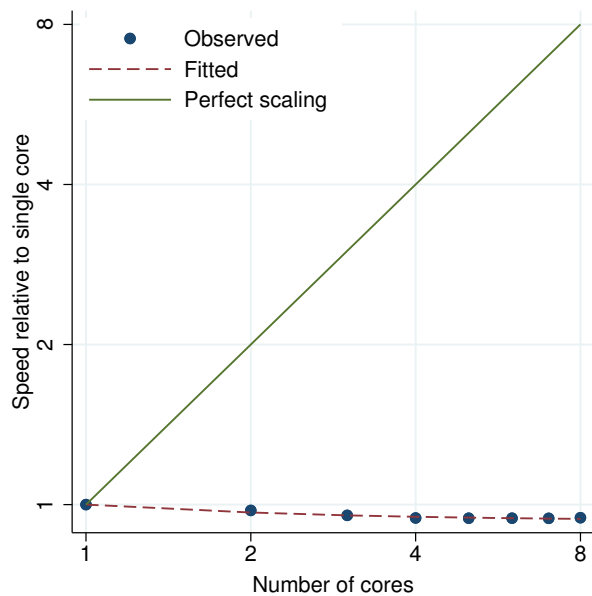


Figure 77. cluster singlelinkage performance plot.

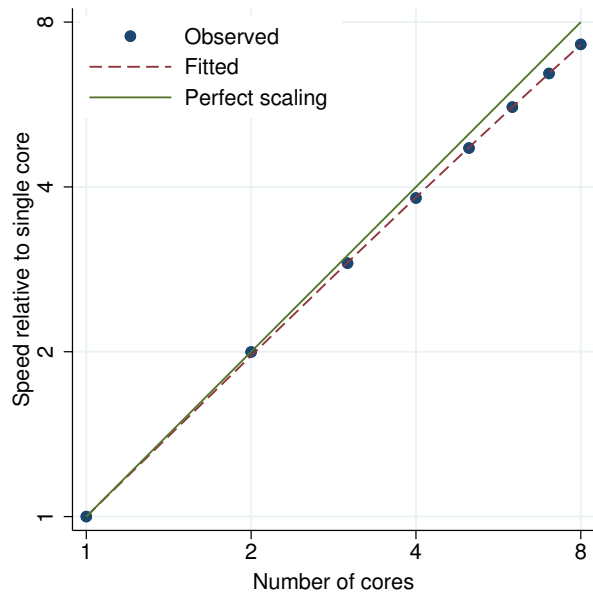


Figure 78. cluster wardslinkage performance plot.

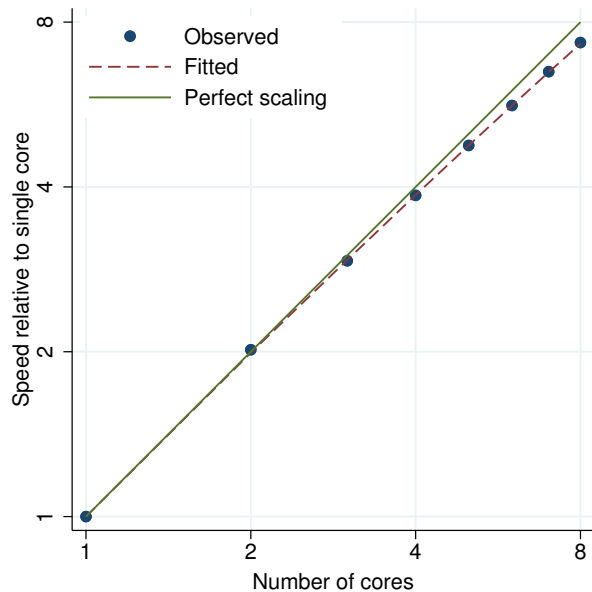
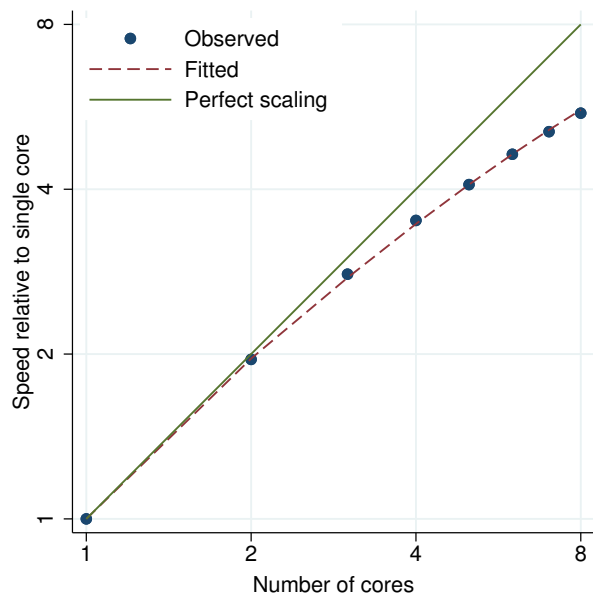
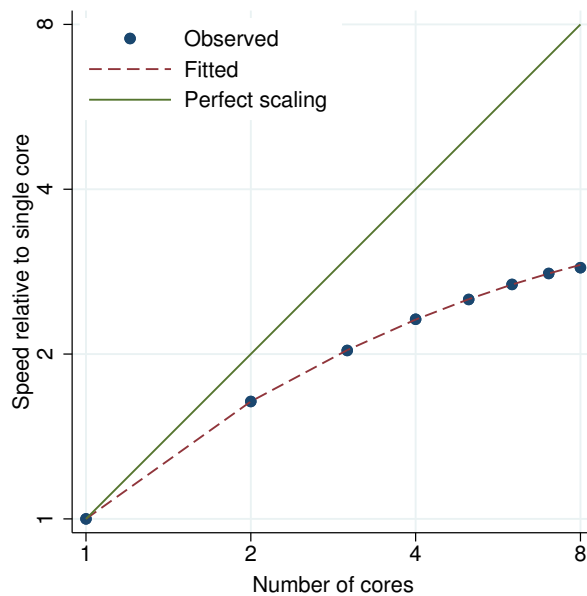
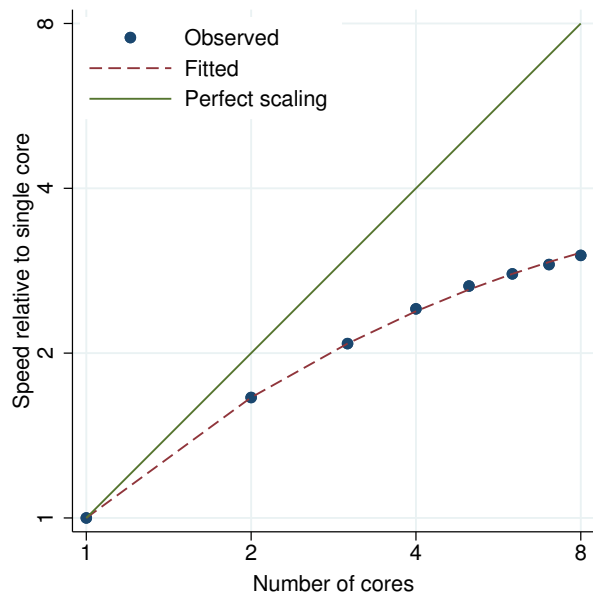
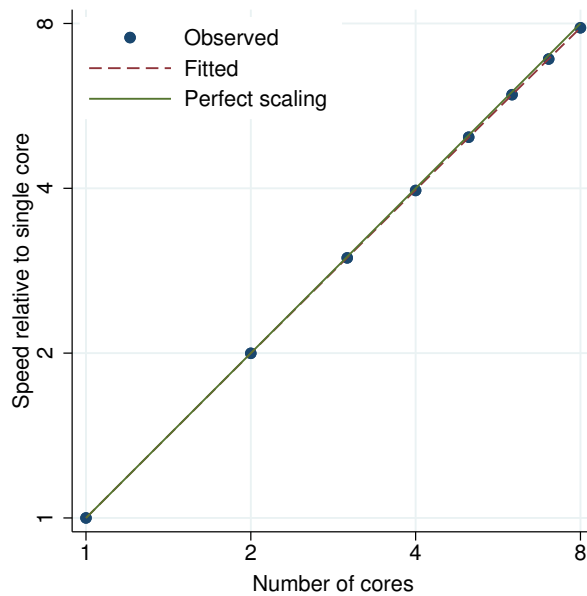


Figure 79. cluster waveragelinkage performance plot.

Figure 80. `cmclogit` performance plot.Figure 81. `cmmprobit` performance plot.Figure 82. `cmroprobit` performance plot.Figure 83. `cnsreg` performance plot.

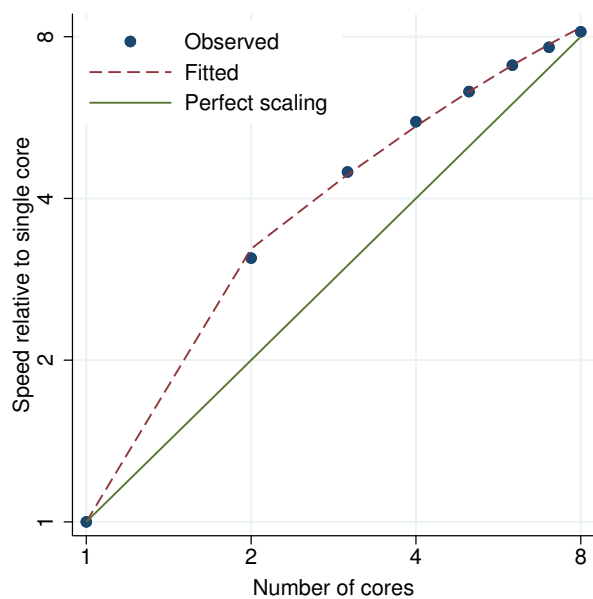


Figure 84. codebook performance plot.

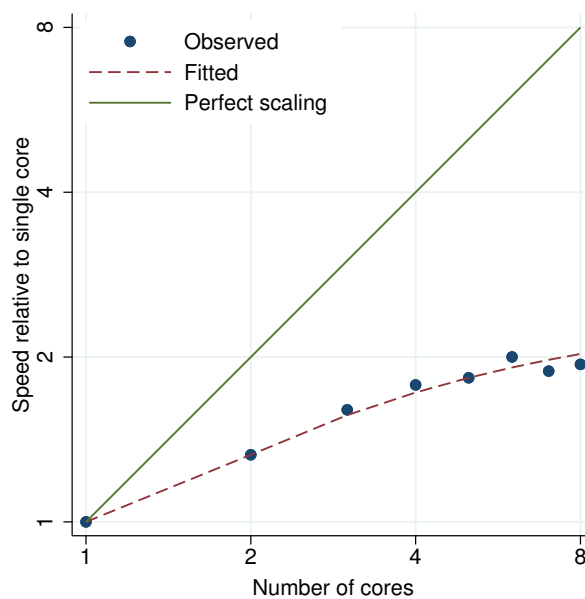


Figure 85. collapse performance plot.

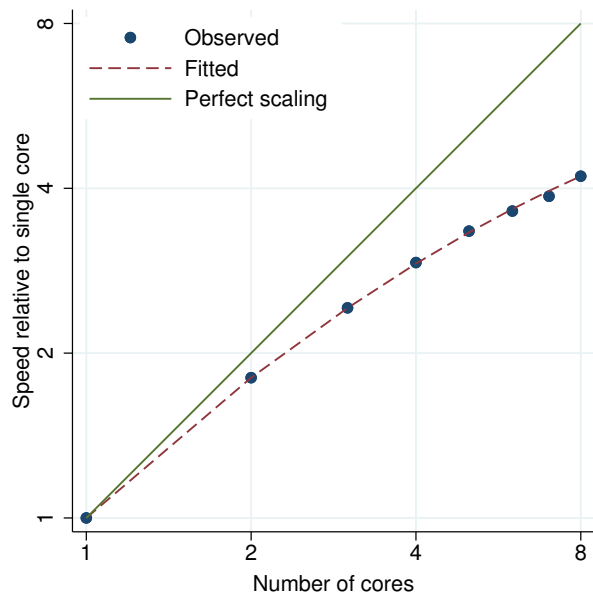


Figure 86. compare performance plot.

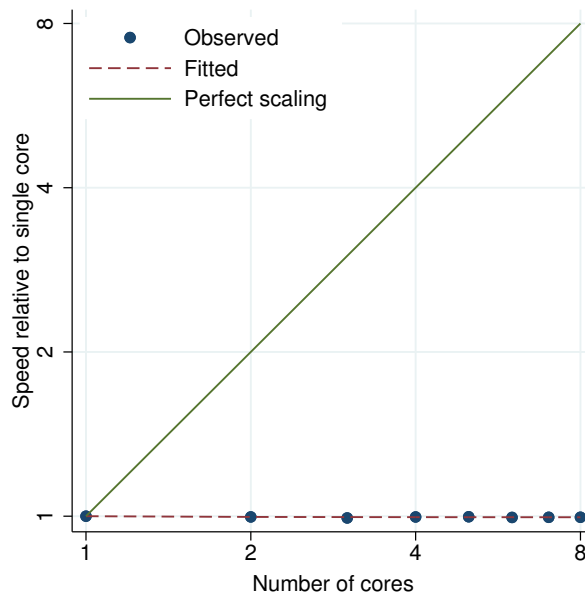
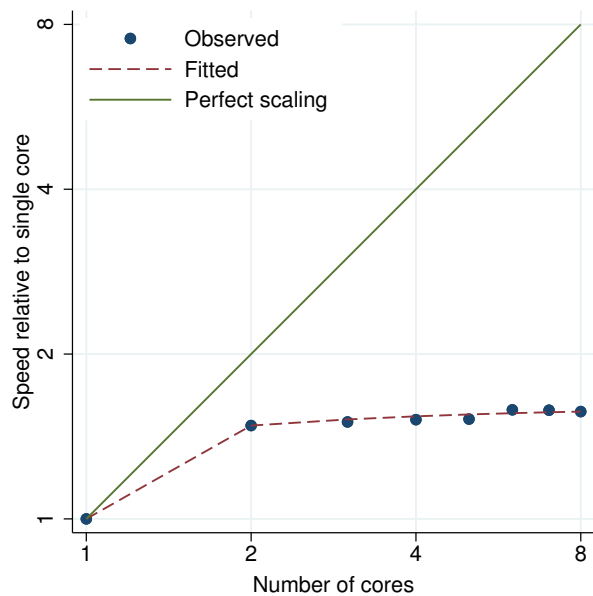
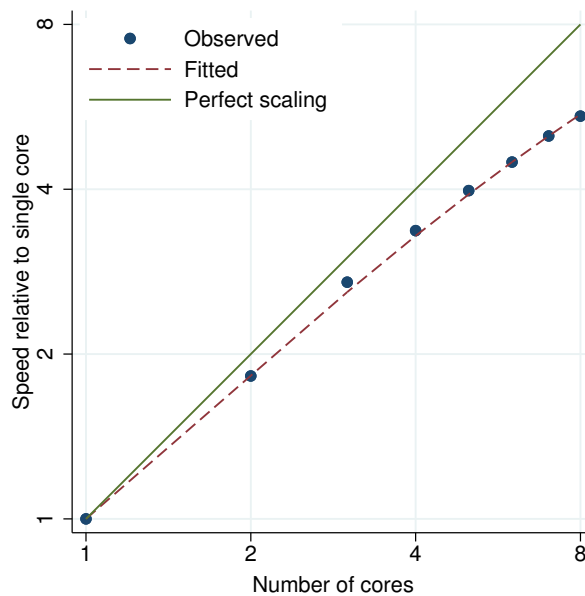
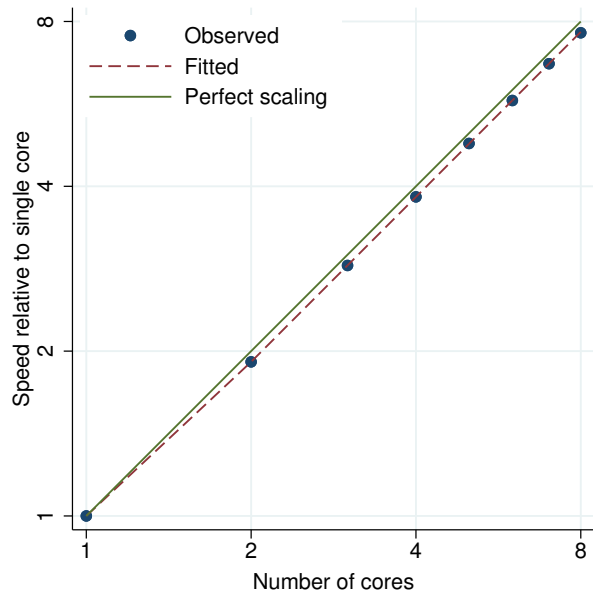
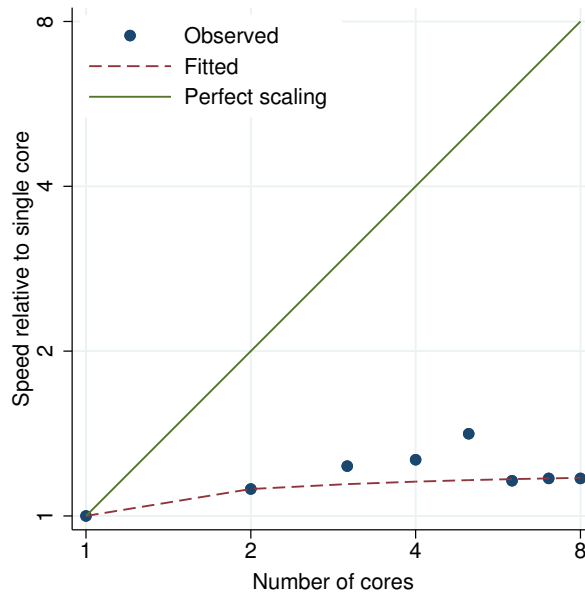
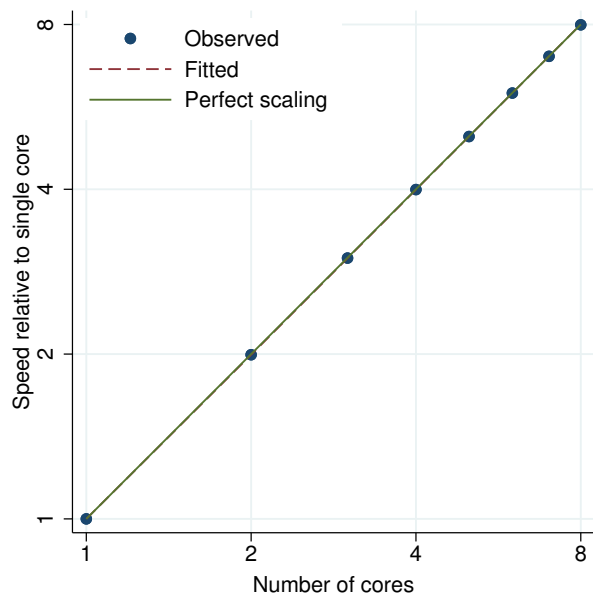
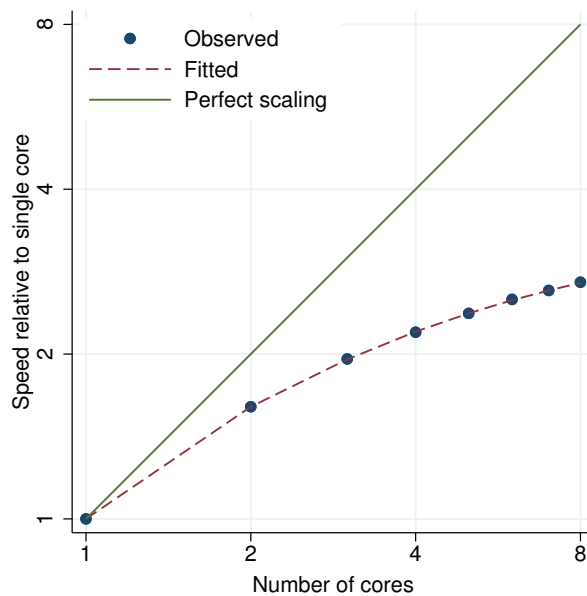
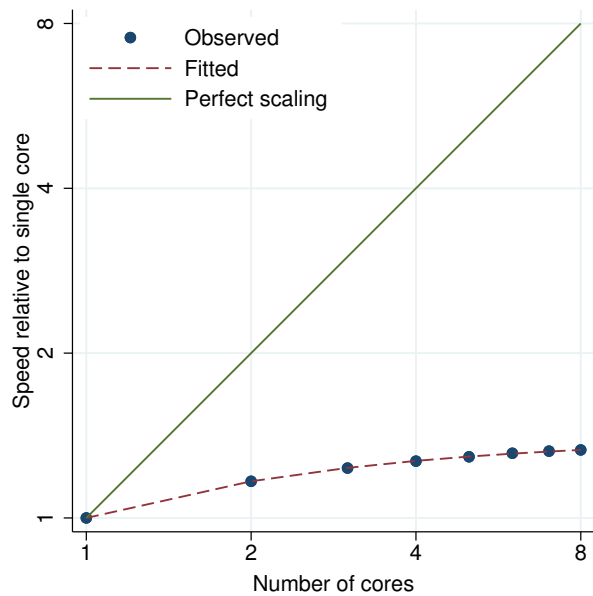
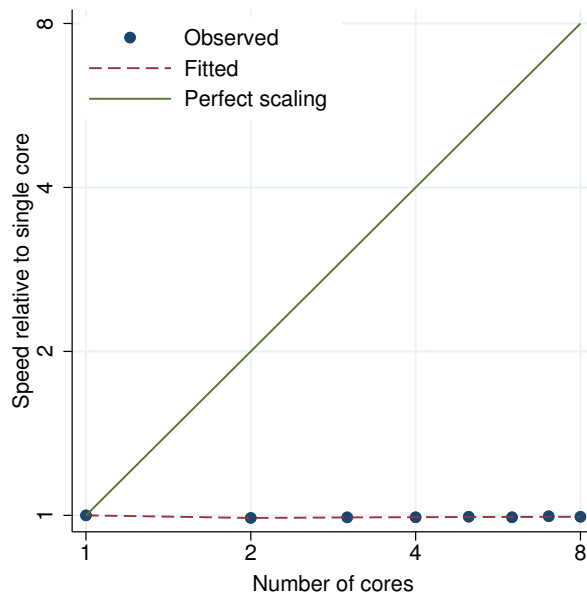
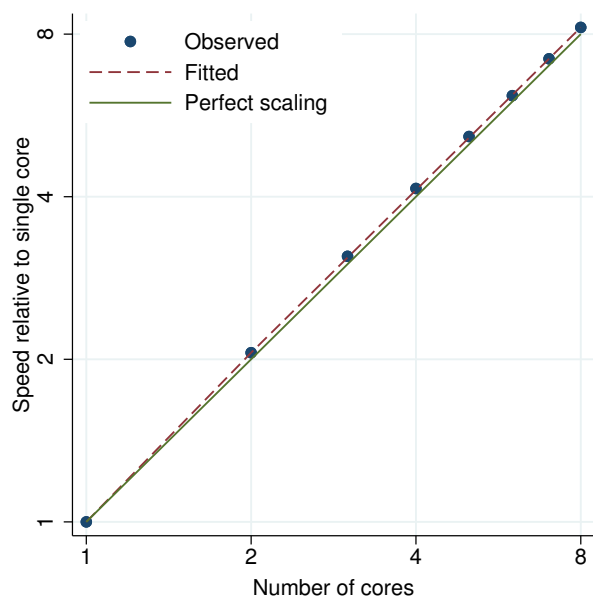
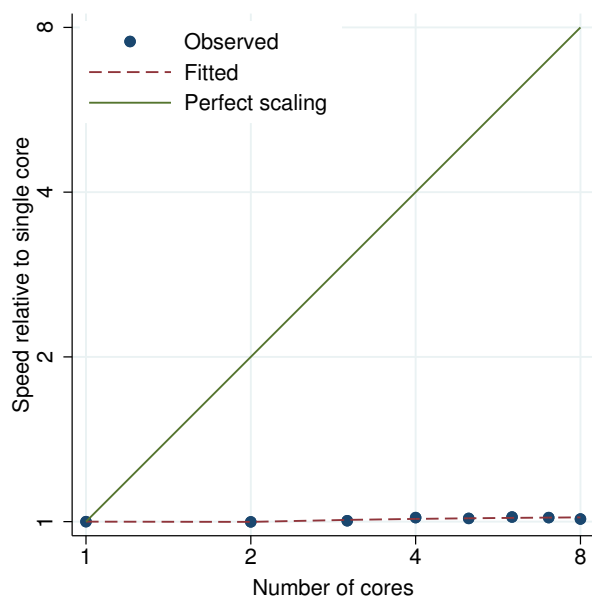
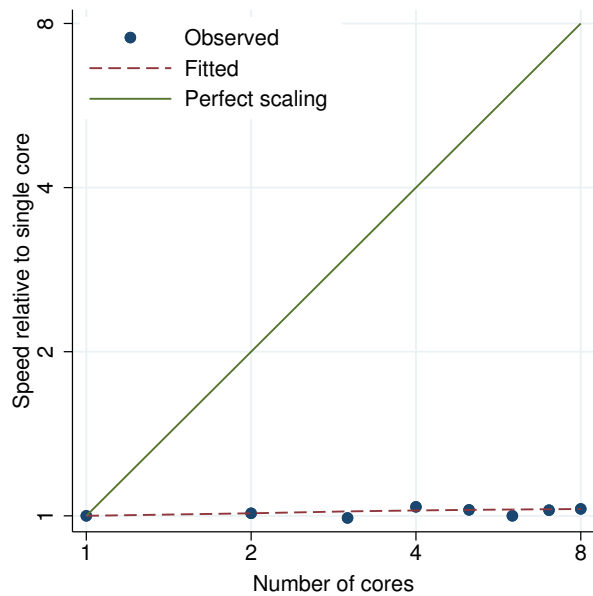
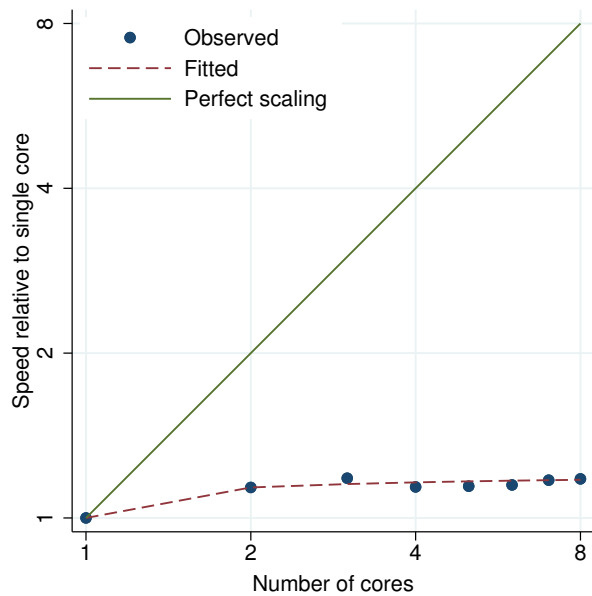


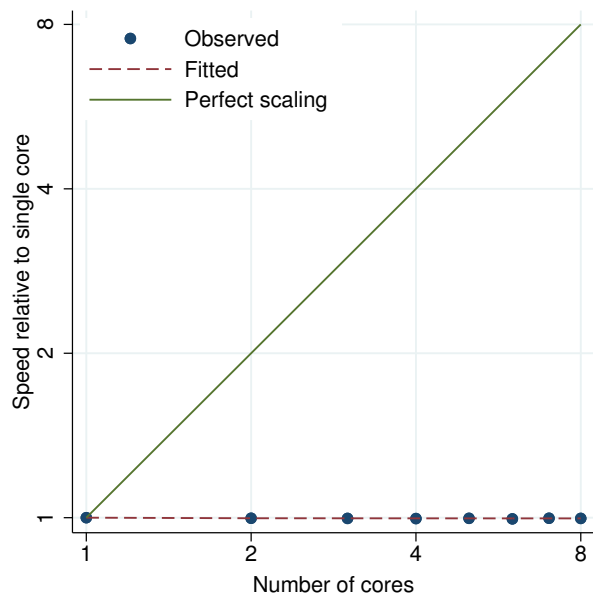
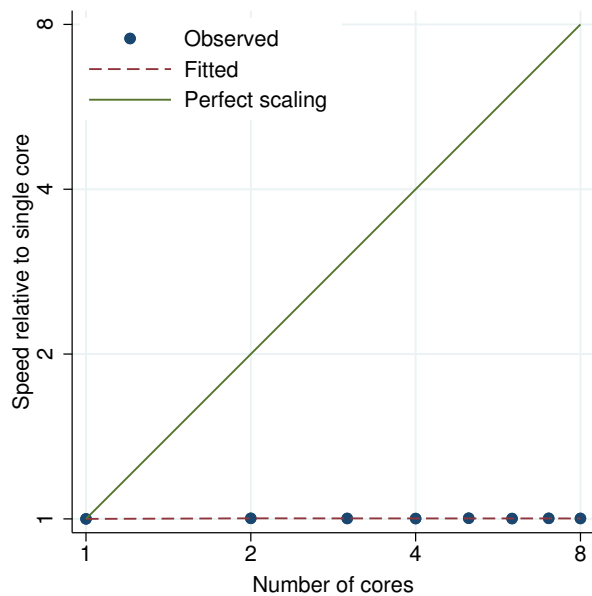
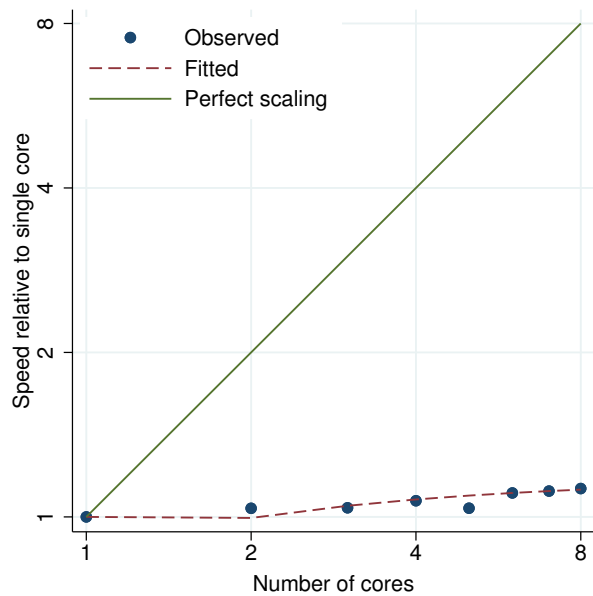
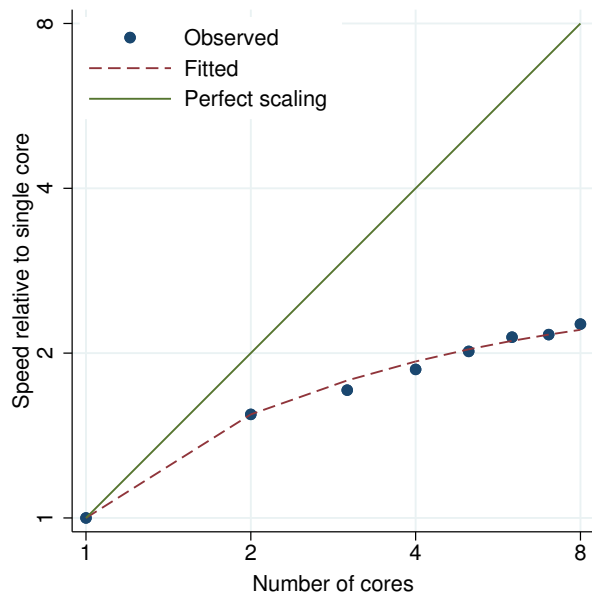
Figure 87. compress performance plot.

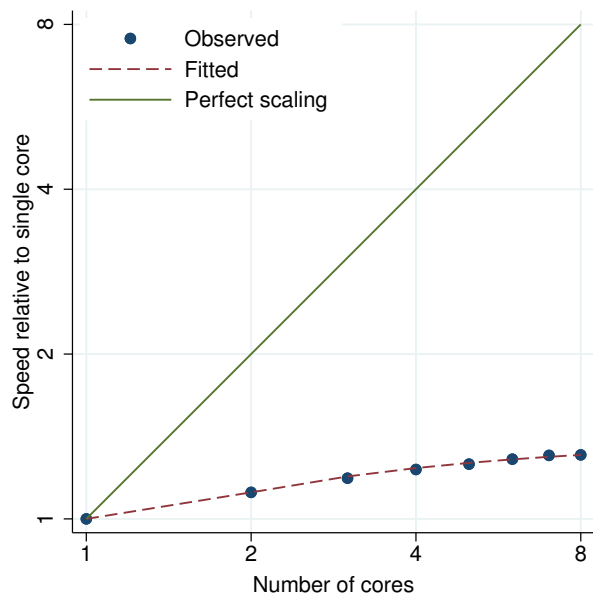
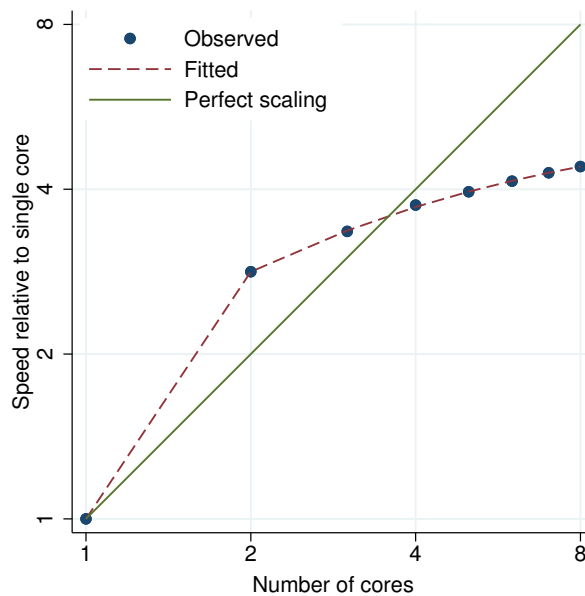
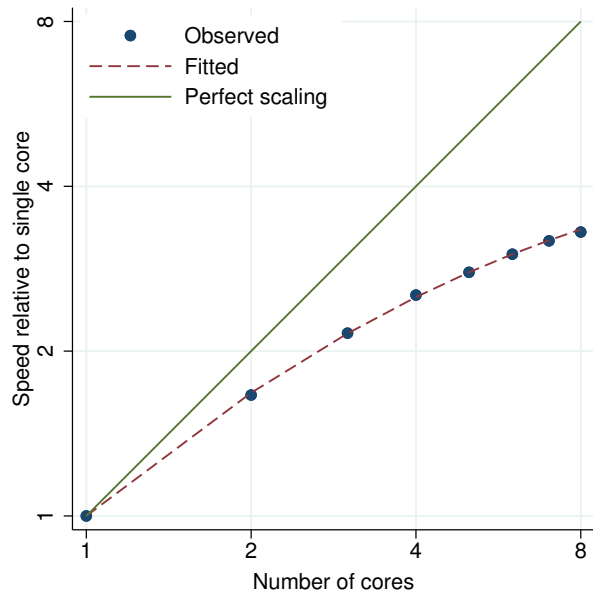
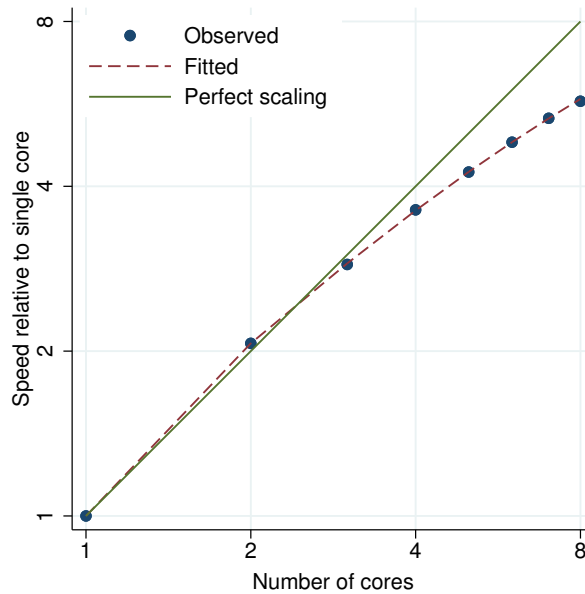
Figure 88. `contract` performance plot.Figure 89. `corr2data` performance plot.Figure 90. `correlate` performance plot.Figure 91. `corrgram` performance plot.

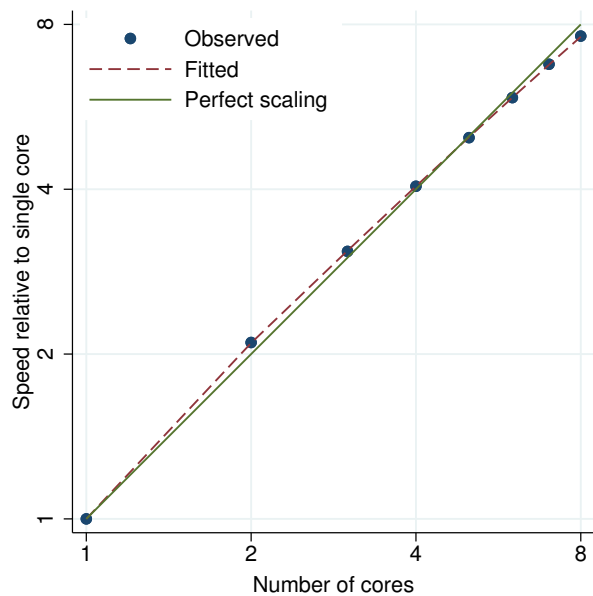
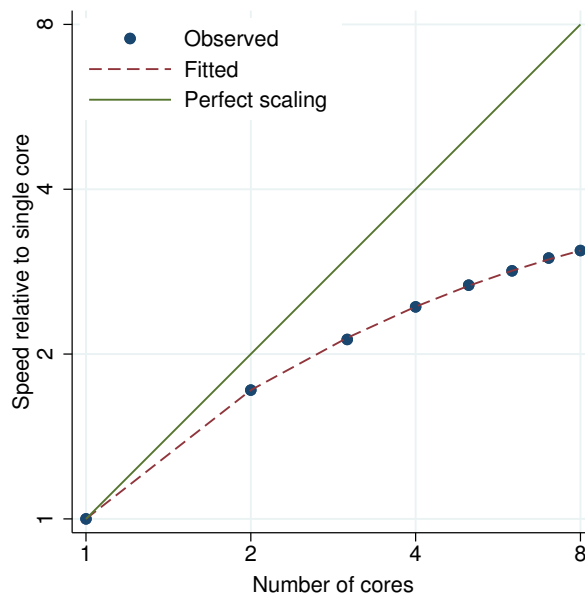
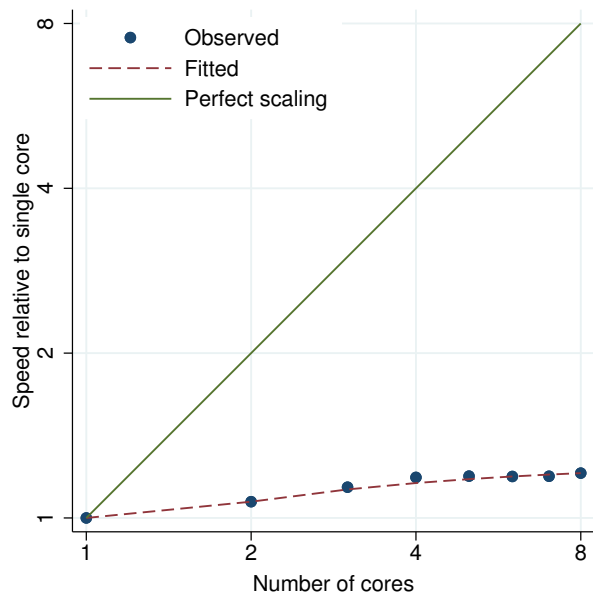
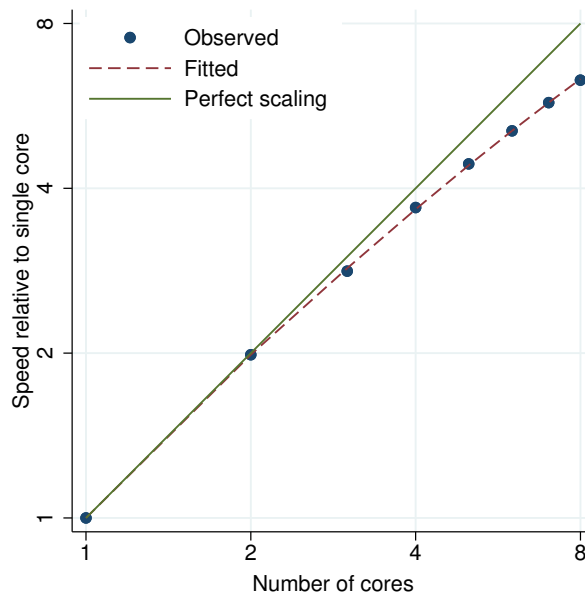


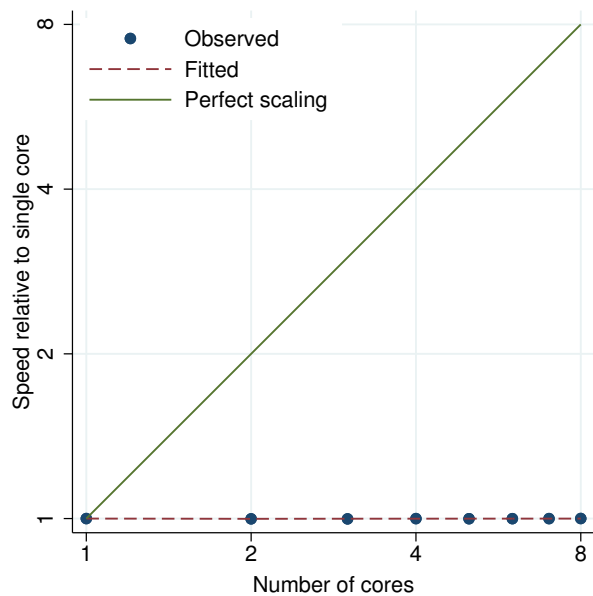
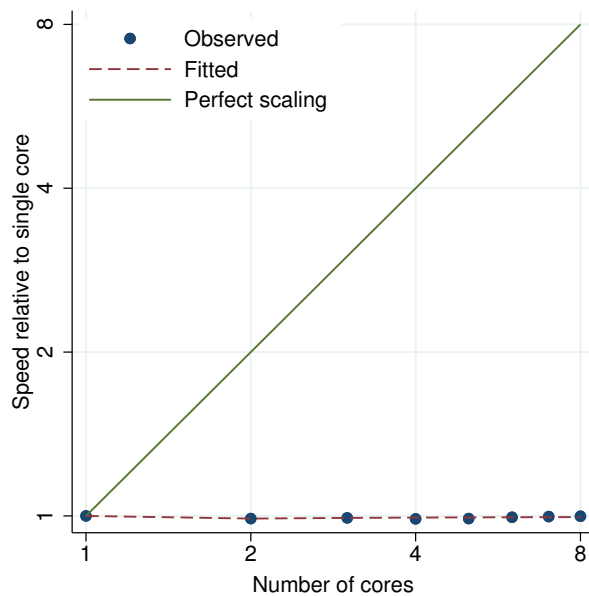
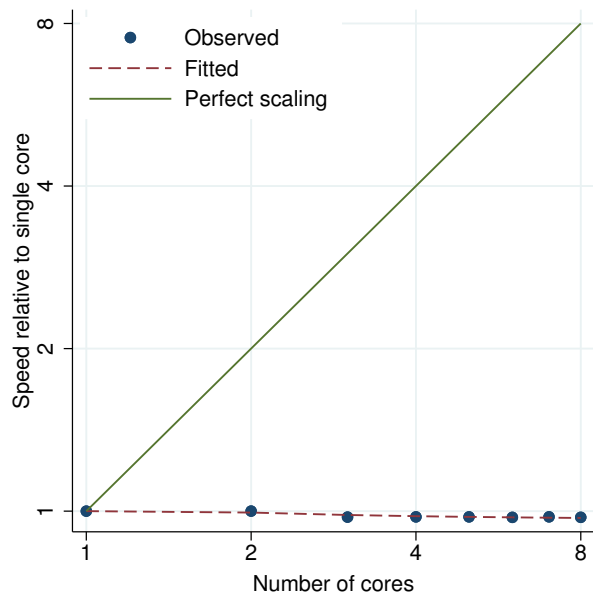
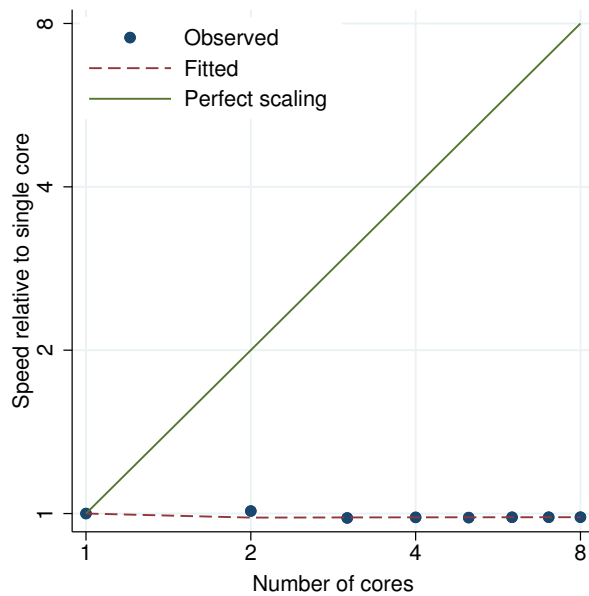
Figure 92. `count` performance plot.Figure 93. `cpoisson` performance plot.Figure 94. `cs` performance plot.Figure 95. `by: cs` performance plot.

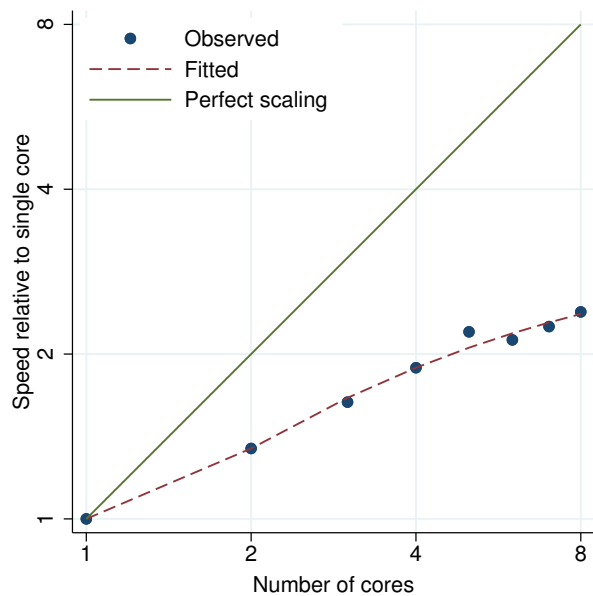
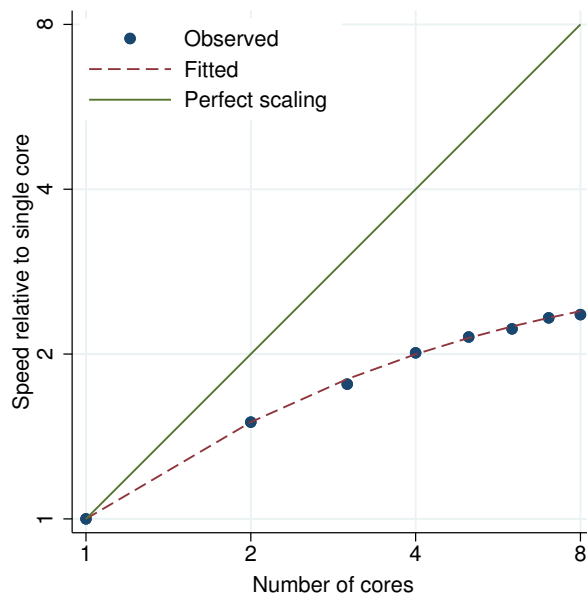
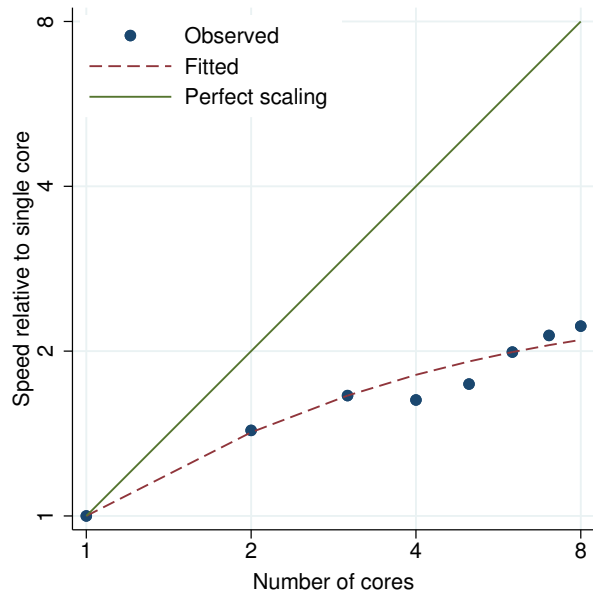
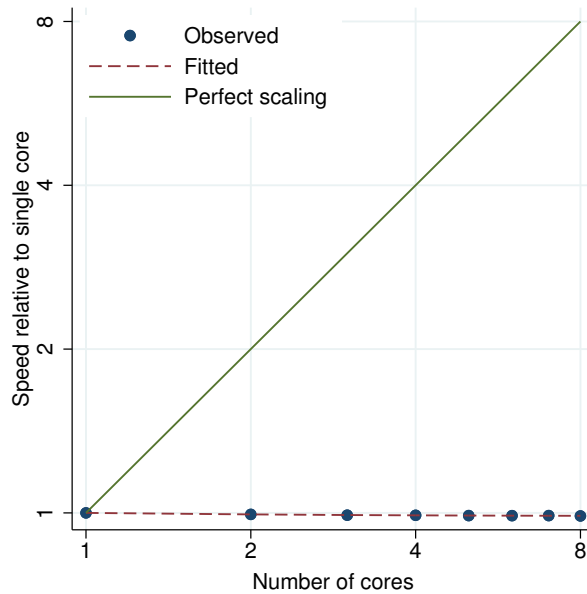
Figure 96. `ctset` performance plot.Figure 97. `cttost` performance plot.Figure 98. `cumul` performance plot.Figure 99. `cusum` performance plot.

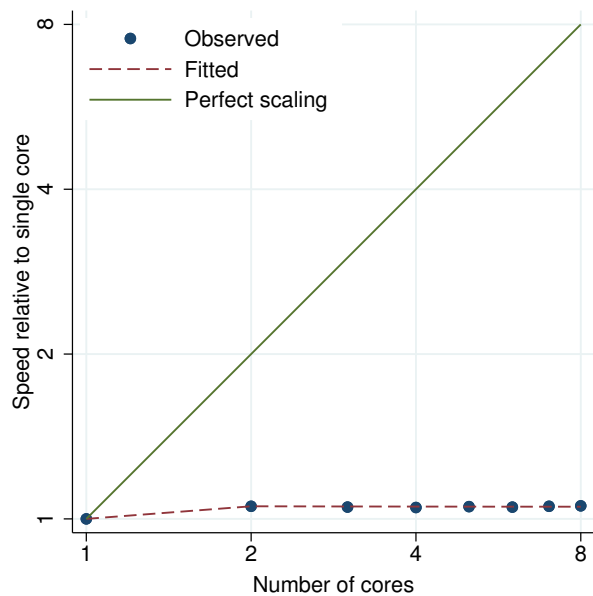
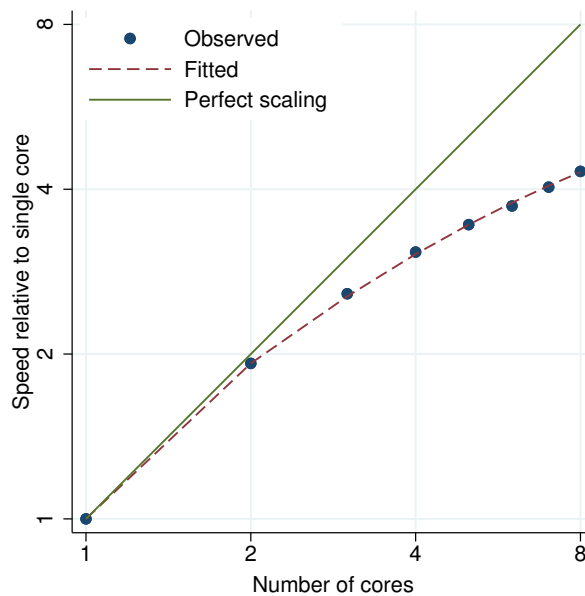
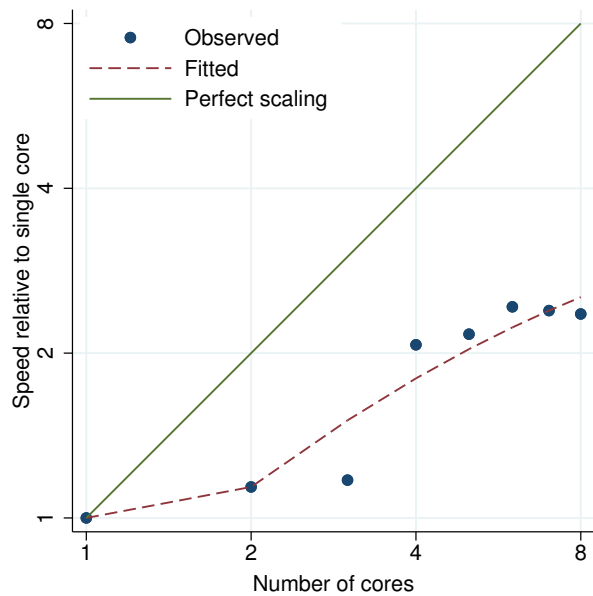
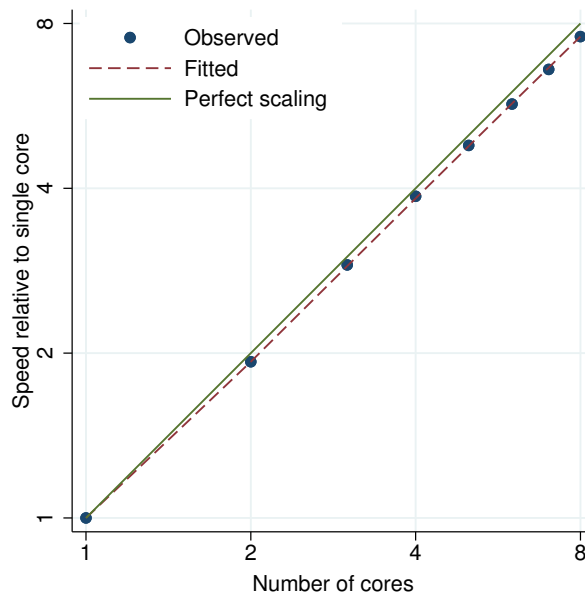
Figure 100. `datasignature` performance plot.Figure 101. `decode` performance plot.Figure 102. `destring` performance plot.Figure 103. `dfactor` performance plot.

Figure 104. `dfpls` performance plot.Figure 105. `dfuller` performance plot.Figure 106. `discrim knn` performance plot.Figure 107. `discrim lda` performance plot.

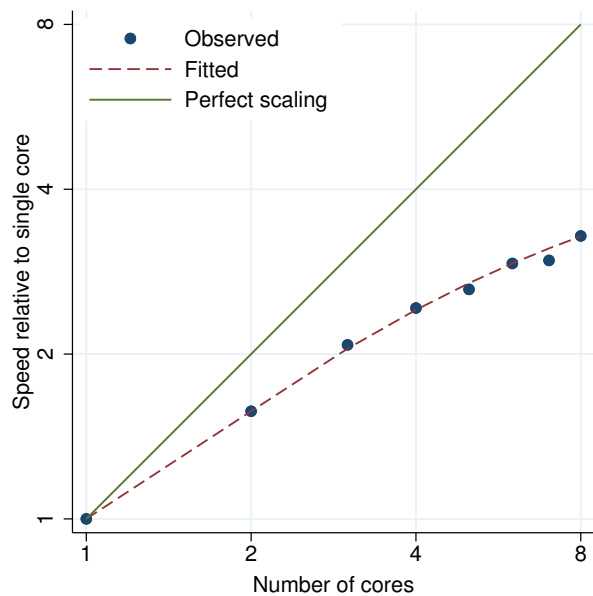
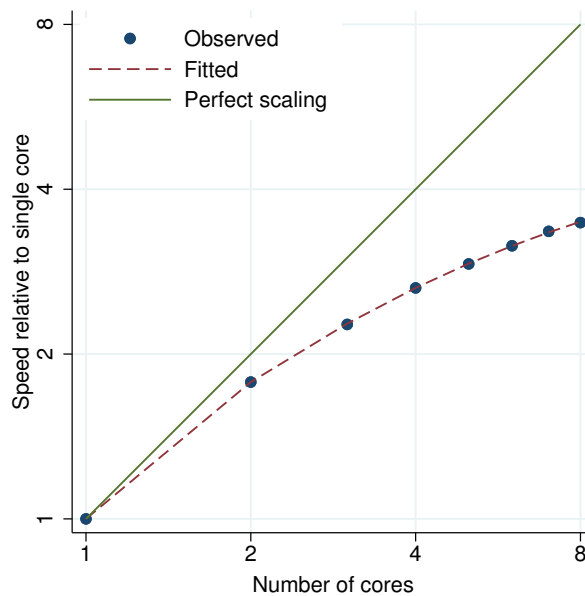
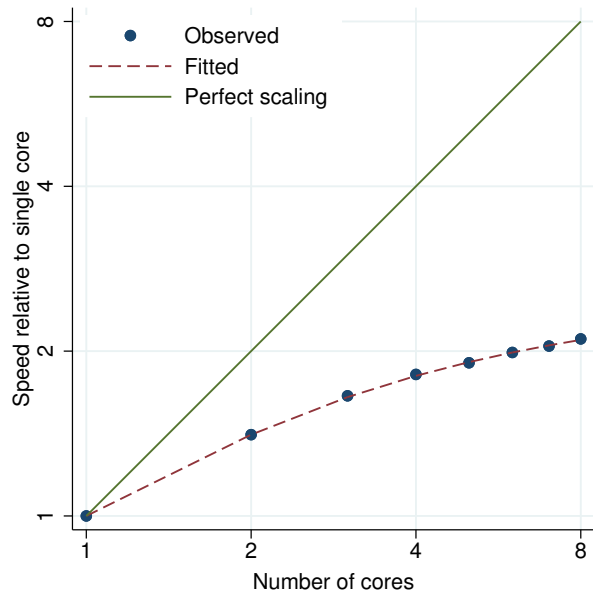
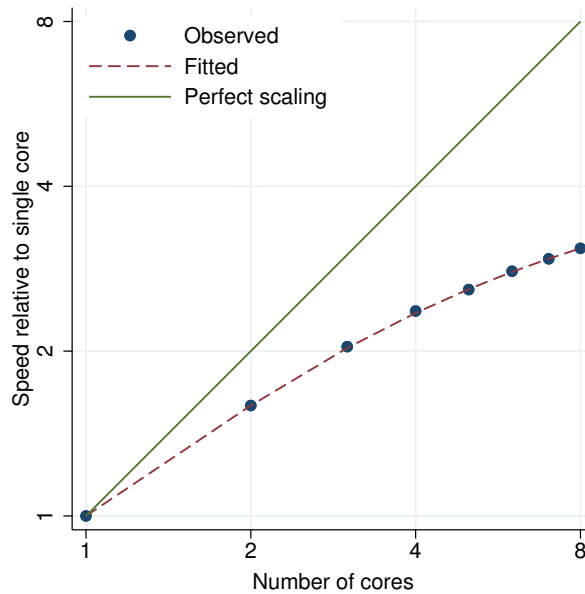
Figure 108. `discrim logistic` performance plot.Figure 109. `discrim qda` performance plot.Figure 110. `dotplot` performance plot.Figure 111. `drawnorm` performance plot.

Figure 112. `drop if exp` performance plot.Figure 113. `drop in range` performance plot.Figure 114. `dsge` performance plot.Figure 115. `dsge1` performance plot.

Figure 116. `dslogit` performance plot.Figure 117. `dspoisson` performance plot.Figure 118. `dsregress` performance plot.Figure 119. `dstdize` performance plot.

Figure 120. `dvehc` performance plot.Figure 121. `egen group()` performance plot.Figure 122. `by: egen mean` performance plot.Figure 123. `eivreg` performance plot.



Figure 124. `encode` performance plot.Figure 125. `eregress` performance plot.Figure 126. `esize twosample` performance plot.Figure 127. `esize unpaired` performance plot.

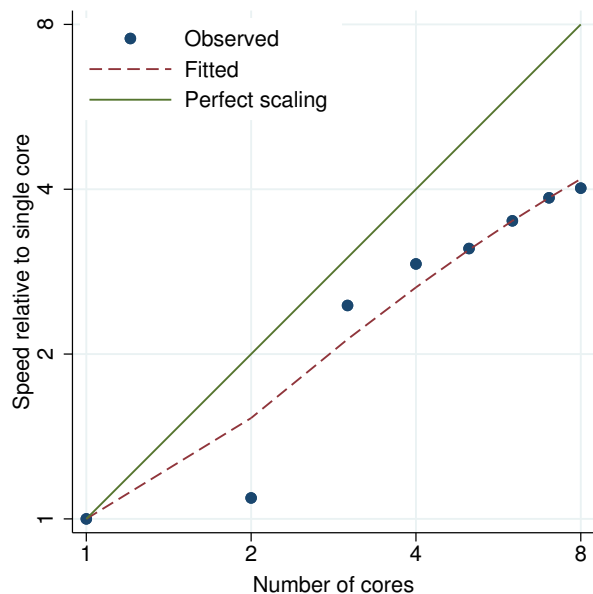


Figure 128. `eteffects` (exponential), `ate` performance plot.

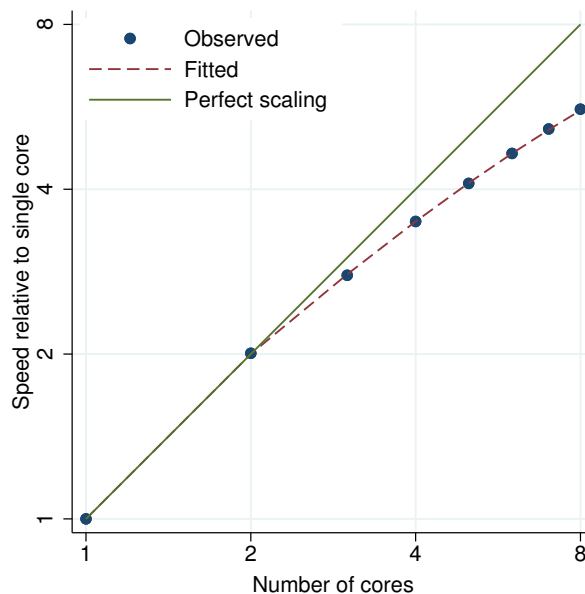


Figure 129. `eteffects` (linear), `ate` performance plot.

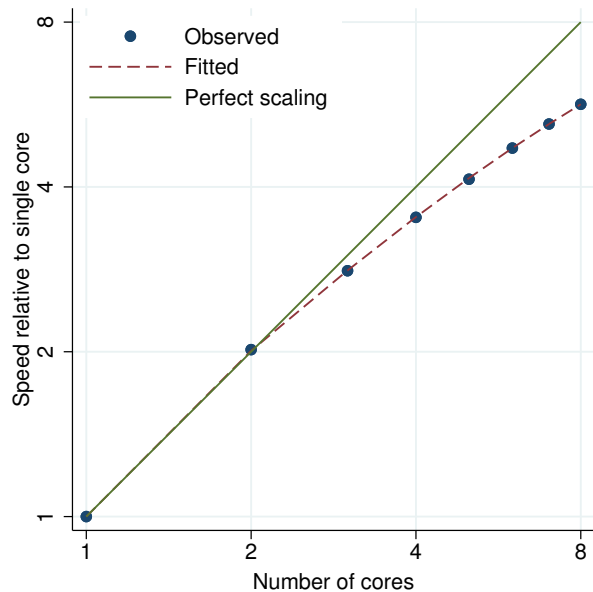


Figure 130. `eteffects` (linear), `pomeans` performance plot.

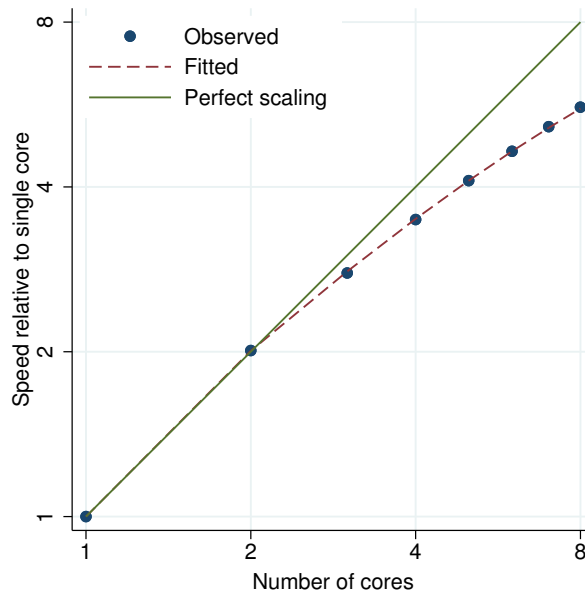
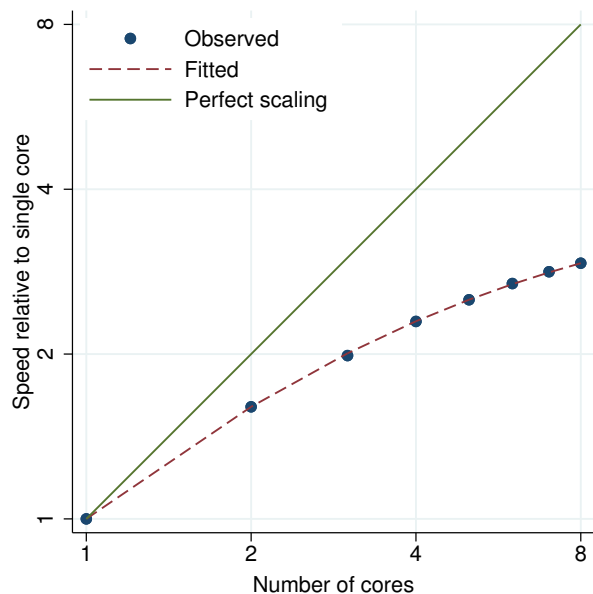
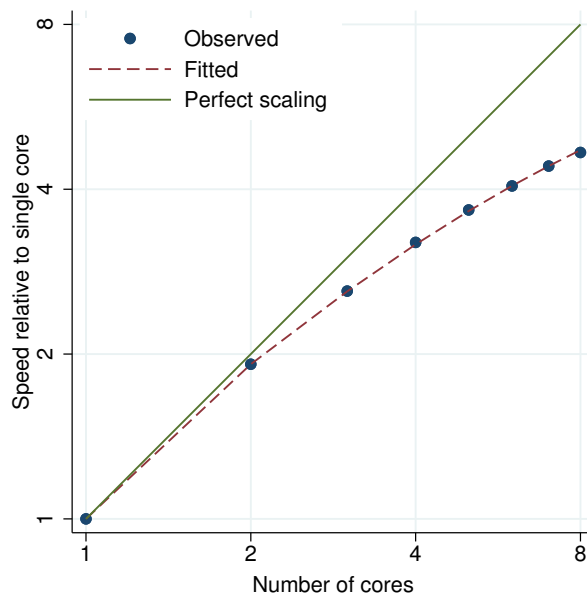
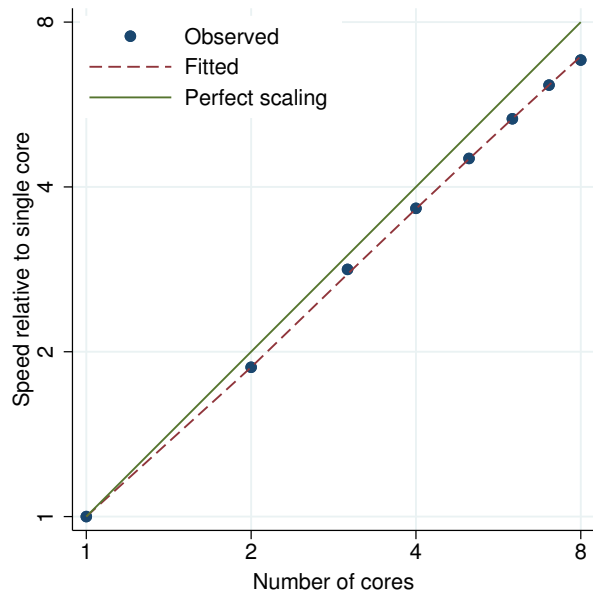
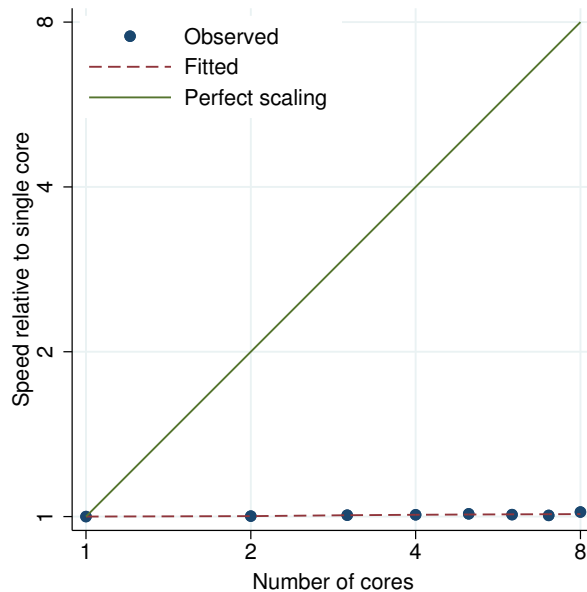
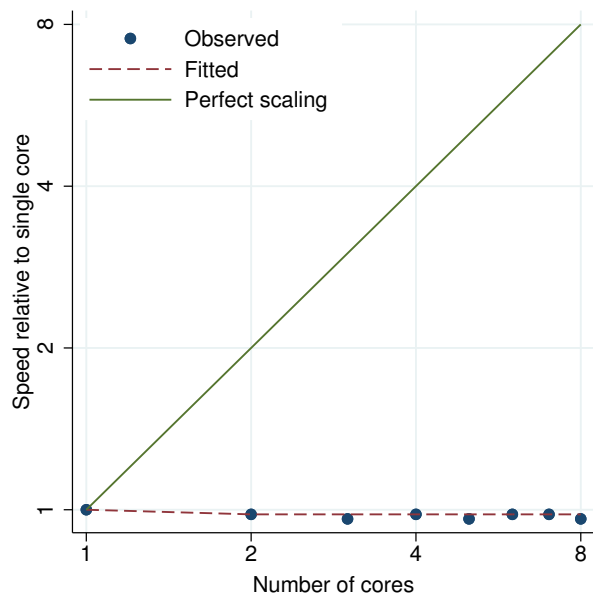
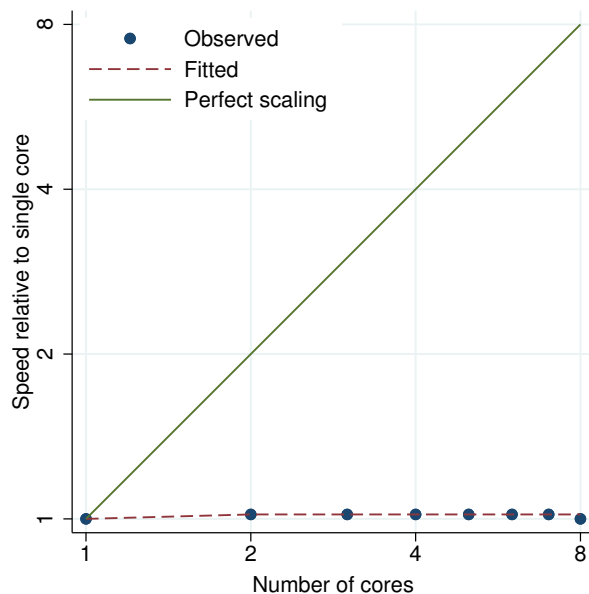
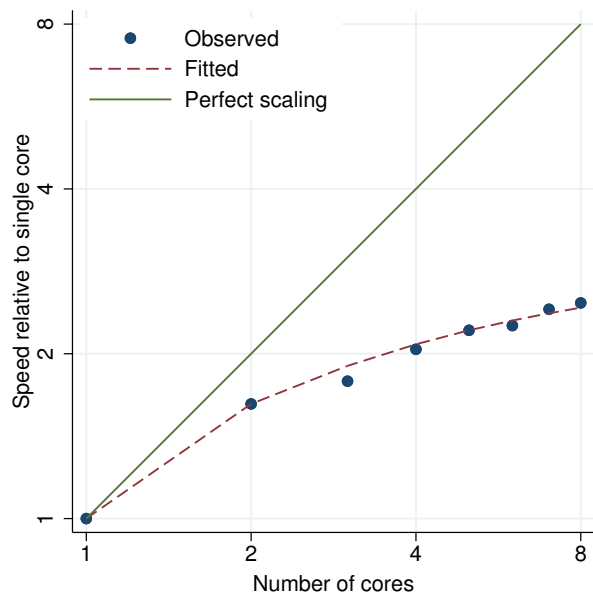
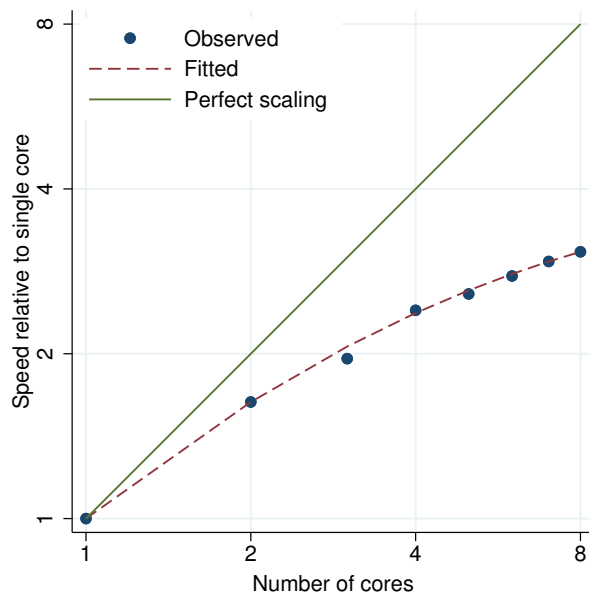
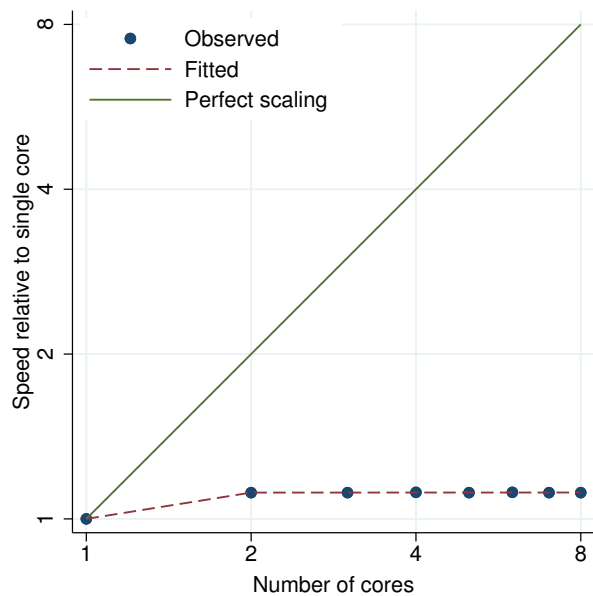
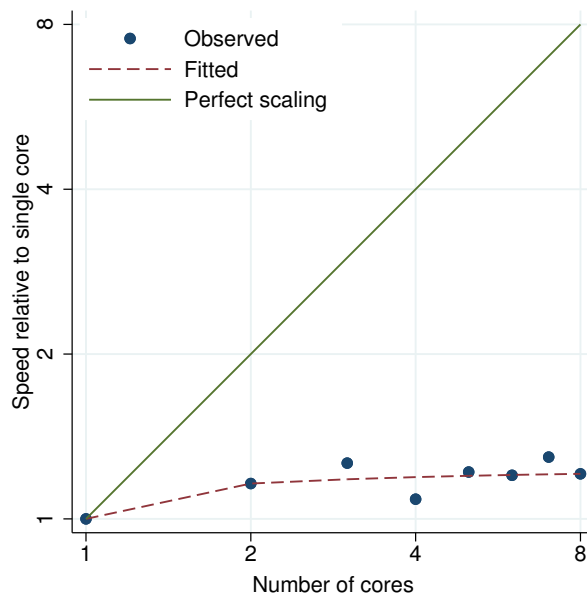
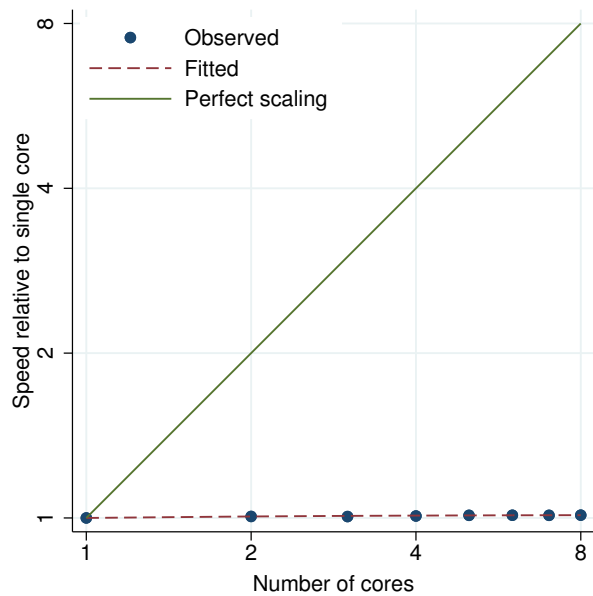
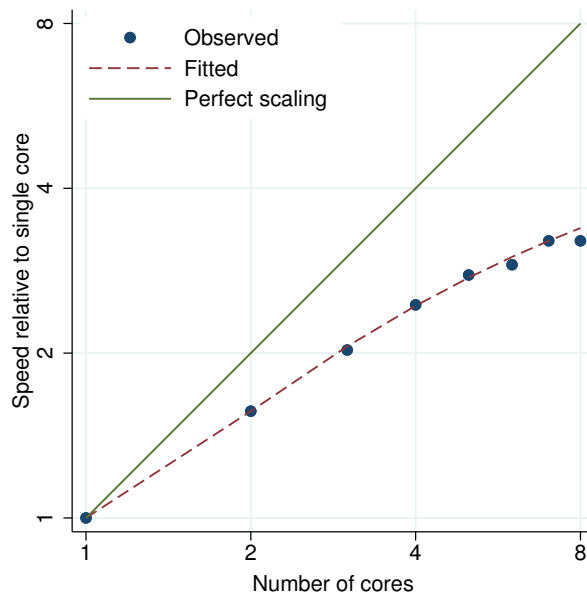


Figure 131. `eteffects` (probit), `ate` performance plot.

Figure 132. `etpoisson` performance plot.Figure 133. `etregress, poutcomes` performance plot.Figure 134. `etregress, twostep` performance plot.Figure 135. `exlogistic` performance plot.

Figure 136. `expand #` performance plot.Figure 137. `expand varname` performance plot.Figure 138. `expandc1 #` performance plot.Figure 139. `expandc1 varname` performance plot.

Figure 140. `expoisson` performance plot.Figure 141. `factor` performance plot.Figure 142. `fcast compute` performance plot.Figure 143. `fillin` performance plot.

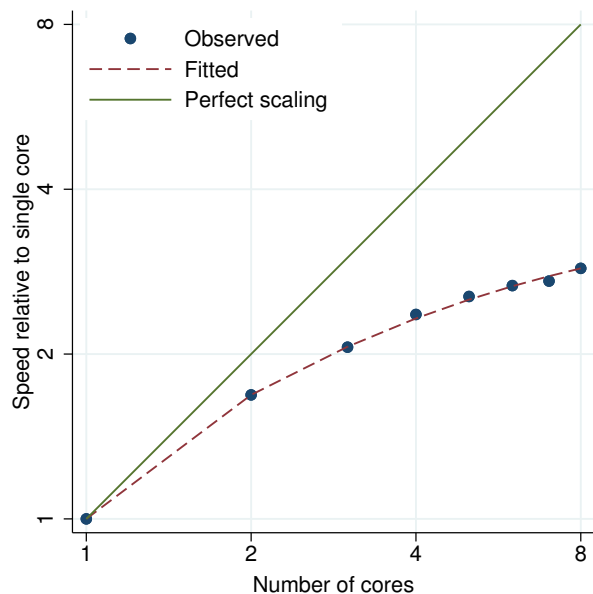


Figure 144. fmm 2: poisson performance plot.

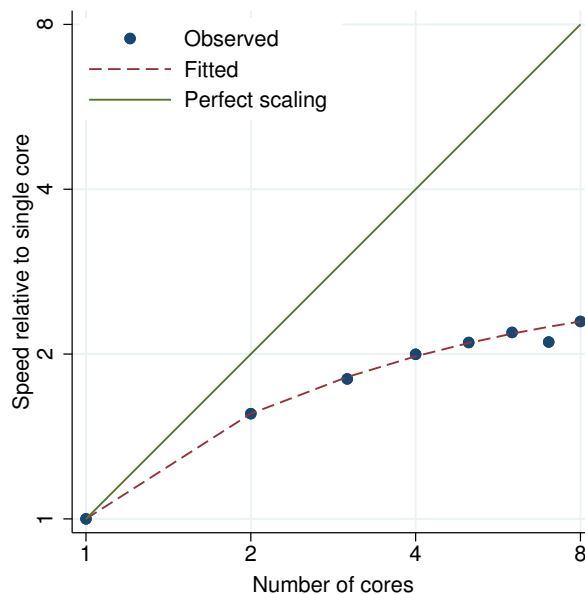


Figure 145. fmm 2: regress performance plot.

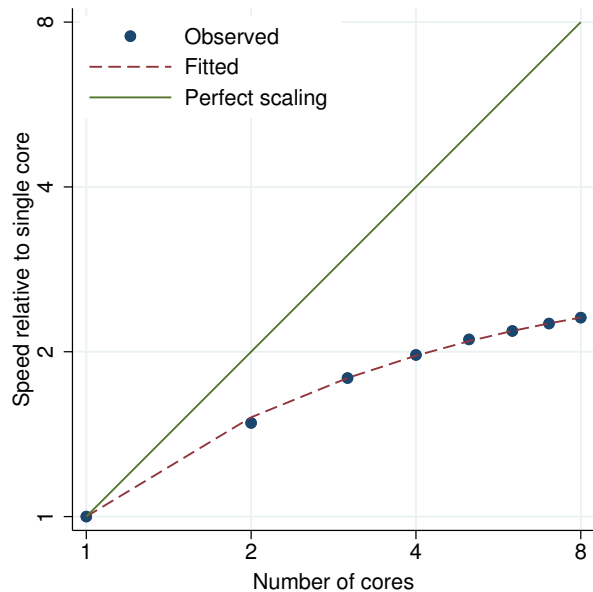


Figure 146. fmm 3: poisson performance plot.

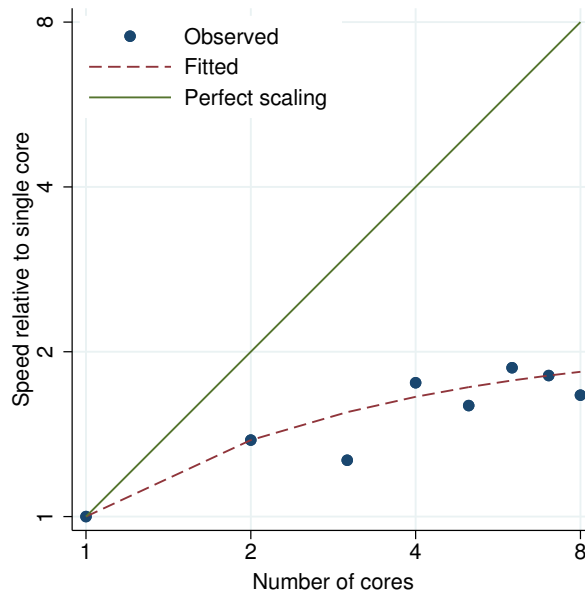
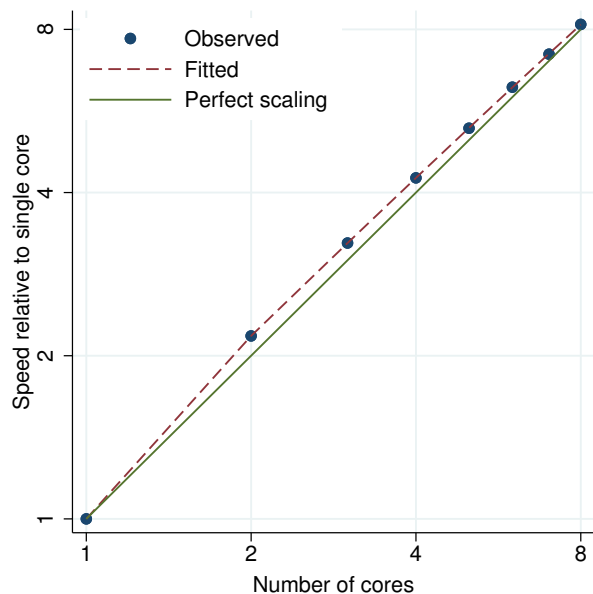
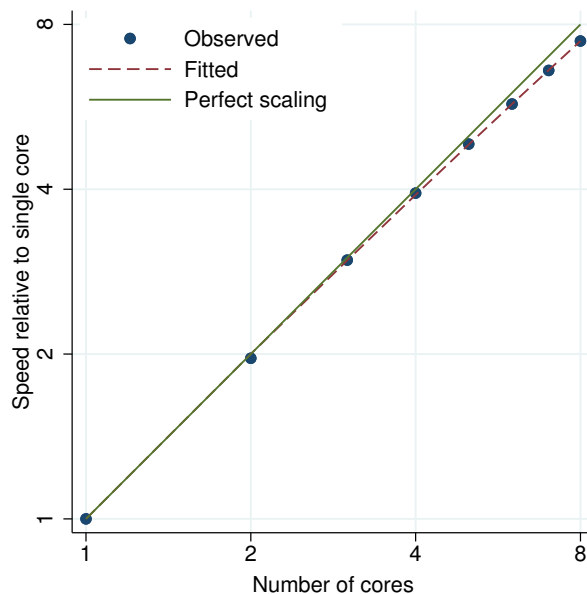
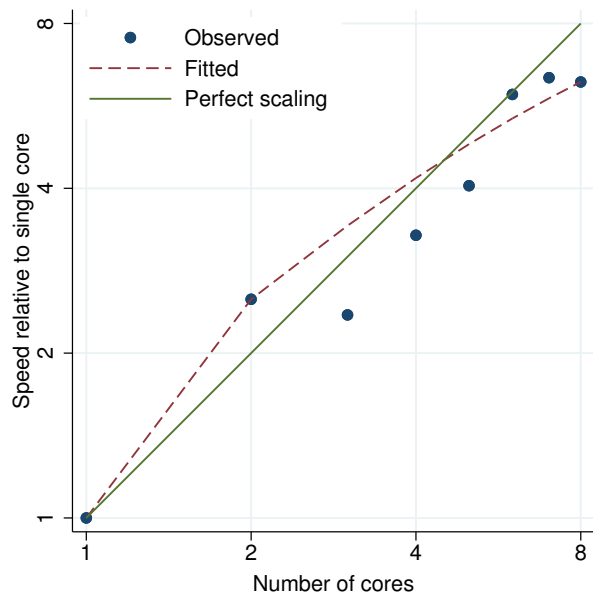
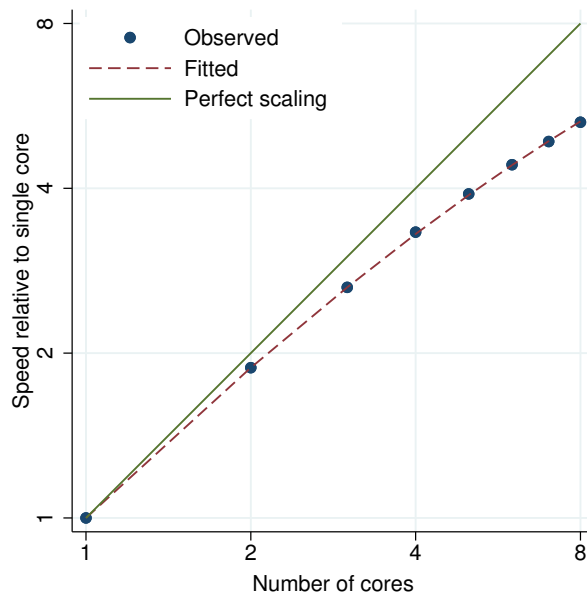


Figure 147. fmm 3: regress performance plot.

Figure 148. `fracreg` probit performance plot.Figure 149. `frontier` performance plot.Figure 150. `fvrevar` (factors) performance plot.Figure 151. `fvrevar` (interaction) performance plot.

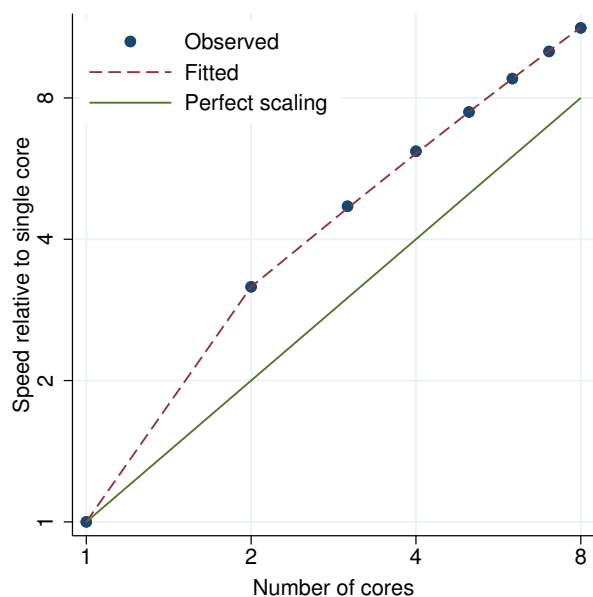


Figure 152. `generate` (small expressions) performance plot.

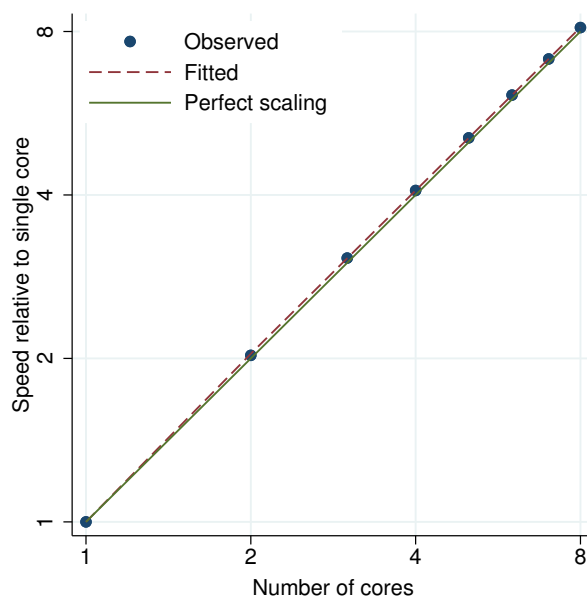


Figure 153. `generate` performance plot.

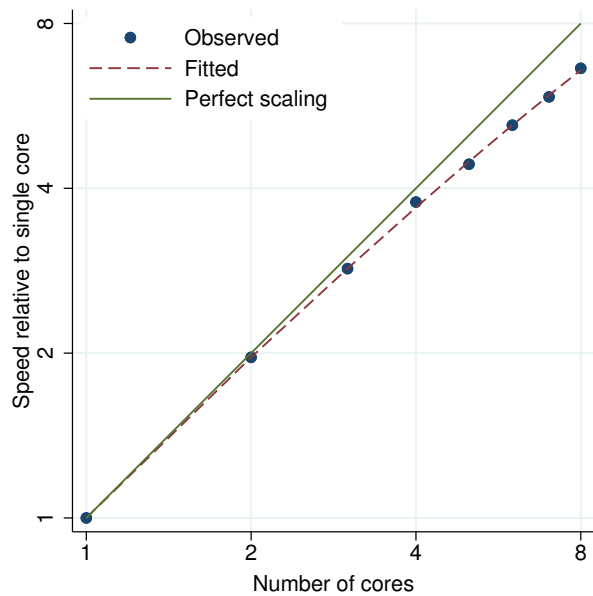


Figure 154. `glm, family(gamma)` performance plot.

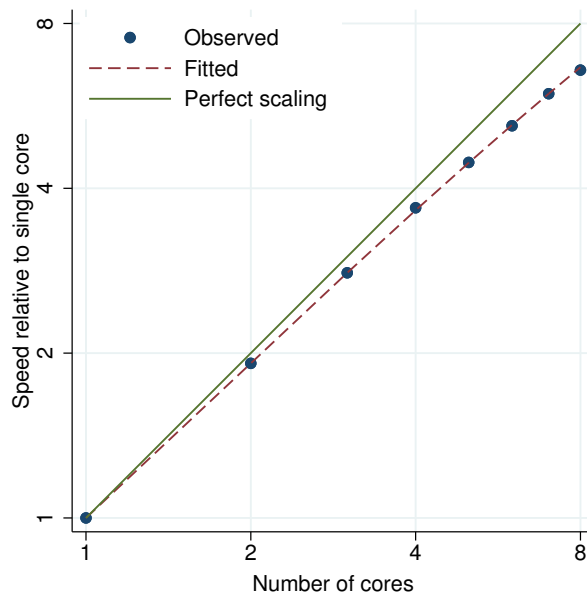


Figure 155. `glm, family(gaussian)` performance plot.



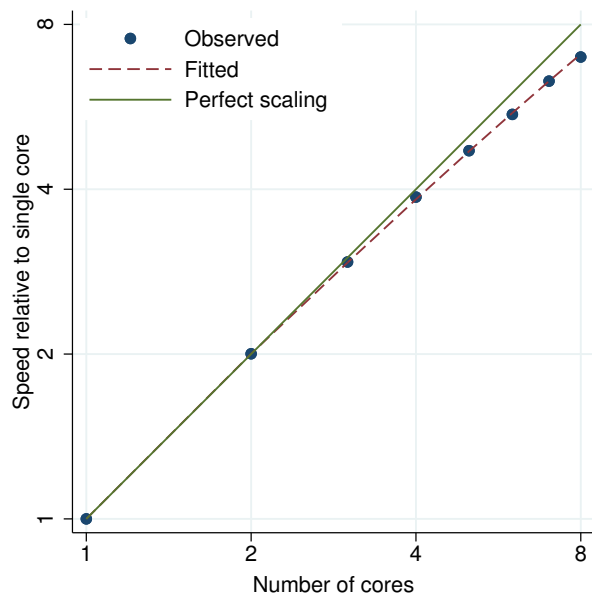


Figure 156. `glm, family(igaussian)` performance plot.

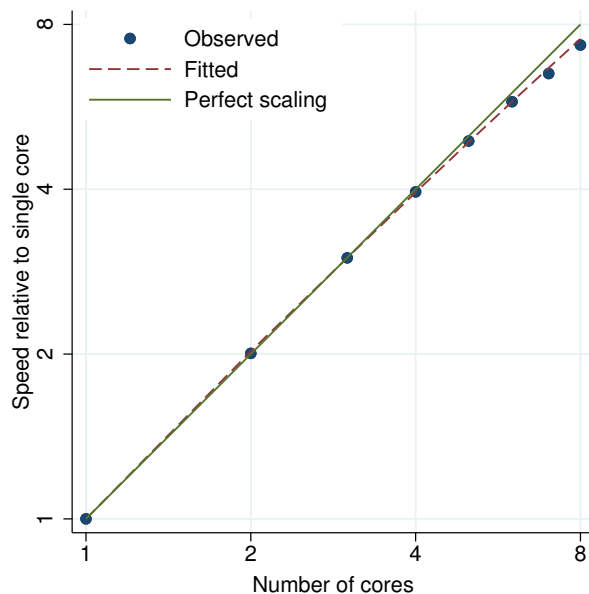


Figure 157. `glm, family(nbinomial)` performance plot.

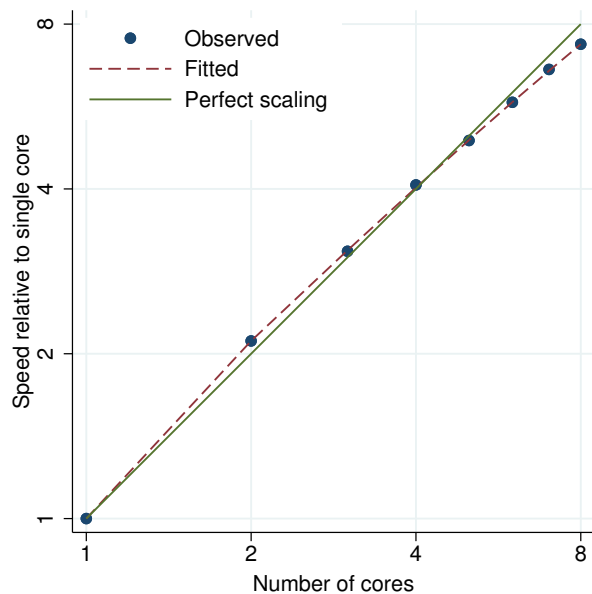


Figure 158. `glm, family(poisson)` performance plot.

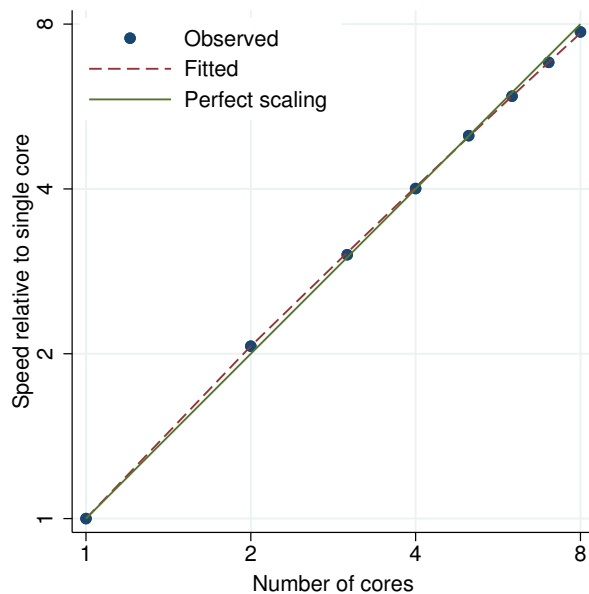
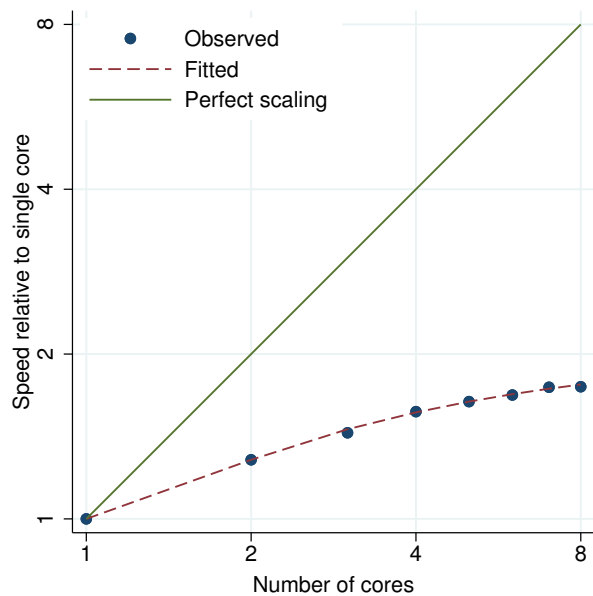
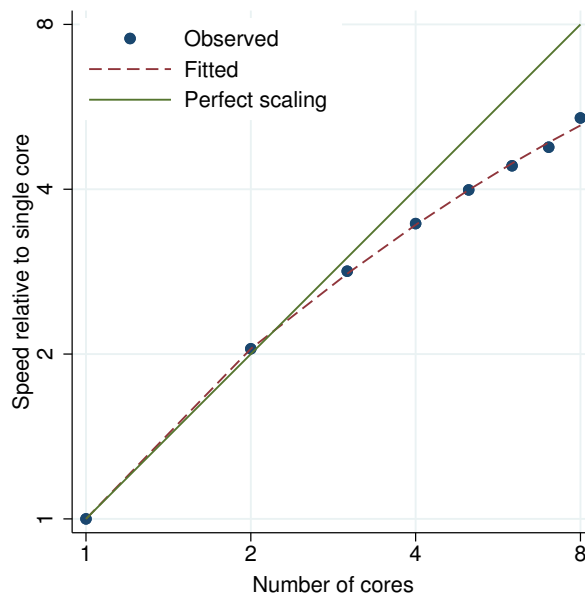
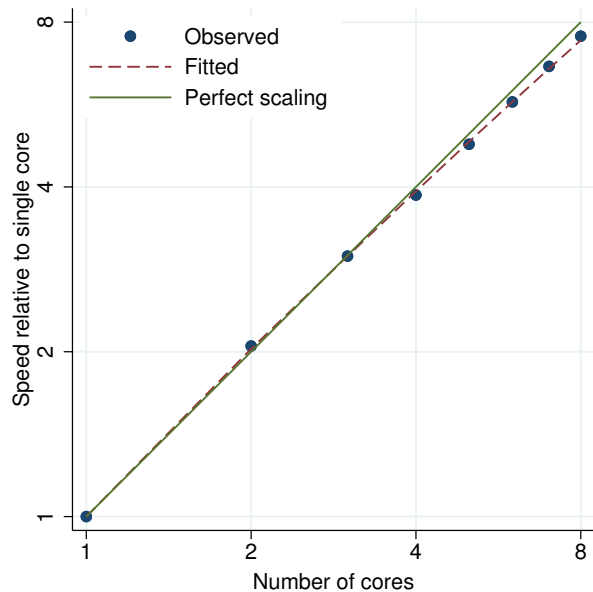
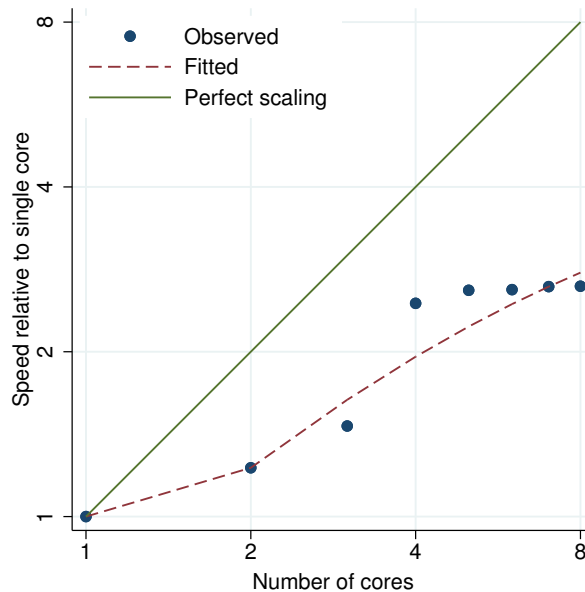


Figure 159. `glogit` performance plot.

Figure 160. `gmm` performance plot.Figure 161. `gmm` (with derivatives) performance plot.Figure 162. `gprobit` performance plot.Figure 163. `graph bar` performance plot.

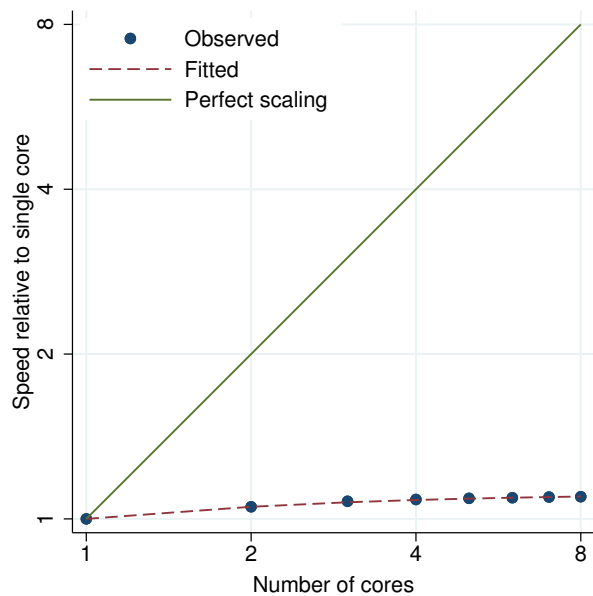


Figure 164. graph box performance plot.

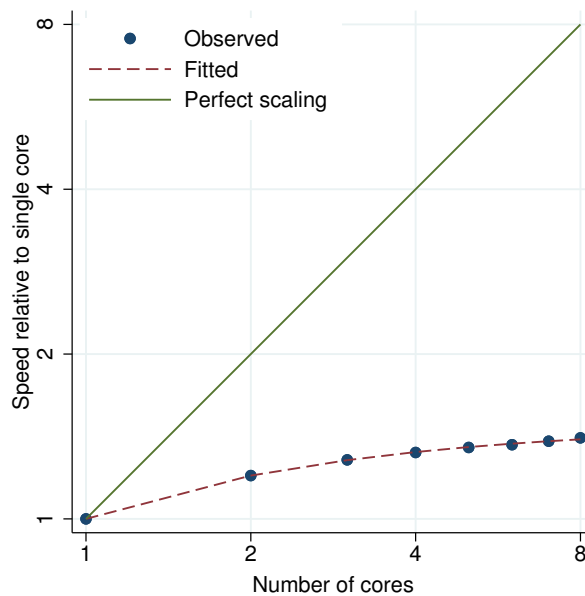


Figure 165. graph pie performance plot.

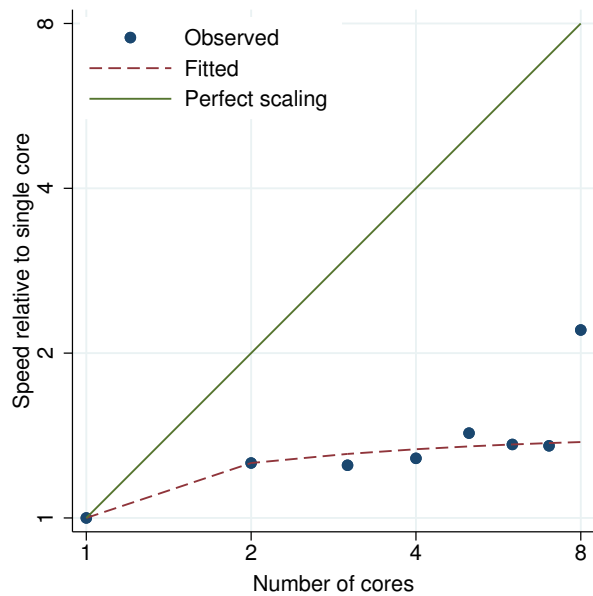


Figure 166. grmeanby performance plot.

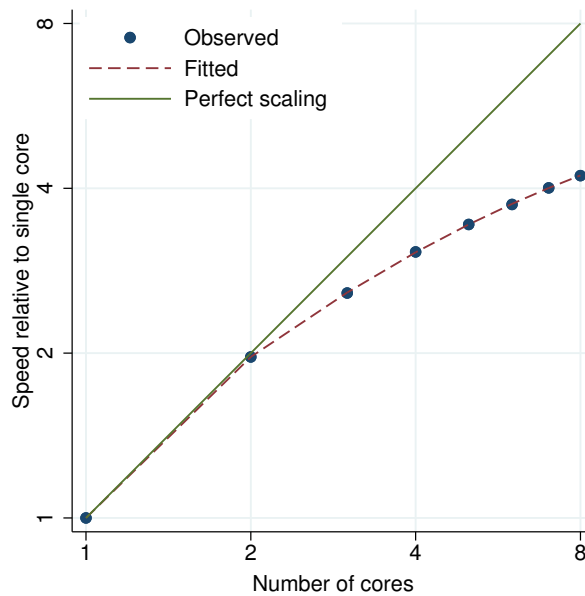


Figure 167. gsem, oprobit (CFA, 2-level) performance plot.

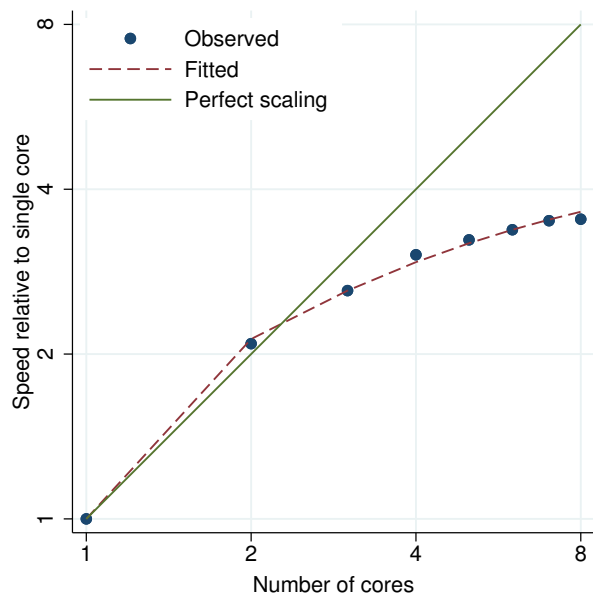


Figure 168. `gsem, oprobit (CFA)` performance plot.

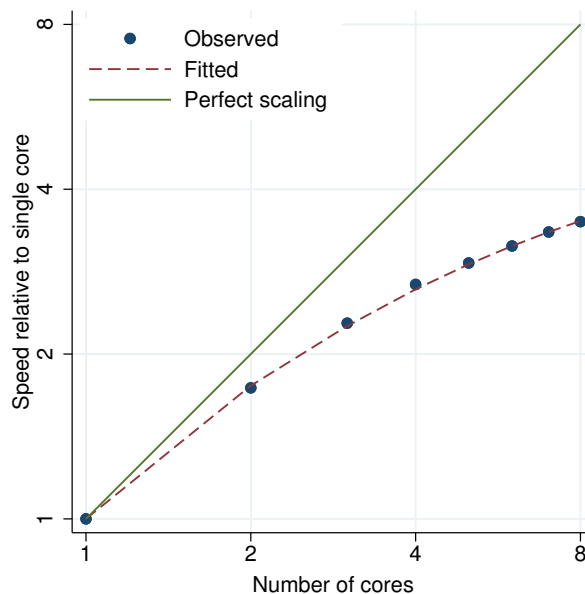


Figure 169. `gsem, logit group()` performance plot.

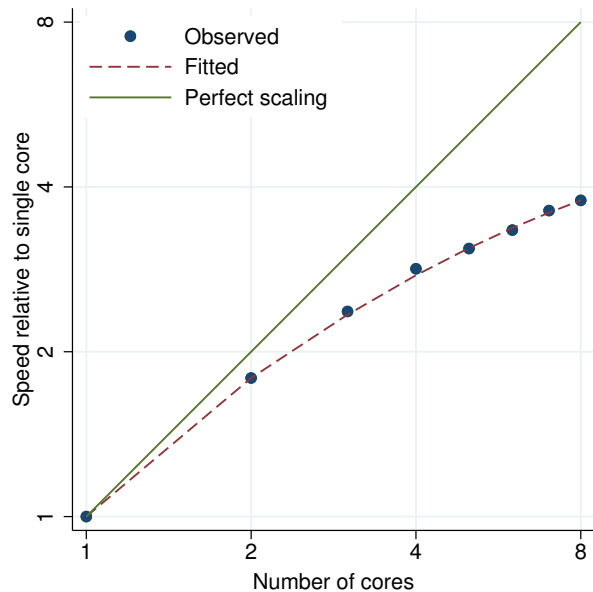


Figure 170. `gsem, group()` performance plot.

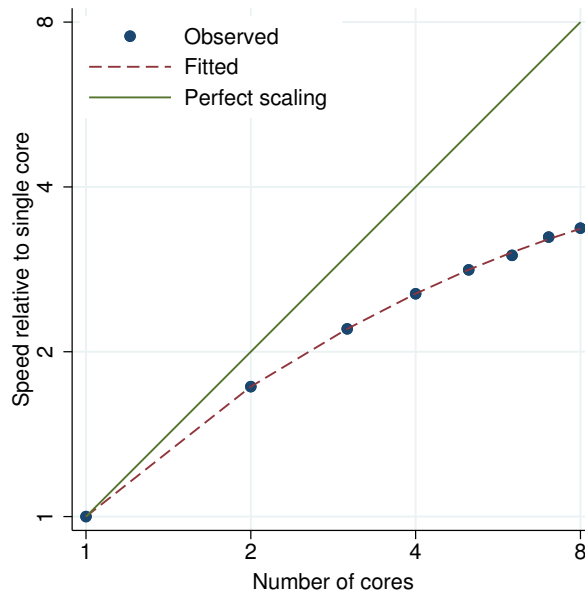
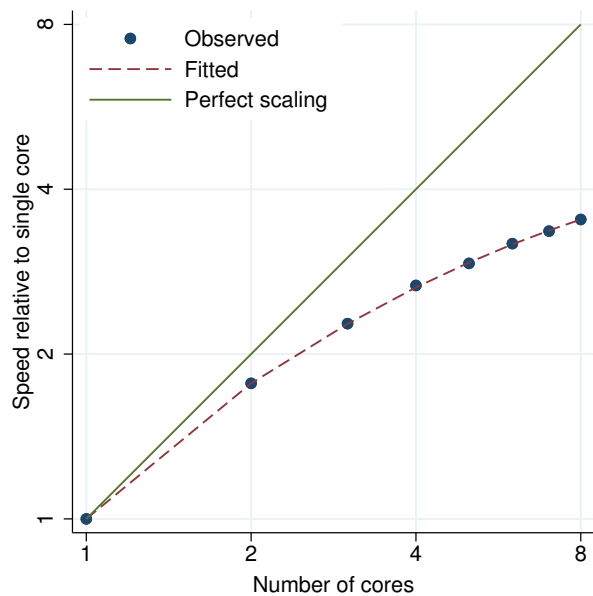
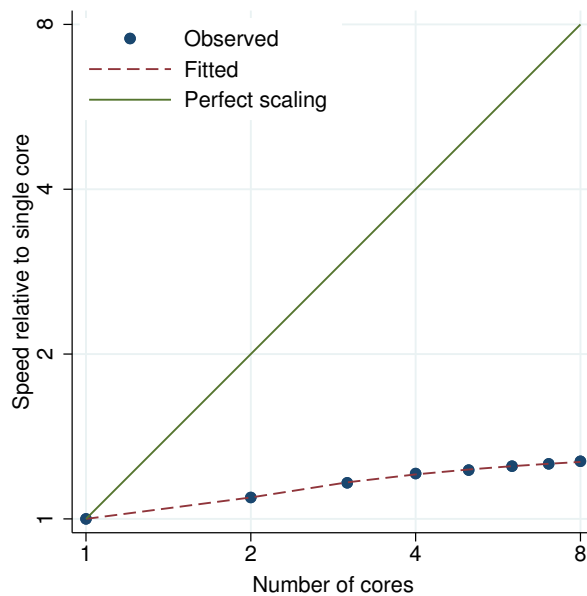
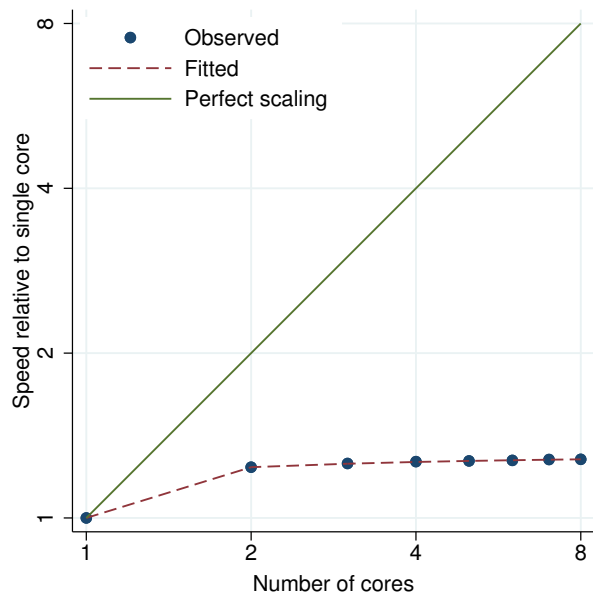
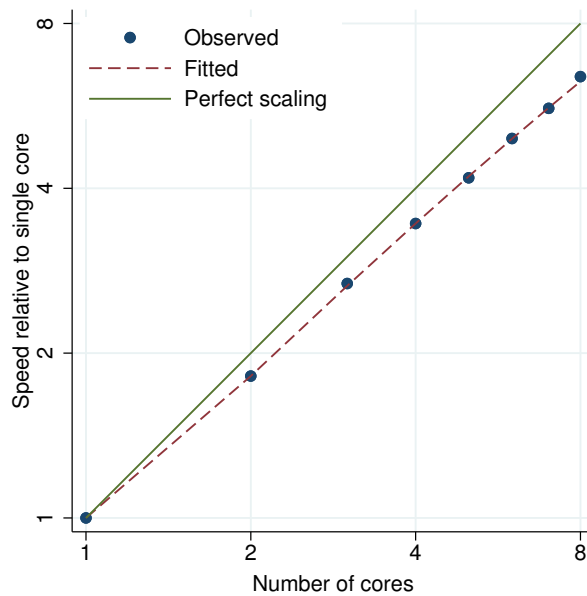


Figure 171. `gsem, ologit group()` performance plot.

Figure 172. `gsem, poisson group()` performance plot.Figure 173. `gsort` performance plot.Figure 174. `hausman` performance plot.Figure 175. `heckman` performance plot.

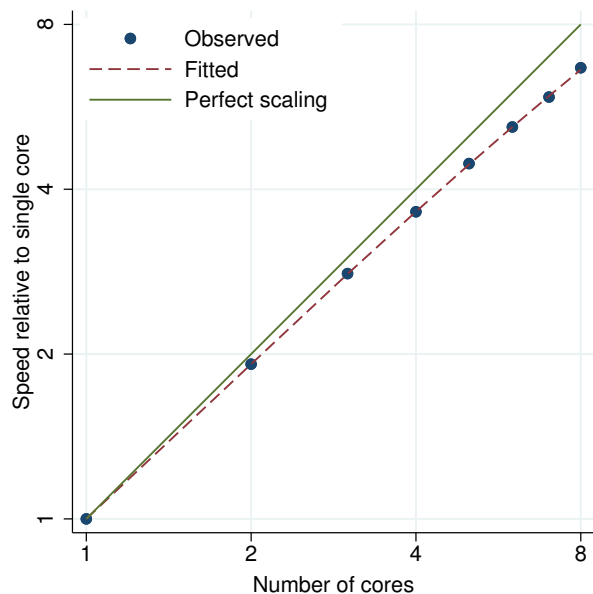


Figure 176. heckman, twostep performance plot.

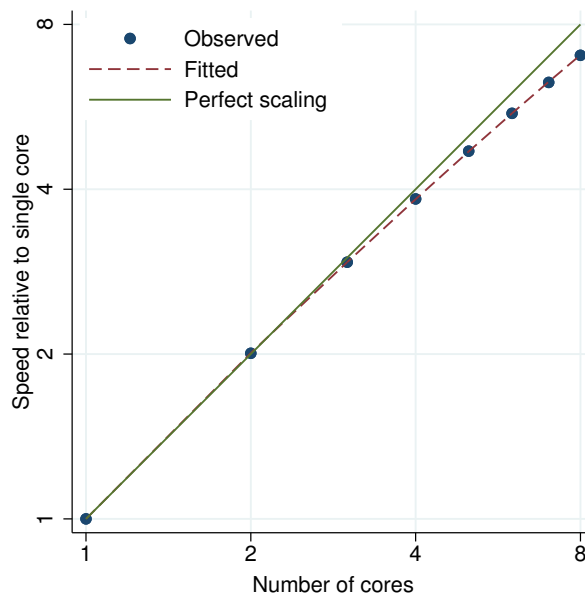


Figure 177. heckprobit performance plot.

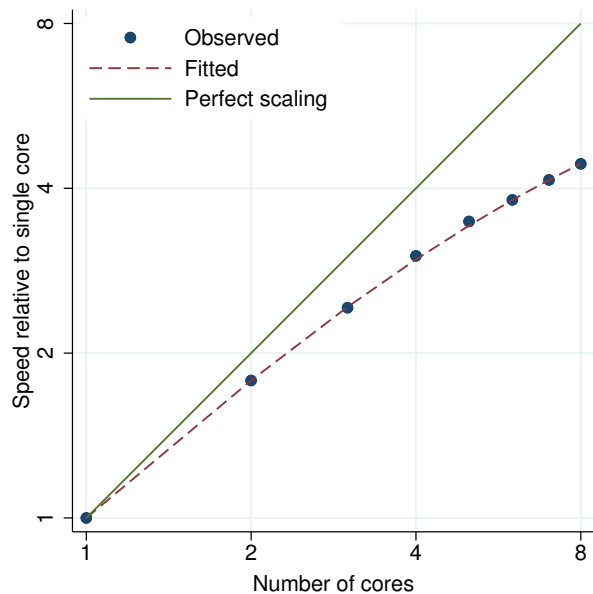


Figure 178. heckpoisson performance plot.

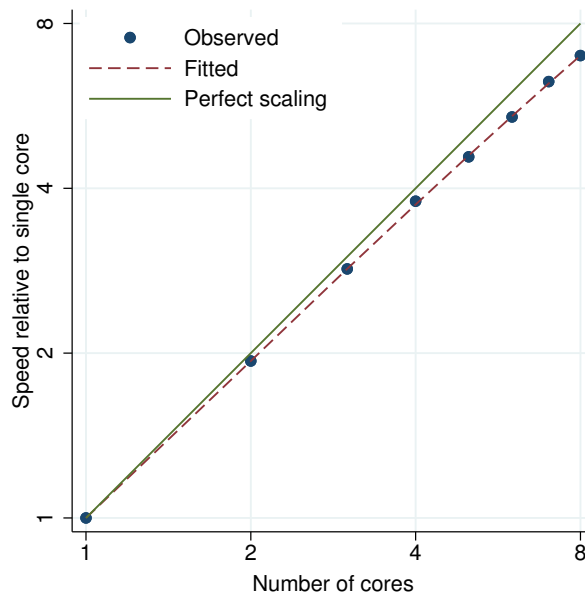
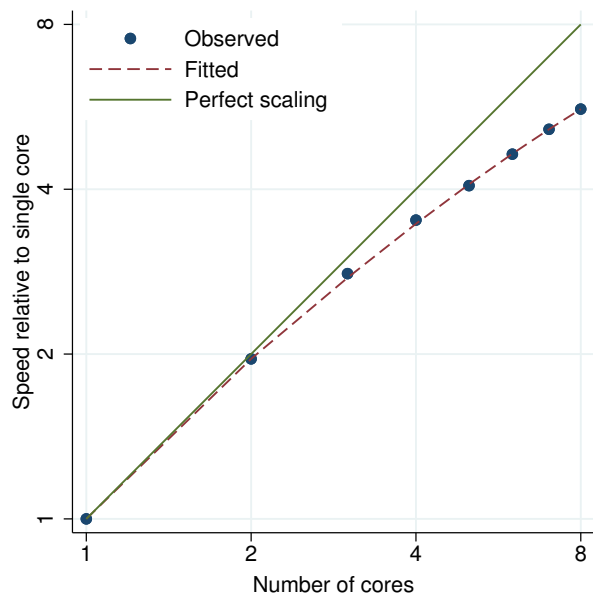
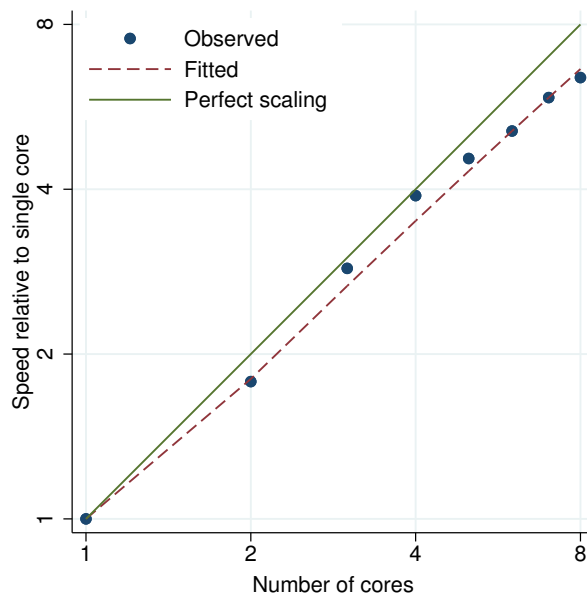
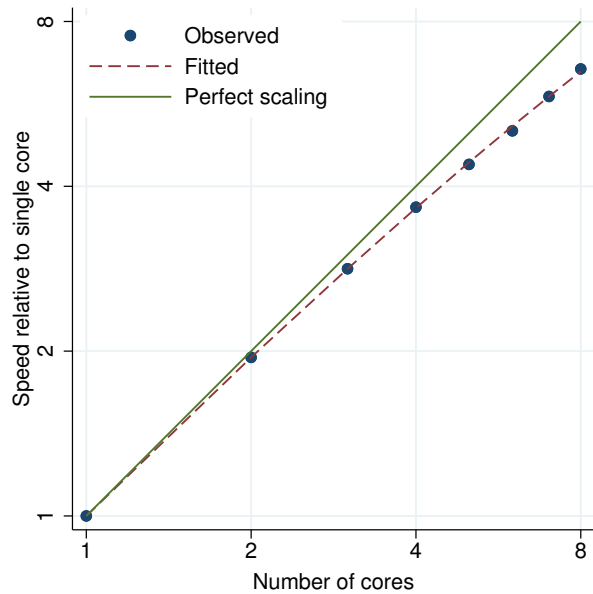
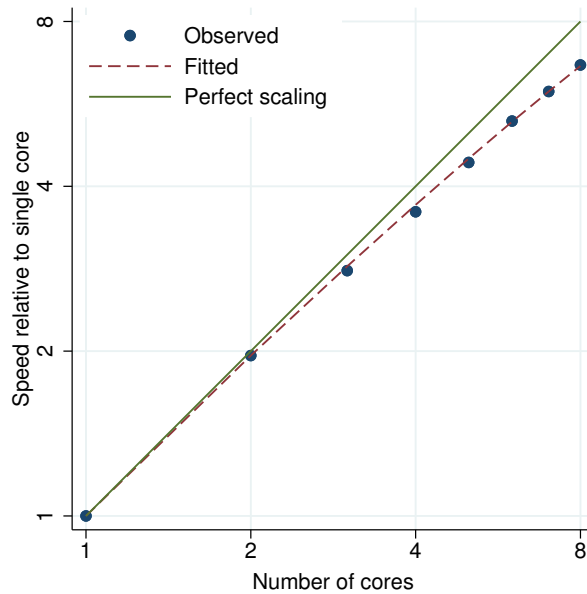


Figure 179. heckprob performance plot.

Figure 180. `hetoprobit` performance plot.Figure 181. `hetprob` performance plot.Figure 182. `hetregress` performance plot.Figure 183. `hetregress, twostep` performance plot.

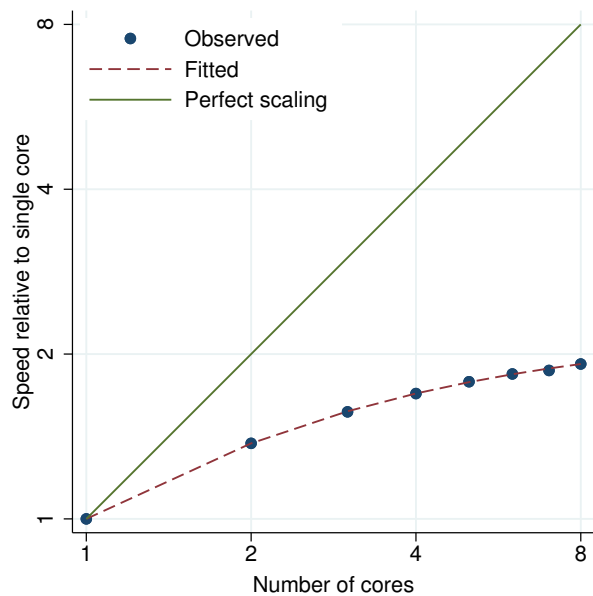


Figure 184. histogram performance plot.

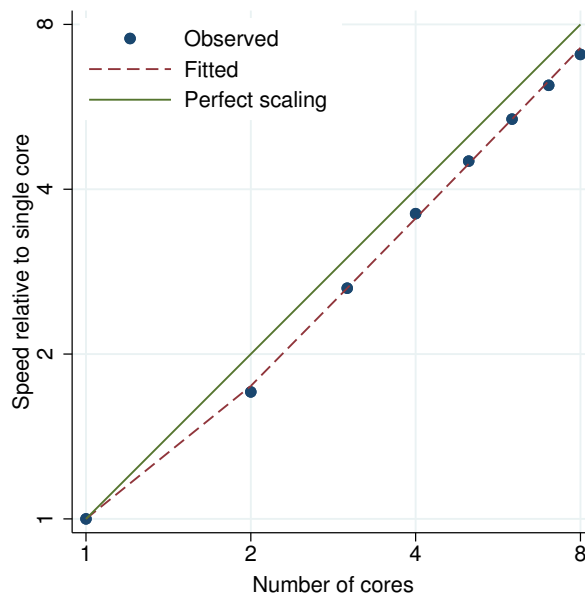


Figure 185. hotellong performance plot.

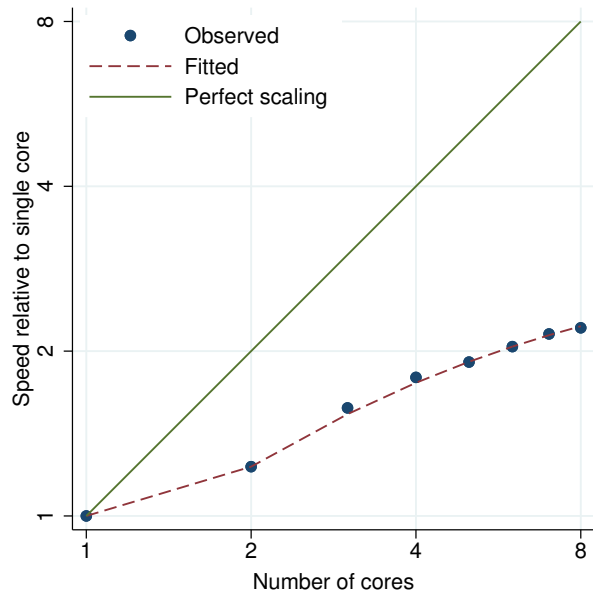


Figure 186. icc, mixed performance plot.

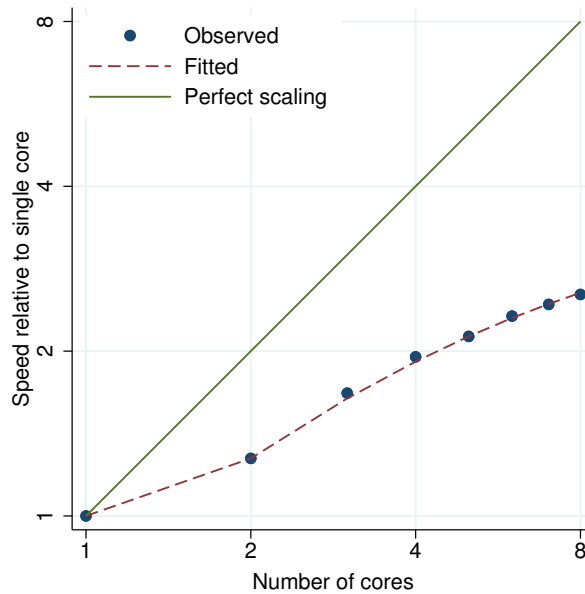
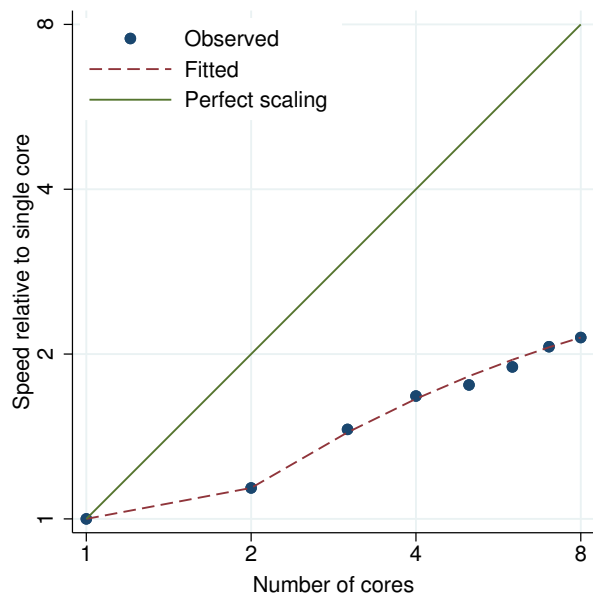
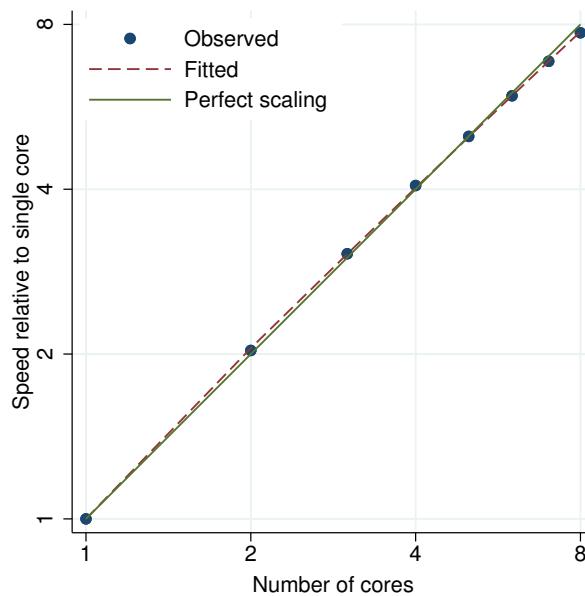
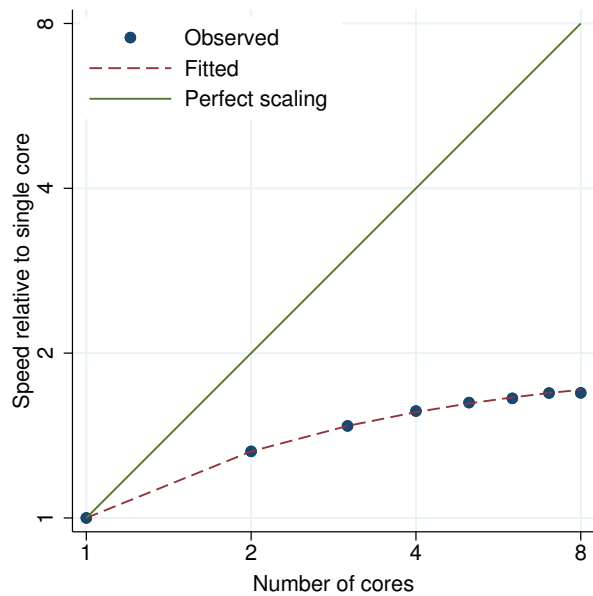
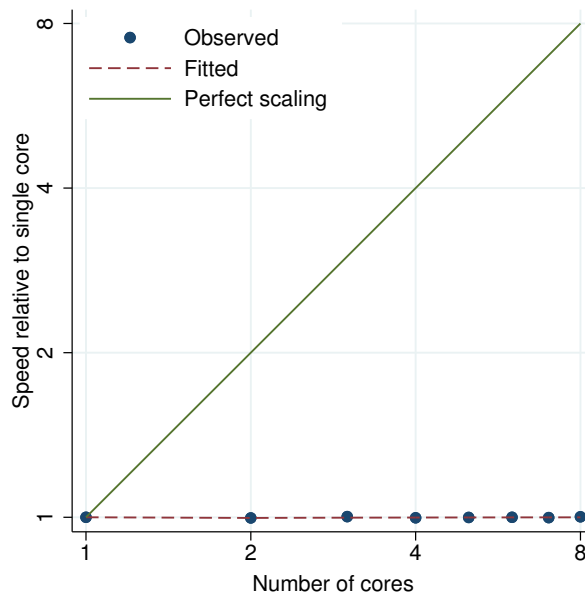
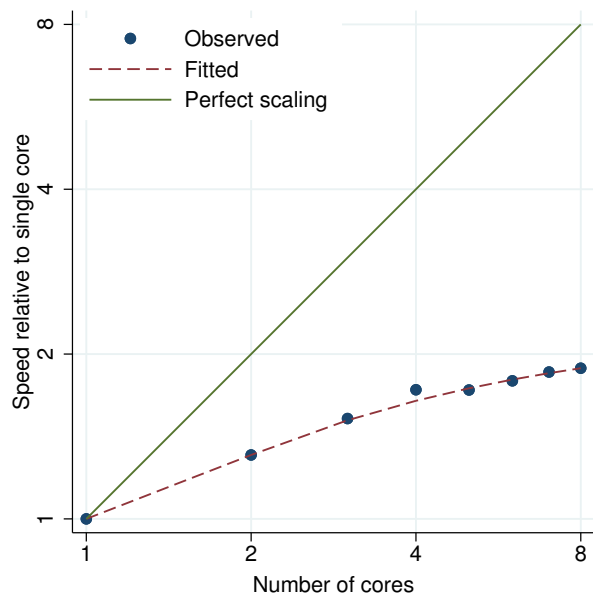
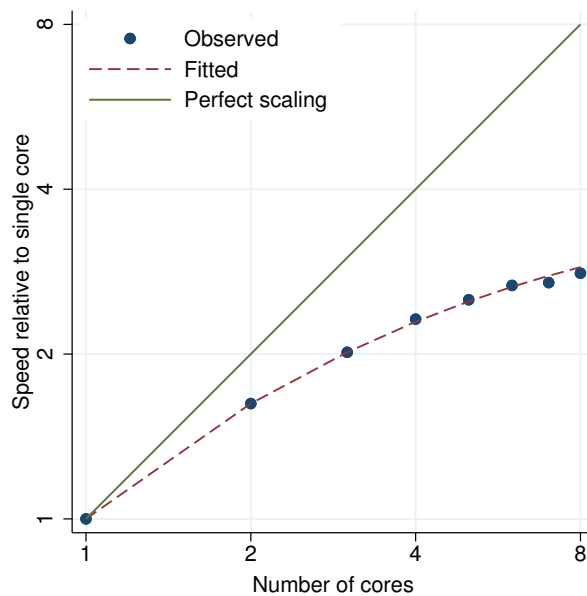
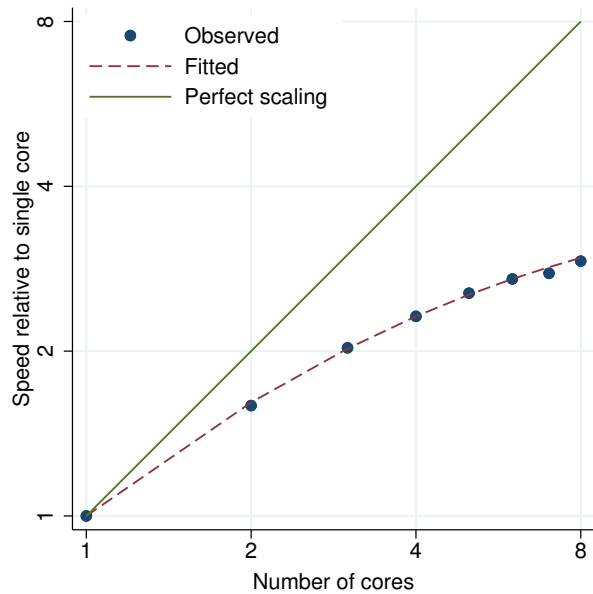
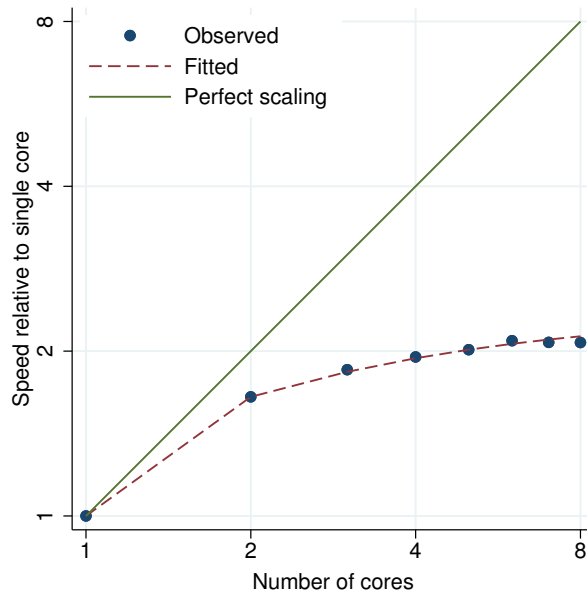
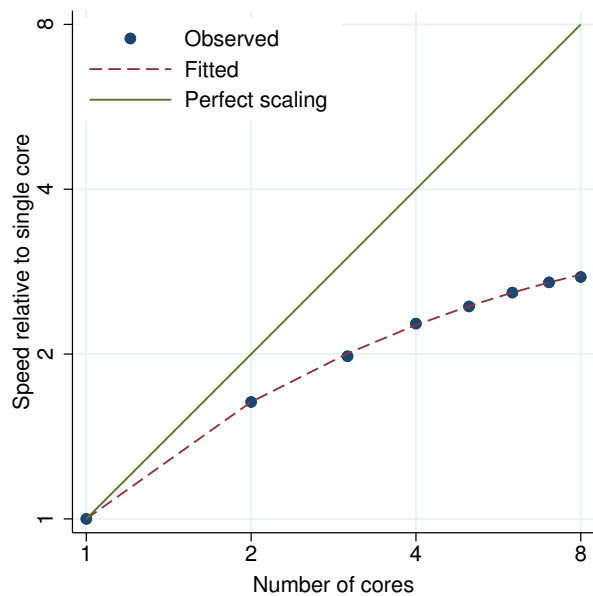
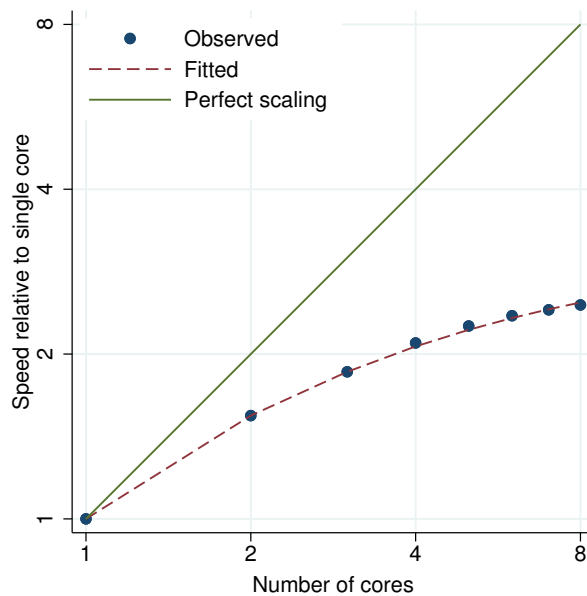
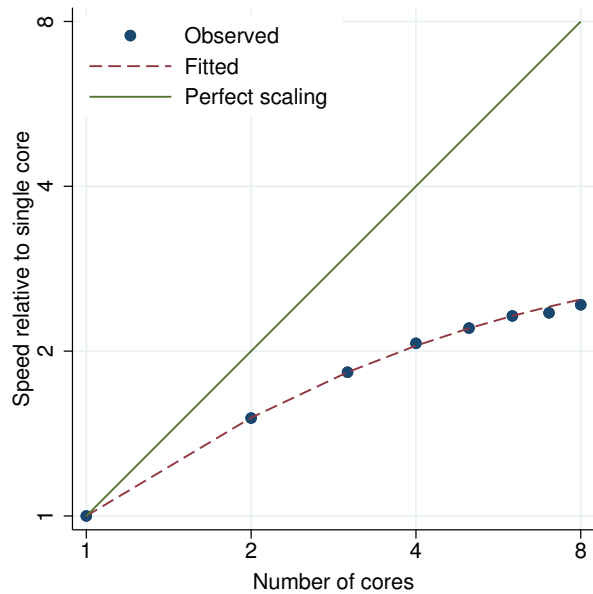
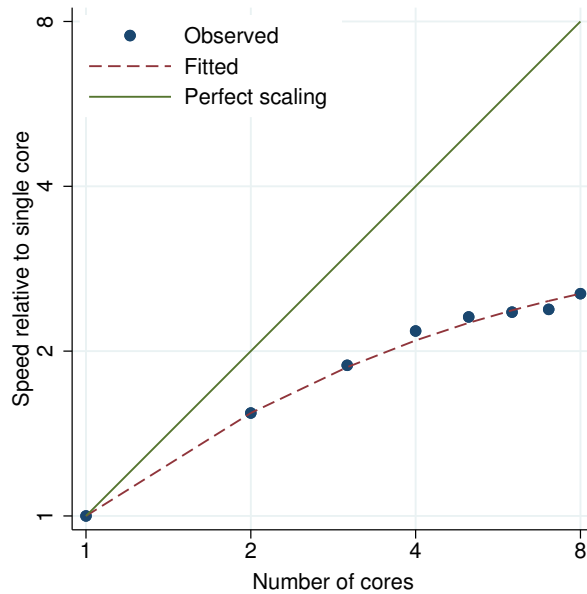


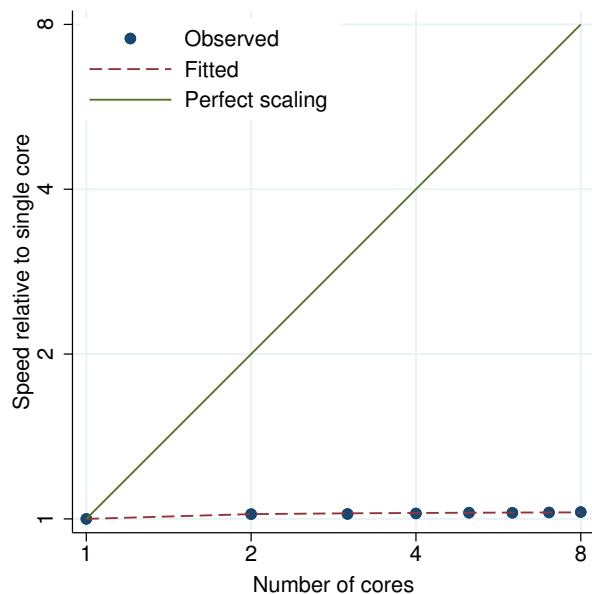
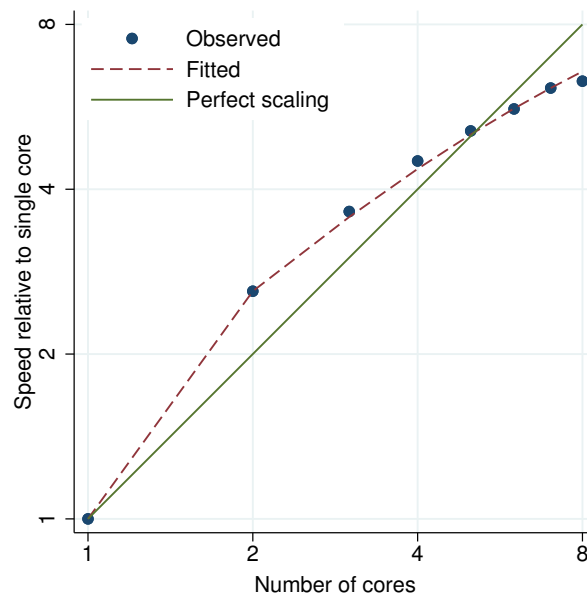
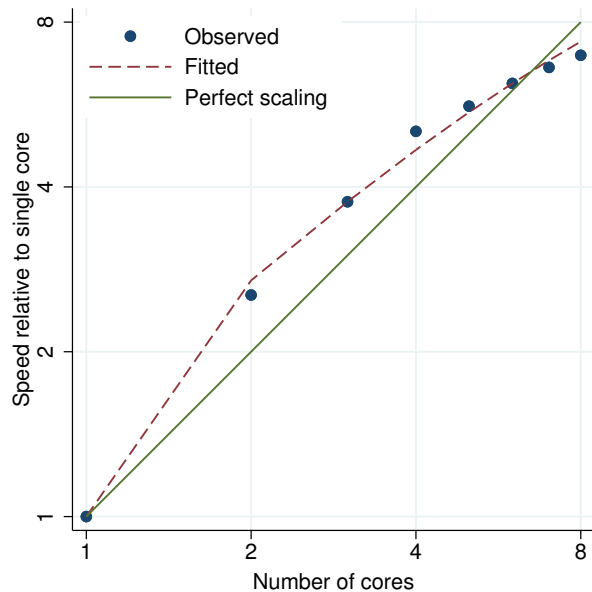
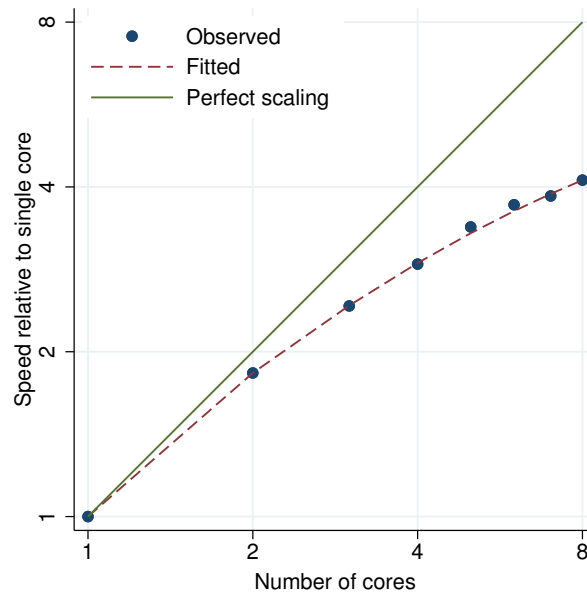
Figure 187. icc (one-way) performance plot.



Figure 188. `icc` (two-way) performance plot.Figure 189. `intreg` performance plot.Figure 190. `ir` performance plot.Figure 191. `by: ir` performance plot.

Figure 192. `irf create` performance plot.Figure 193. `irt 1pl` performance plot.Figure 194. `irt 2pl` performance plot.Figure 195. `irt 3pl` performance plot.

Figure 196. `irt grm` performance plot.Figure 197. `irt nrm` performance plot.Figure 198. `irt pcm` performance plot.Figure 199. `irt rsm` performance plot.

Figure 200. `istdize` performance plot.Figure 201. `ivpoisson cfunction` performance plot.Figure 202. `ivpoisson gmm`, additive performance plot.Figure 203. `ivpoisson gmm`, multiplicative performance plot.

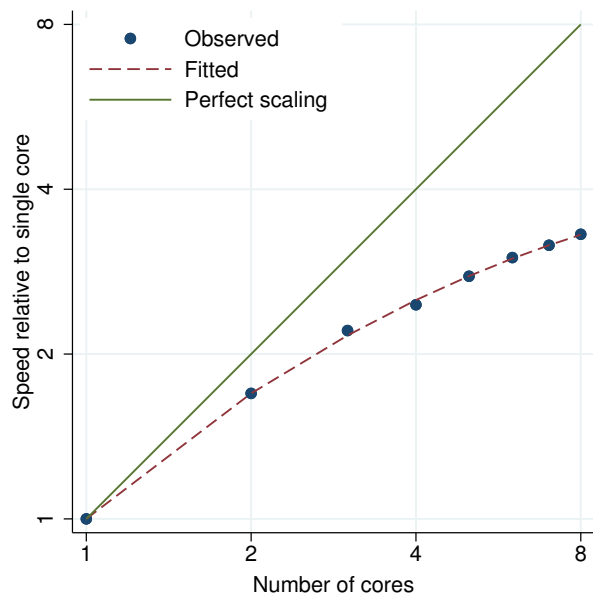


Figure 204. ivprobit performance plot.

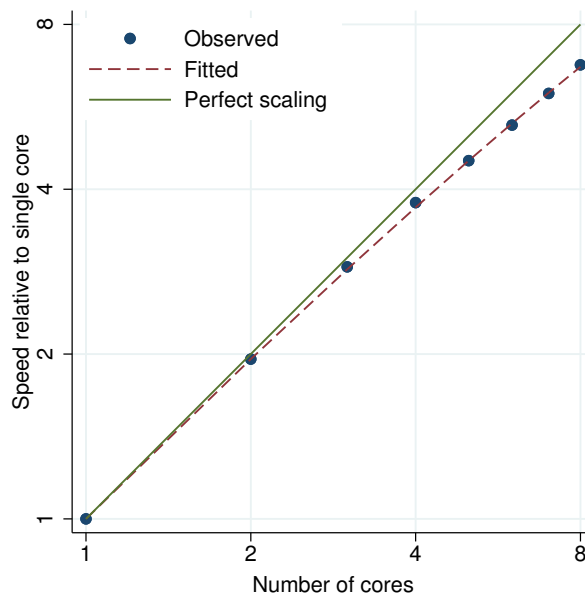


Figure 205. ivregress 2sls performance plot.

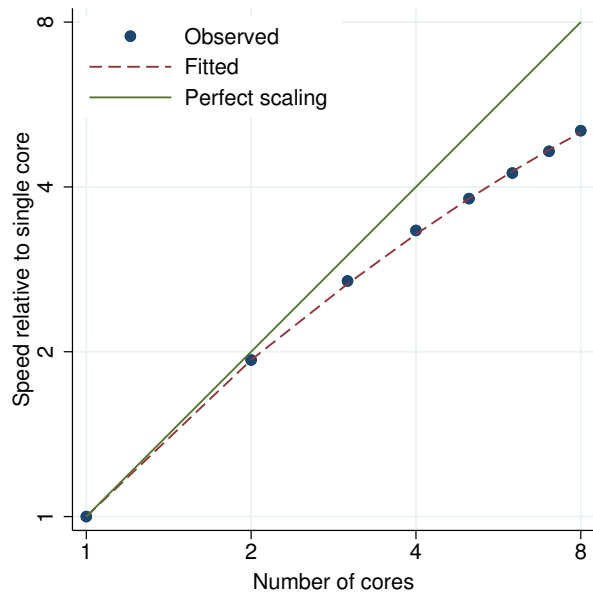


Figure 206. ivregress gmm performance plot.

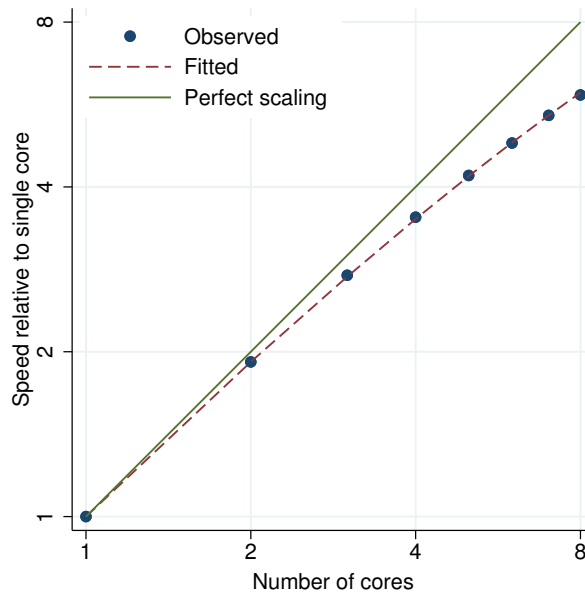
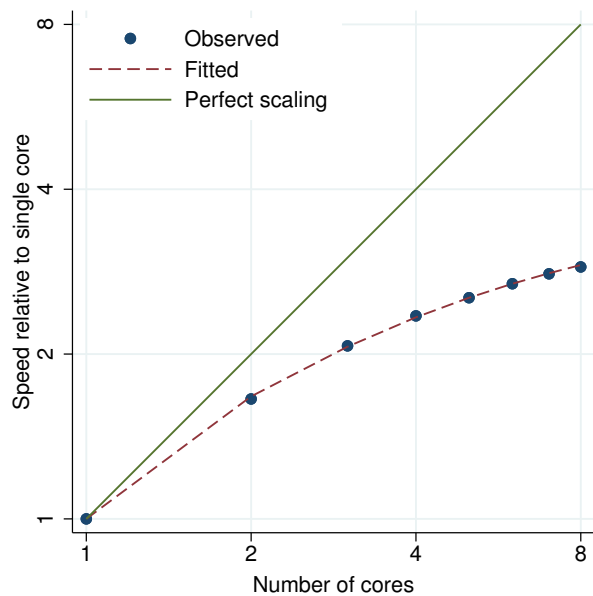
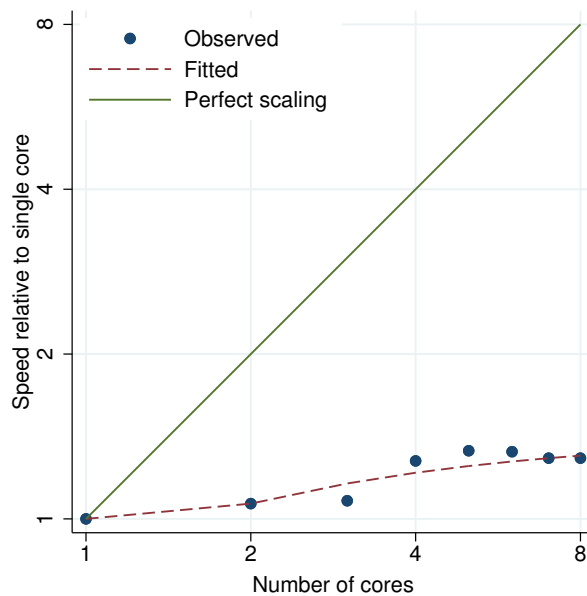
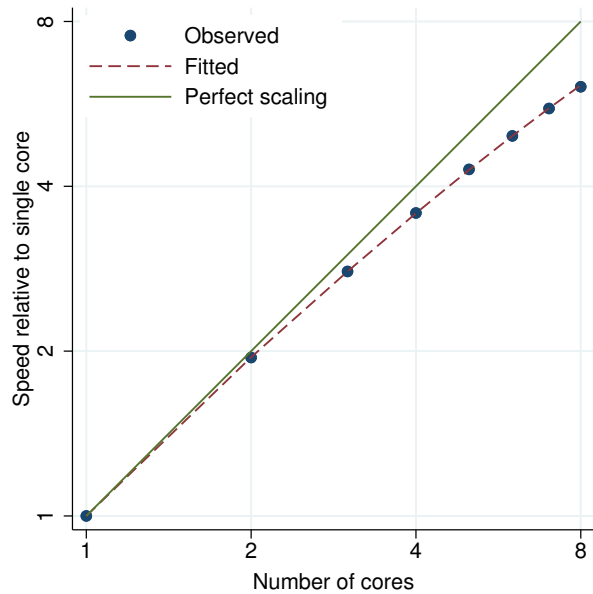
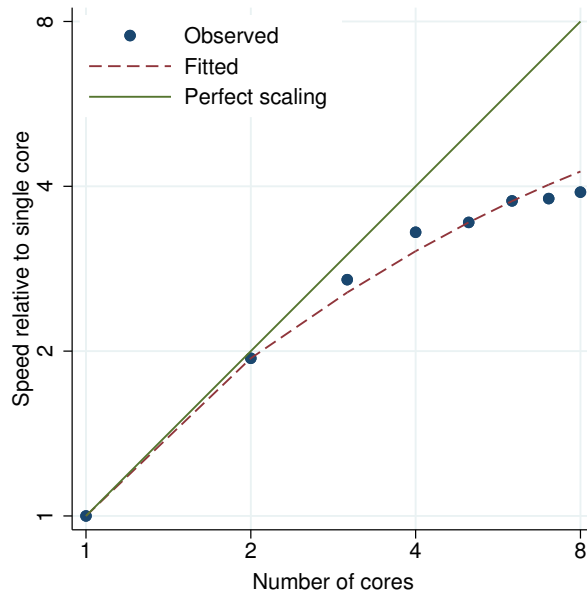
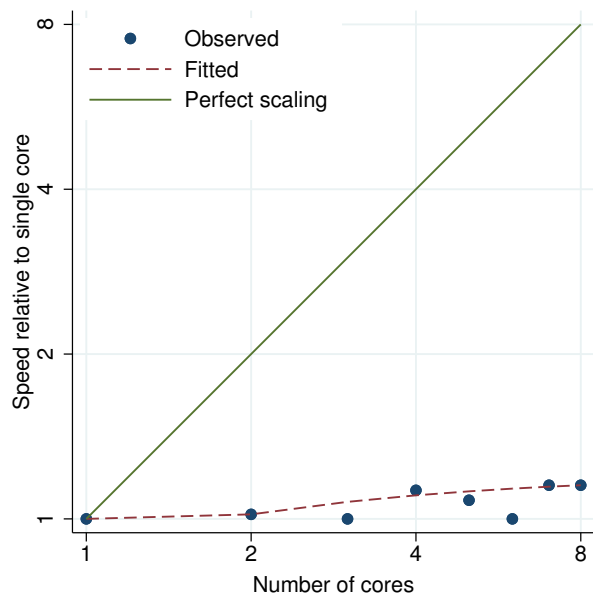
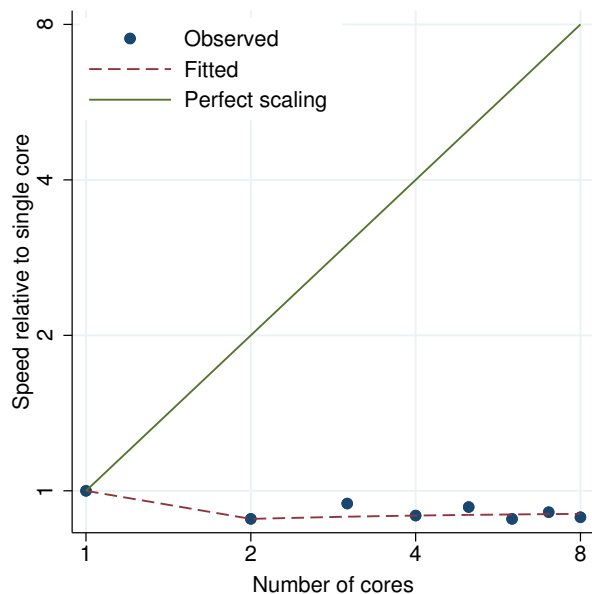
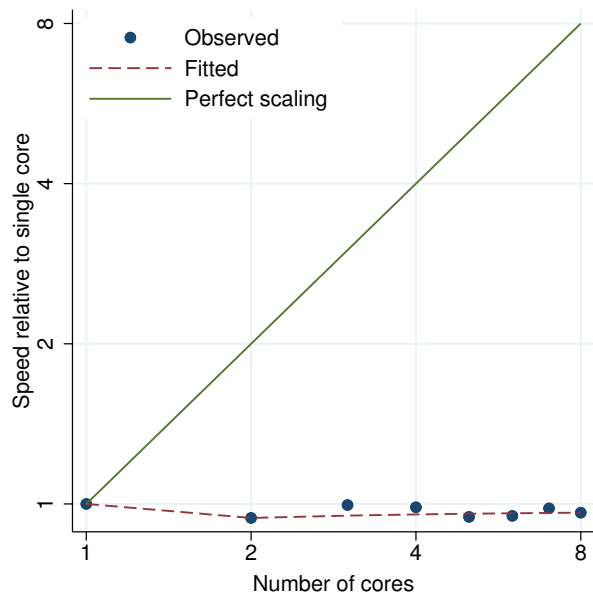
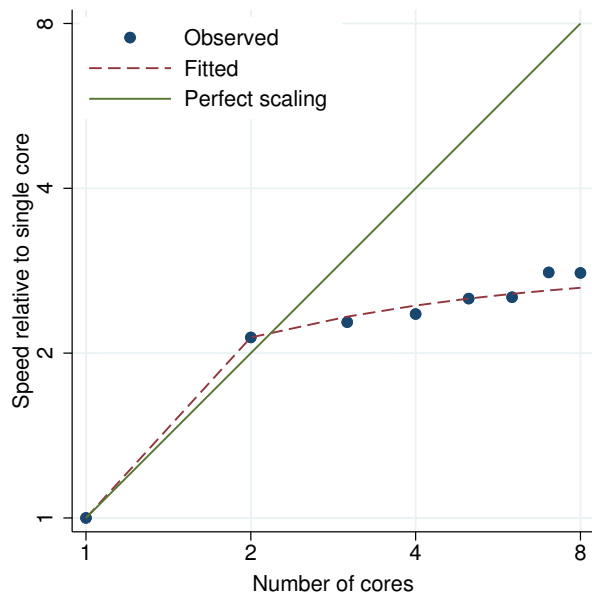
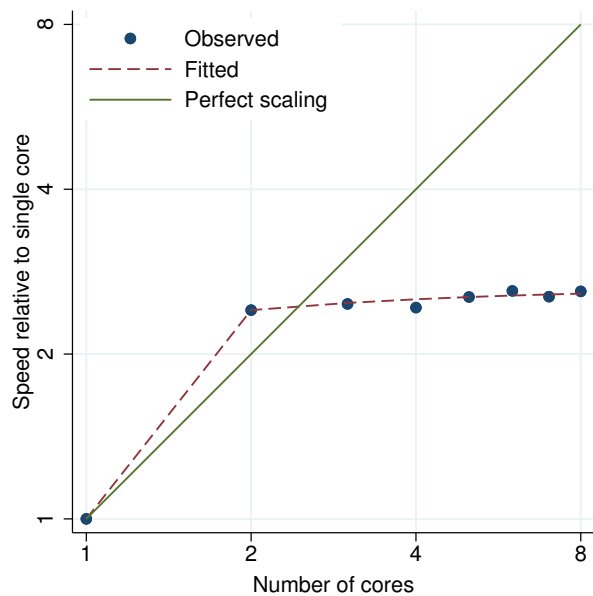
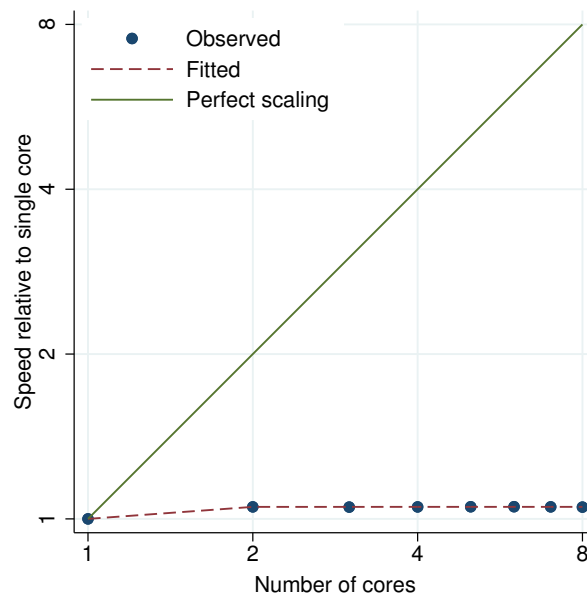
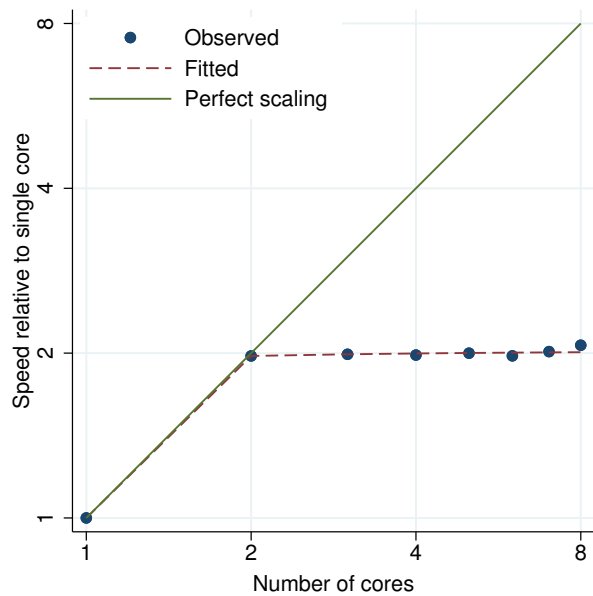
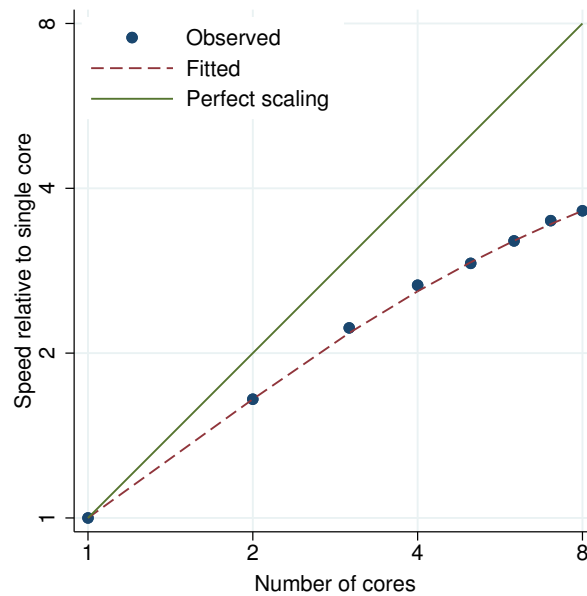


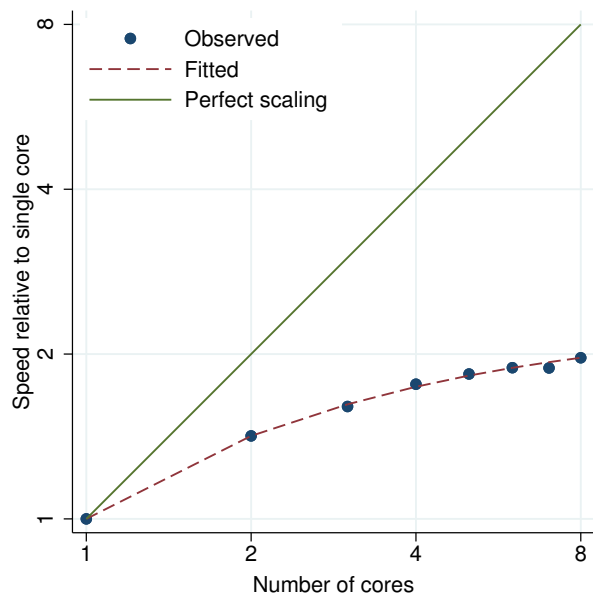
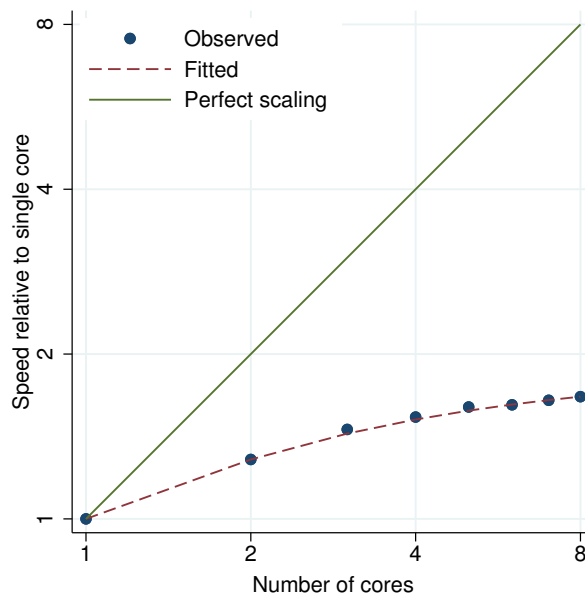
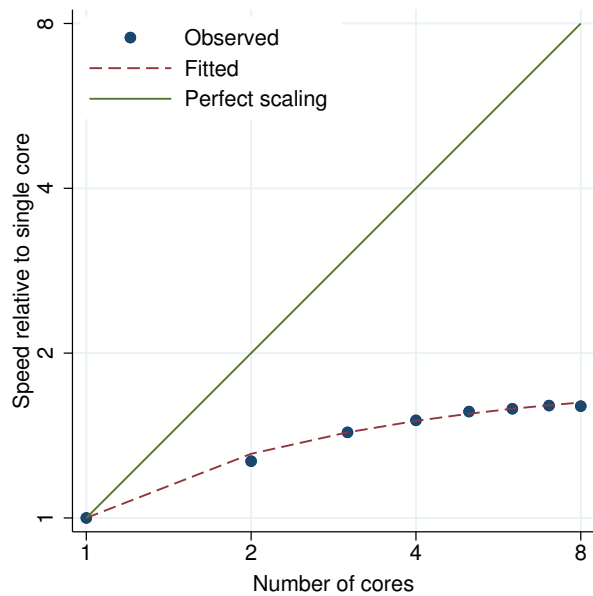
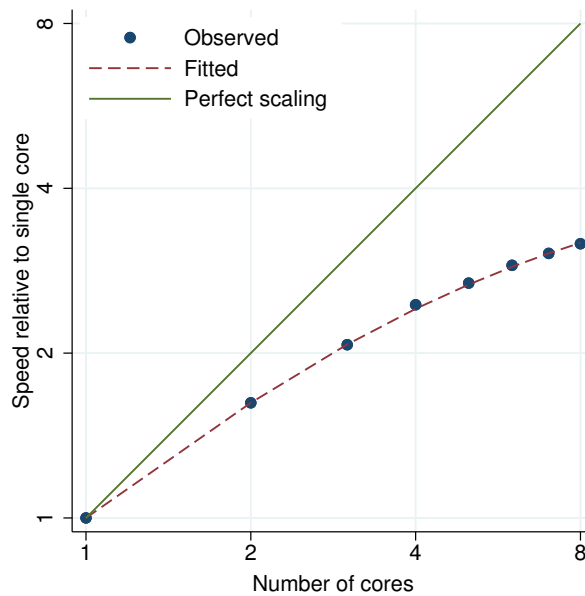
Figure 207. ivregress liml performance plot.

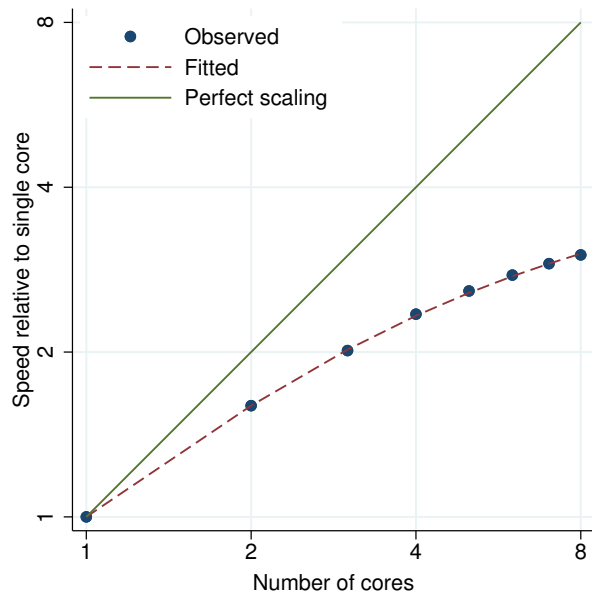
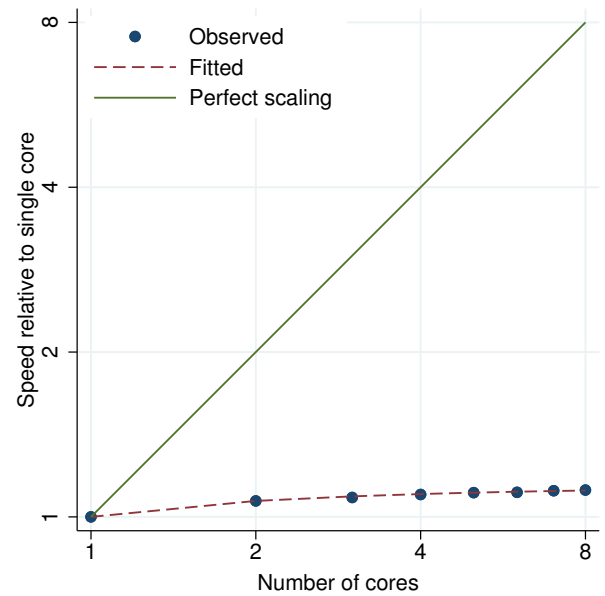
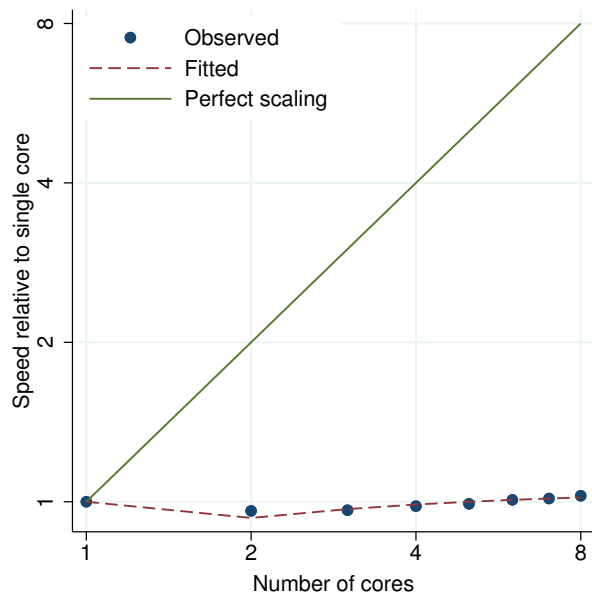
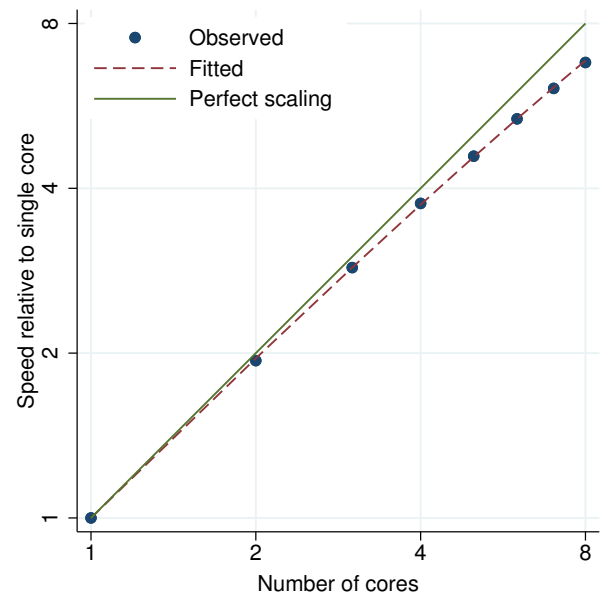
Figure 208. `ivtobit` performance plot.Figure 209. `kap` performance plot.Figure 210. `kappa` performance plot.Figure 211. `kdensity` performance plot.

Figure 212. `keep if exp` performance plot.Figure 213. `keep in range` performance plot.Figure 214. `keep varlist` performance plot.Figure 215. `ksmirnov` performance plot.

Figure 216. `ksmirnov, by()` performance plot.Figure 217. `ktau` performance plot.Figure 218. `kwallis` performance plot.Figure 219. `ladder` performance plot.



Figure 220. `lasso_linear` performance plot.Figure 221. `lasso_logit` performance plot.Figure 222. `lasso_poisson` performance plot.Figure 223. `gsem, lclass(C 2)` performance plot.

Figure 224. `gsem, lclass(C 3)` performance plot.Figure 225. `levelsof` performance plot.Figure 226. `loadingplot` performance plot.Figure 227. `logistic` performance plot.

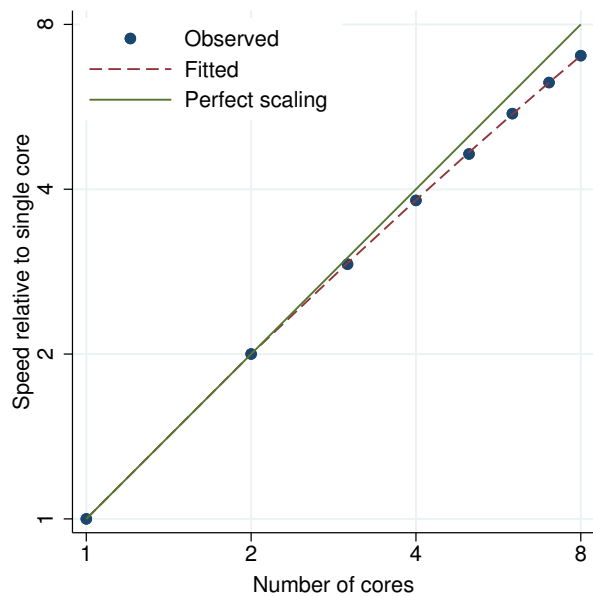


Figure 228. logit performance plot.

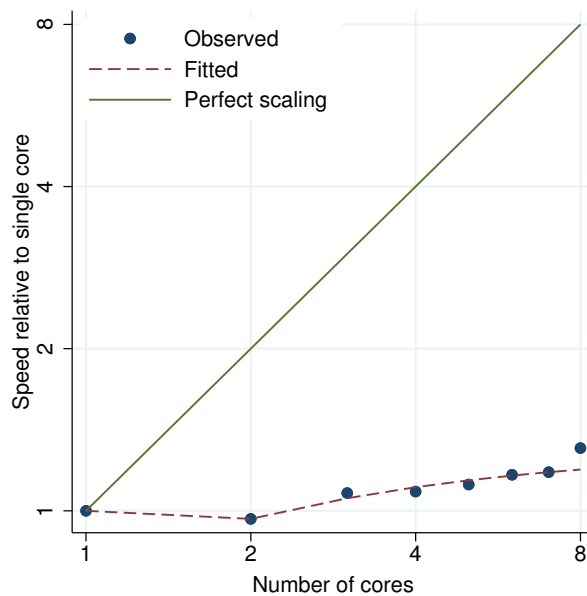


Figure 229. loneyway performance plot.

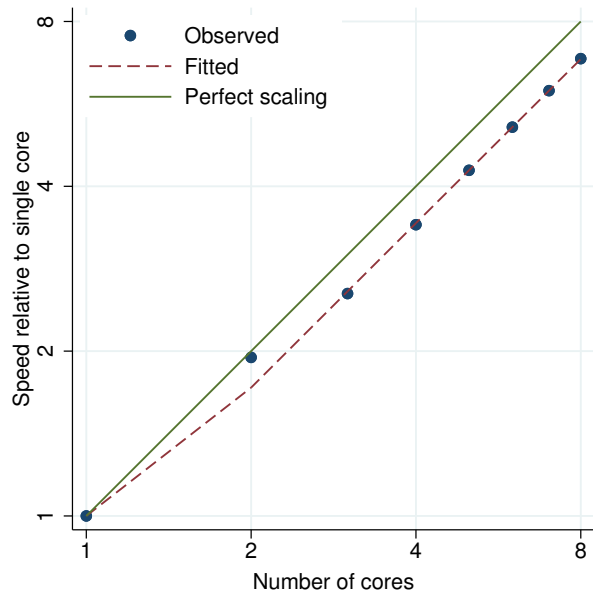


Figure 230. lowess performance plot.

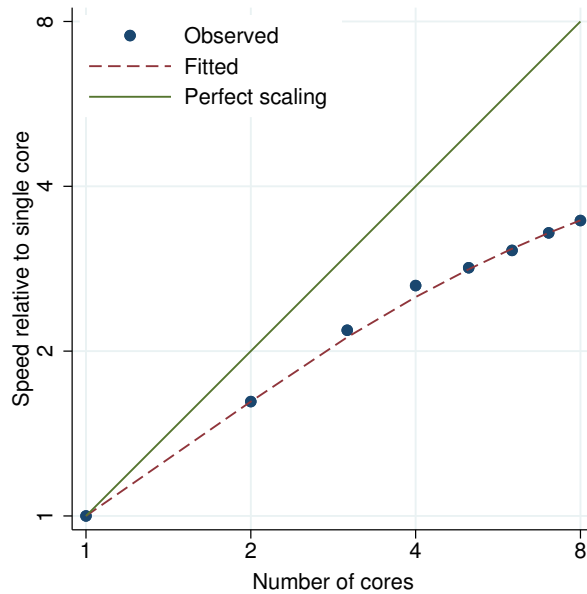
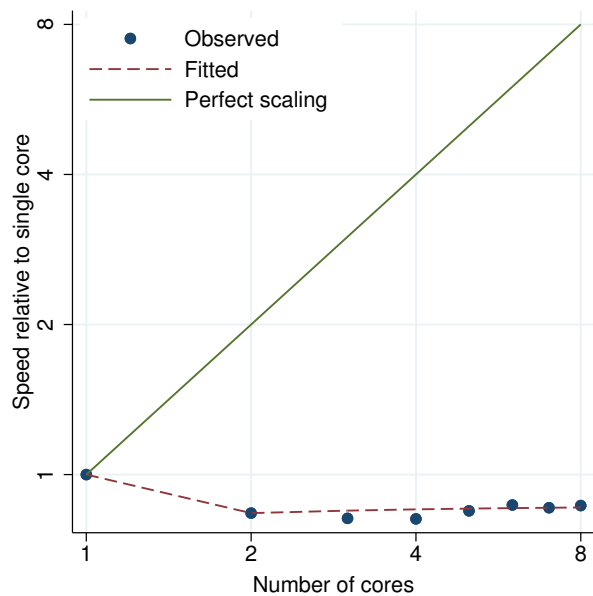
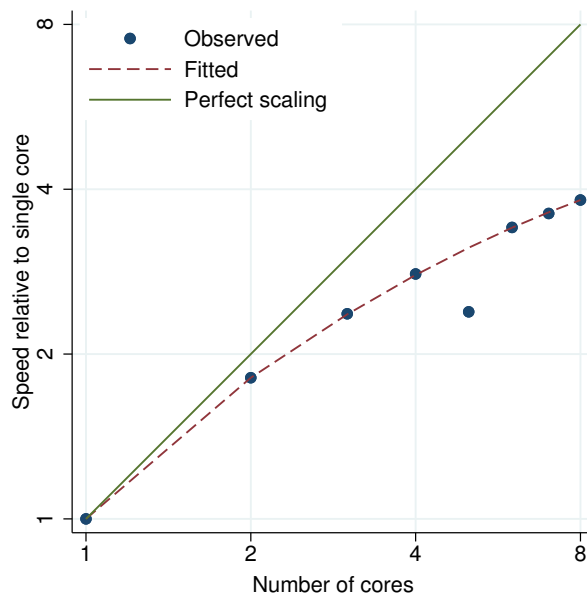
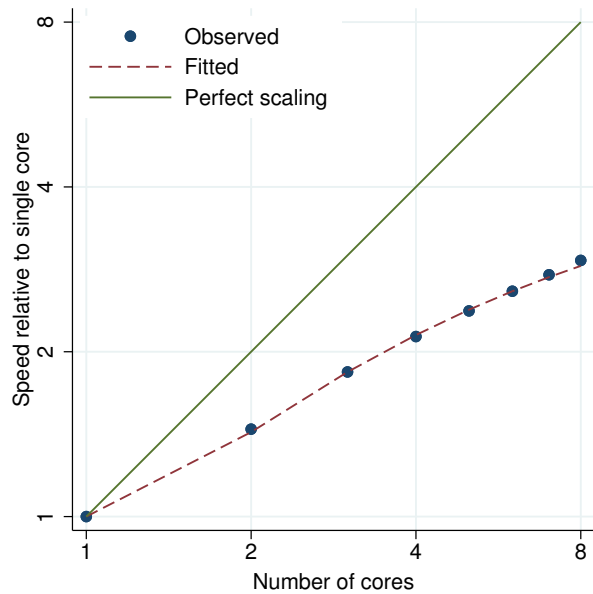
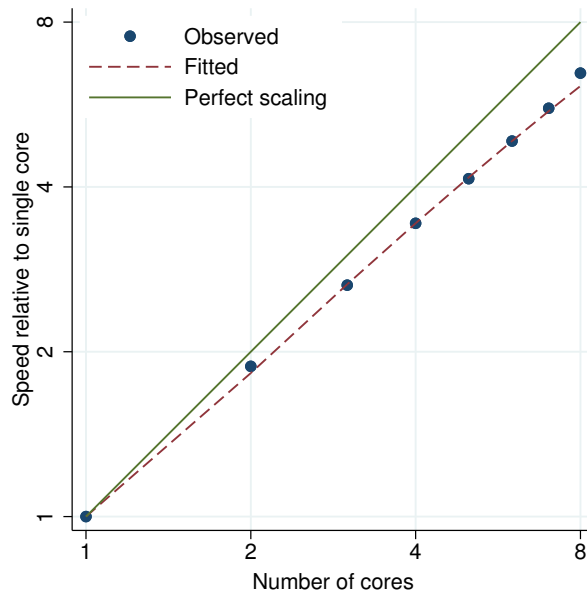


Figure 231. lpoly performance plot.

Figure 232. `1table` performance plot.Figure 233. `manova` (one-way) performance plot.Figure 234. `manova` (two-way) performance plot.Figure 235. `margins` performance plot.

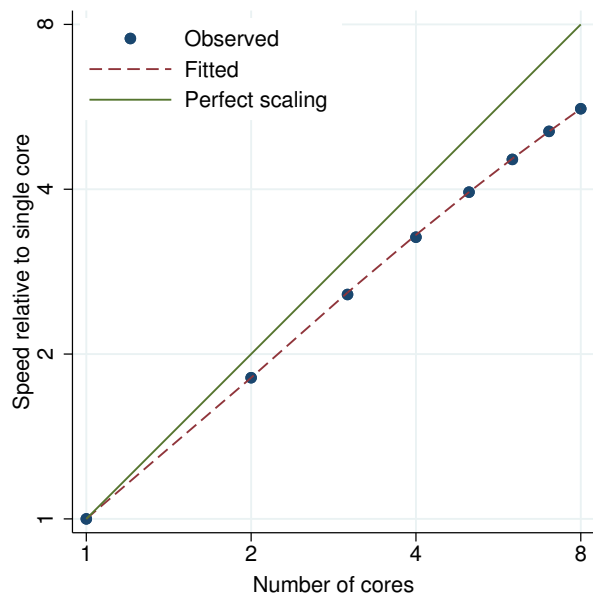


Figure 236. `margins, dydx() exp()` performance plot.

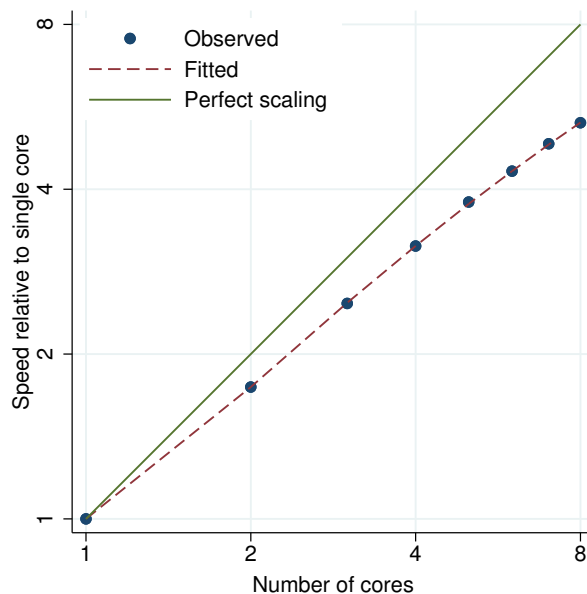


Figure 237. `margins, dydx()` performance plot.

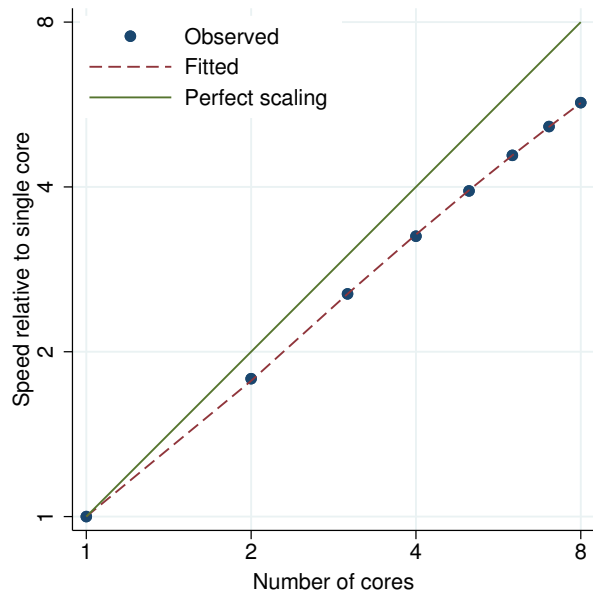


Figure 238. `margins, exp()` performance plot.

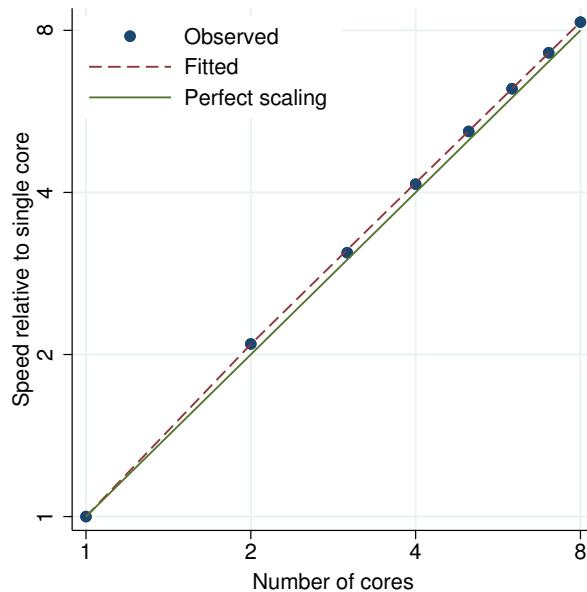


Figure 239. `markout` performance plot.

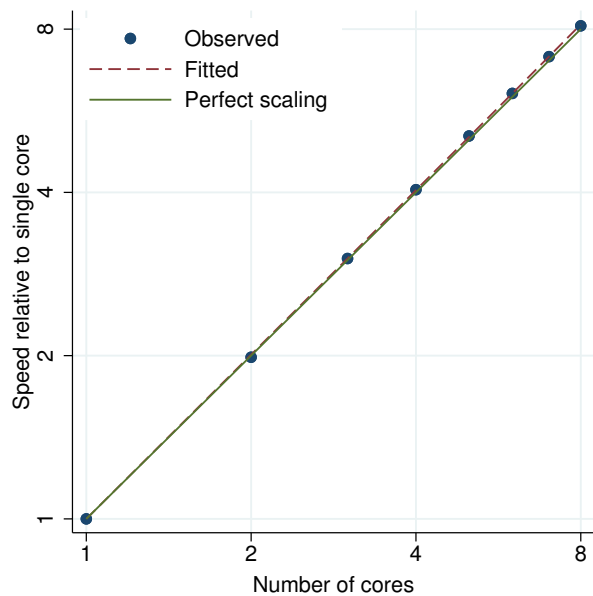


Figure 240. marksample performance plot.

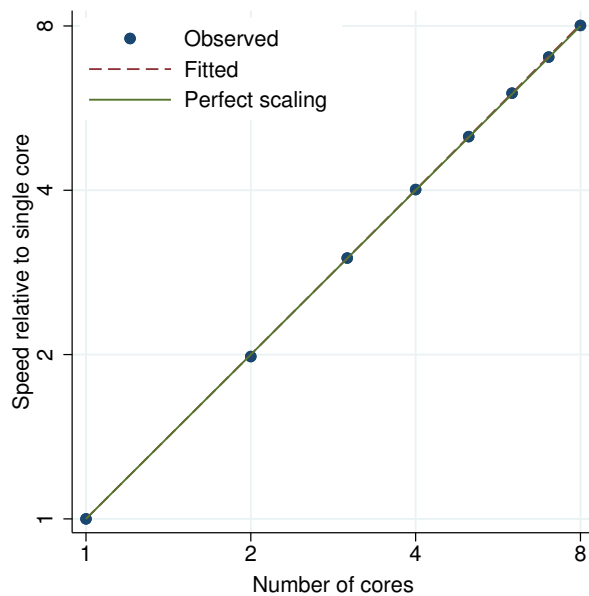
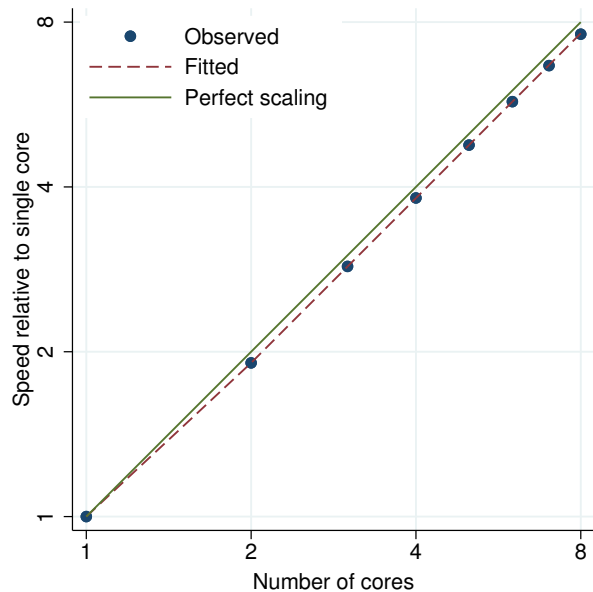
Figure 241. marksample if *exp* performance plot.

Figure 242. matrix accum performance plot.

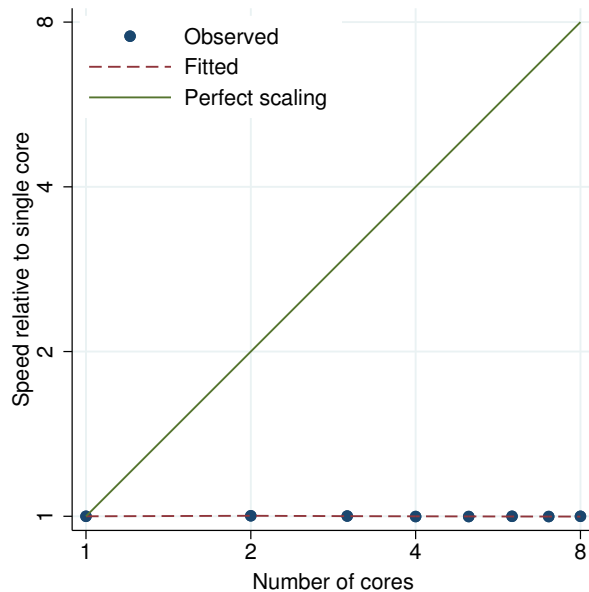
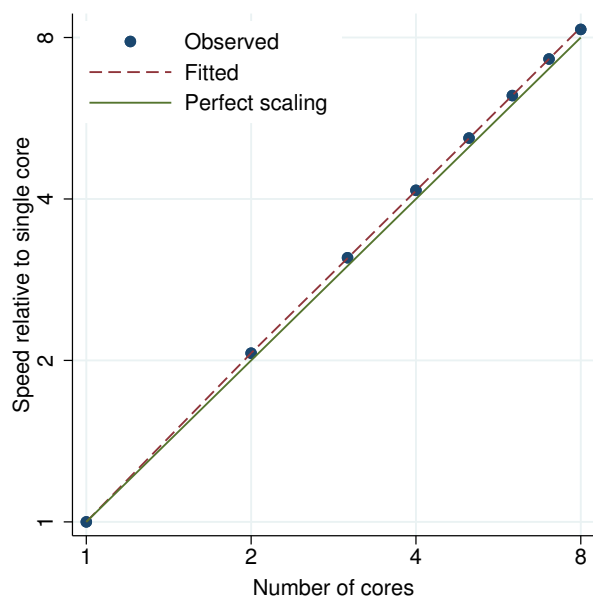
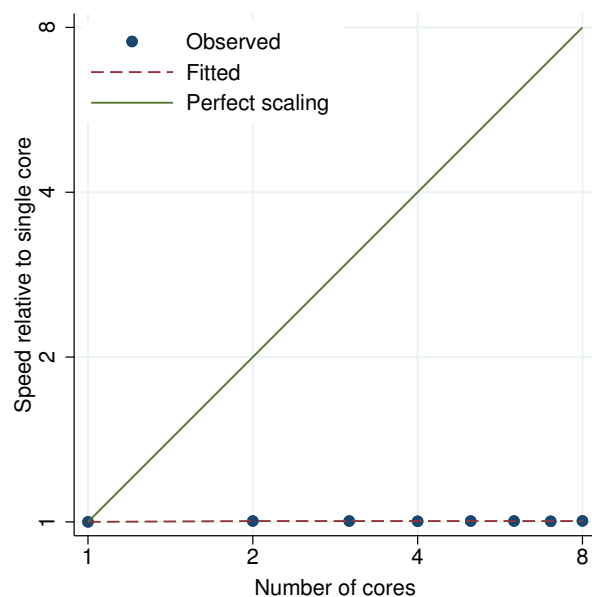
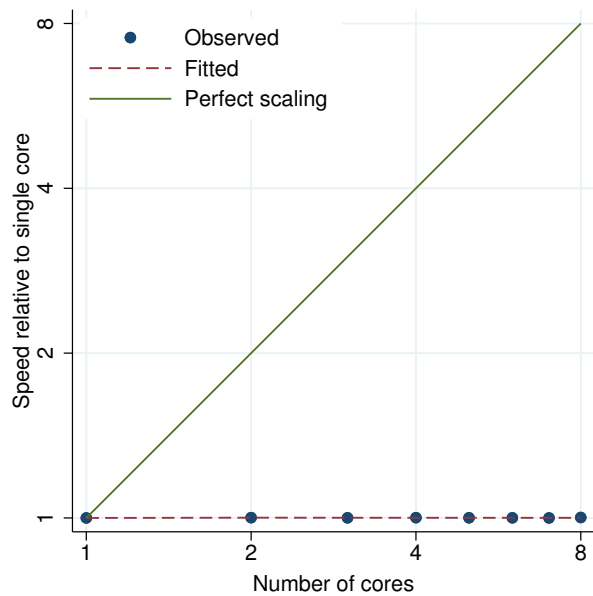
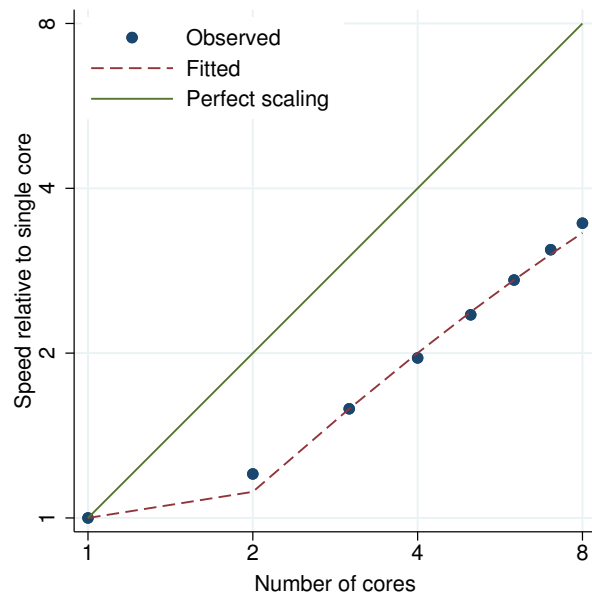


Figure 243. matrix eigenvalues performance plot.

Figure 244. `matrix score` performance plot.Figure 245. `matrix svd` performance plot.Figure 246. `matrix symeigen` performance plot.Figure 247. `matrix syminv` performance plot.

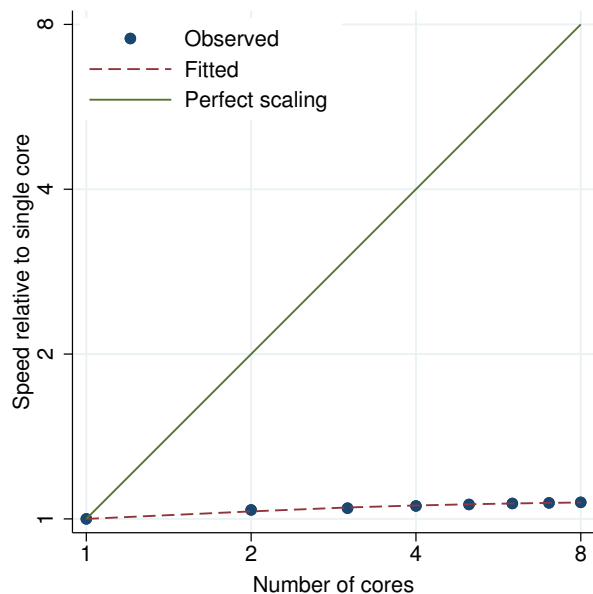


Figure 248. mca performance plot.

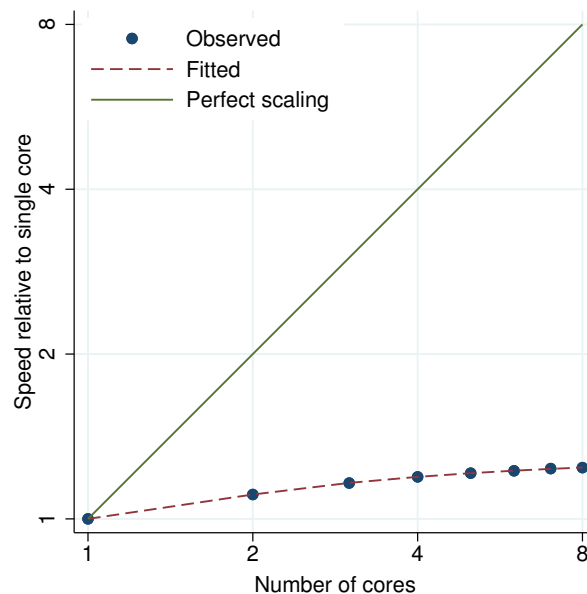


Figure 249. mcc performance plot.

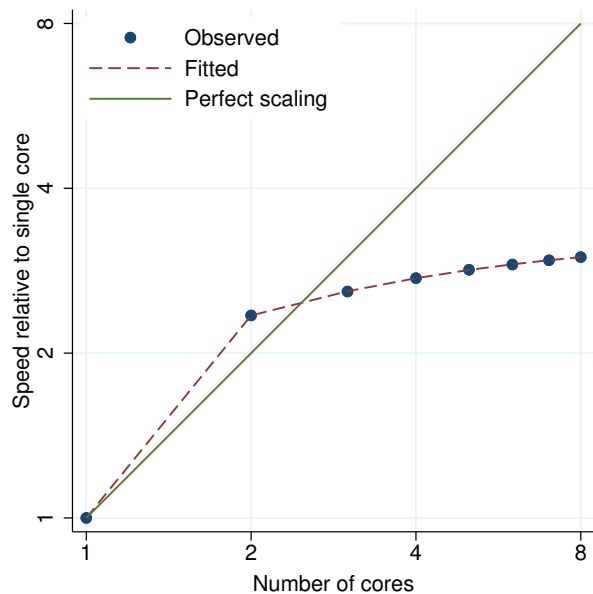


Figure 250. mds performance plot.

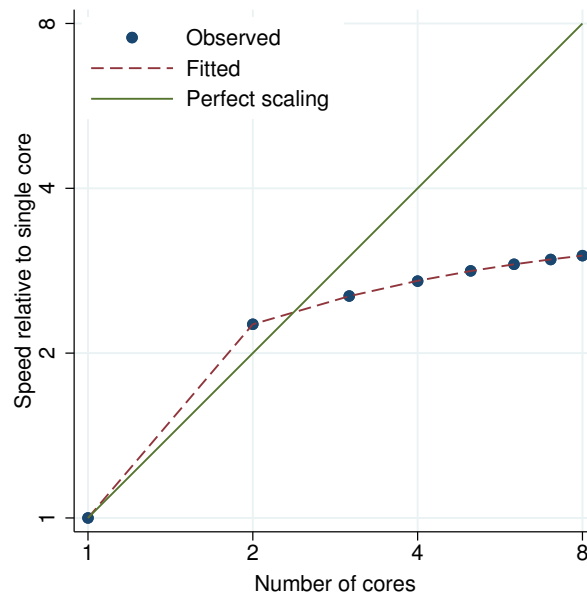


Figure 251. mdslong performance plot.



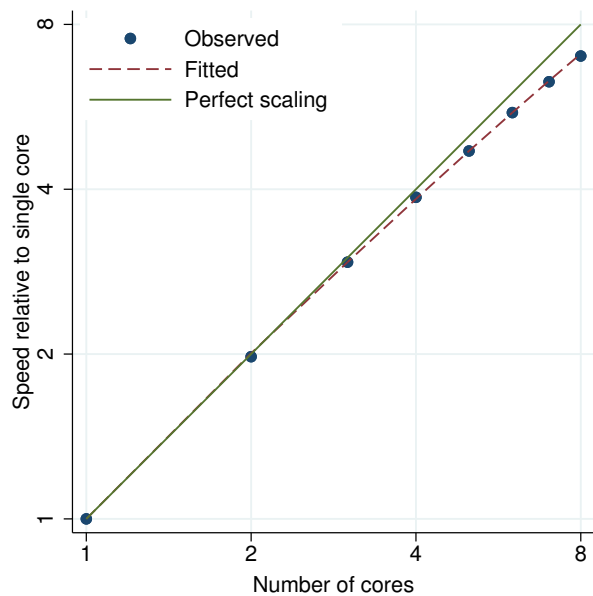


Figure 252. mean performance plot.

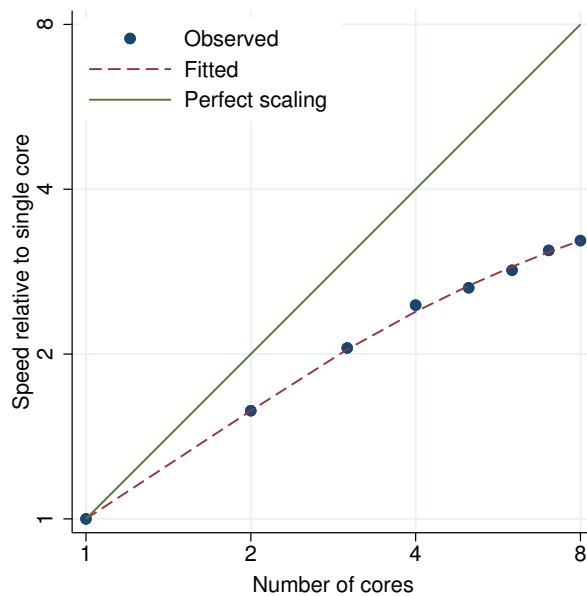


Figure 253. mecloglog performance plot.

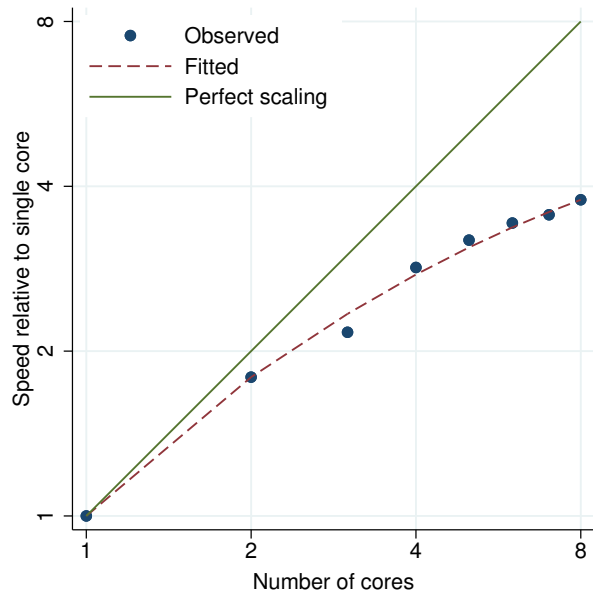


Figure 254. median performance plot.

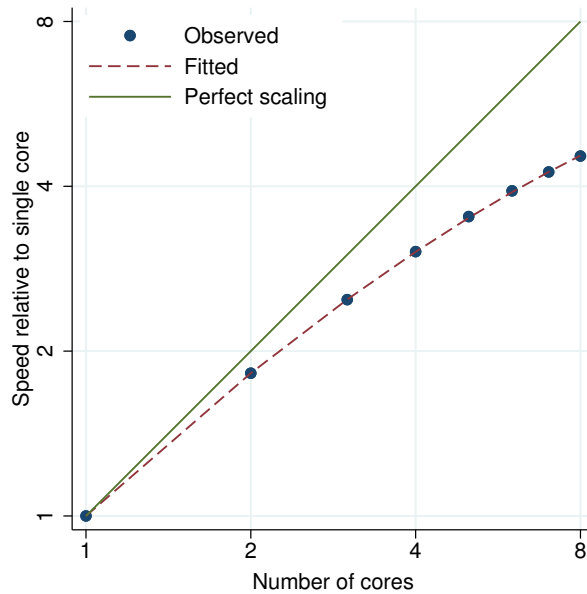


Figure 255. meintreg performance plot.

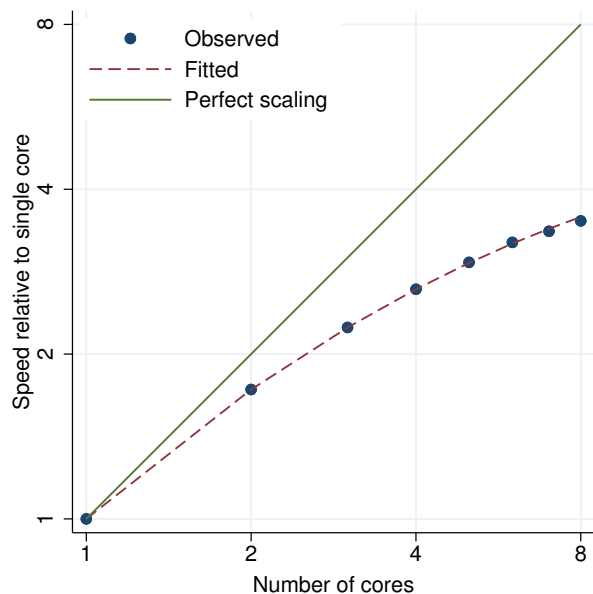


Figure 256. melogit performance plot.

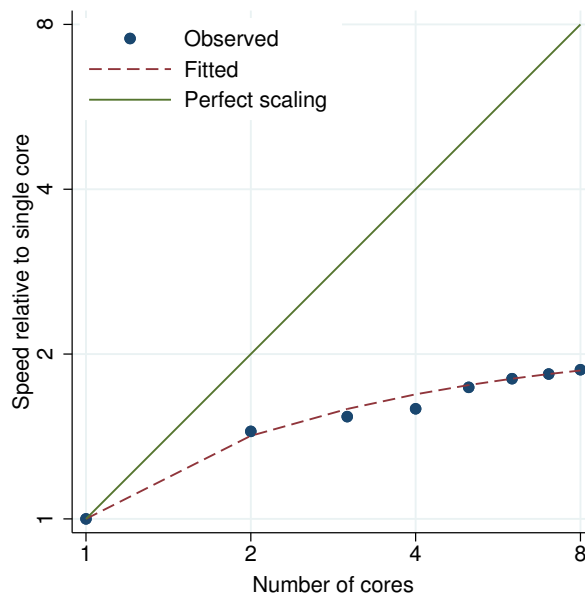
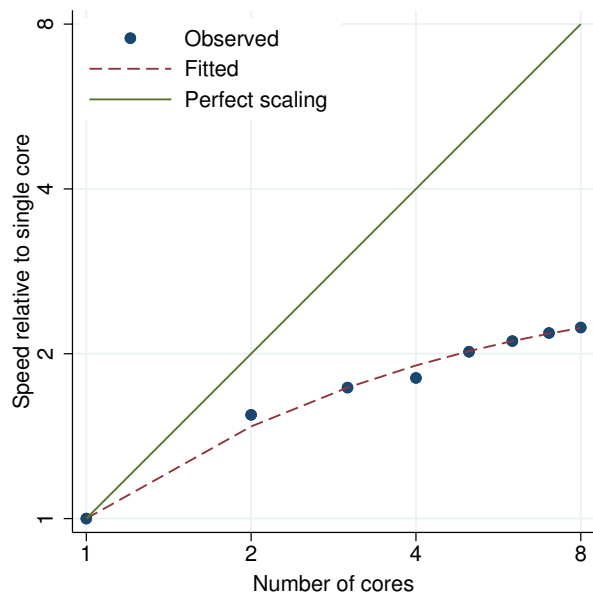
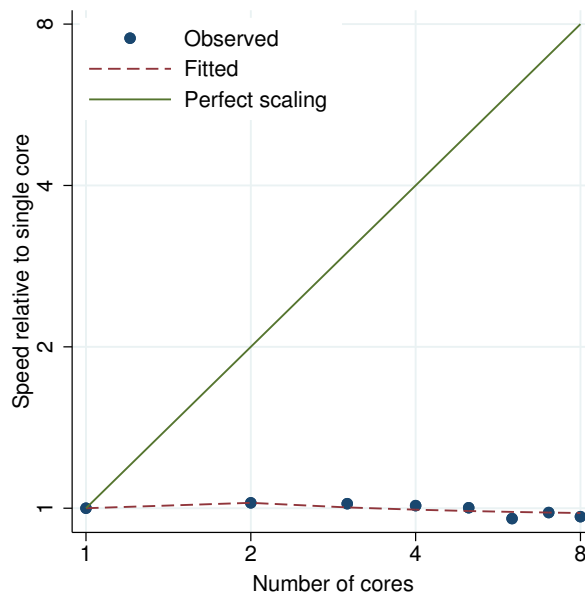
Figure 257. menbreg,  
dispersion(constant) performance plot.Figure 258. menbreg, dispersion(mean)  
performance plot.

Figure 259. menl performance plot.

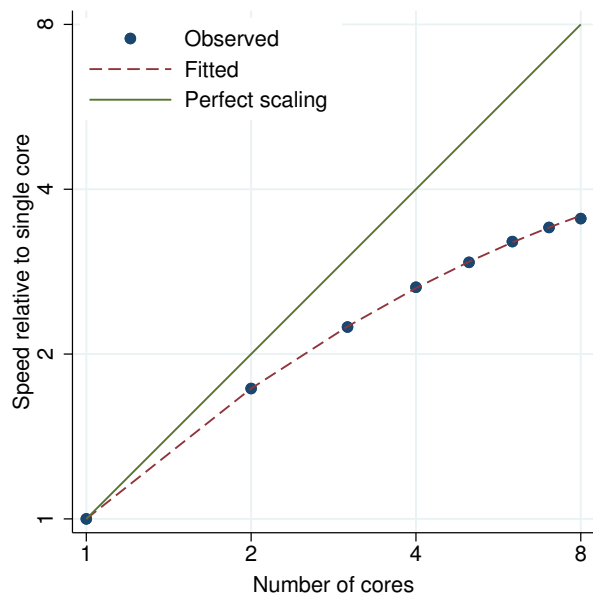


Figure 260. meologit performance plot.

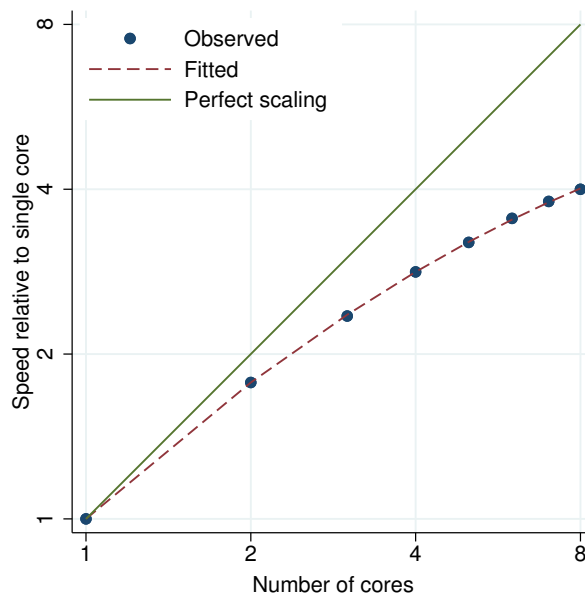


Figure 261. meoprobit performance plot.

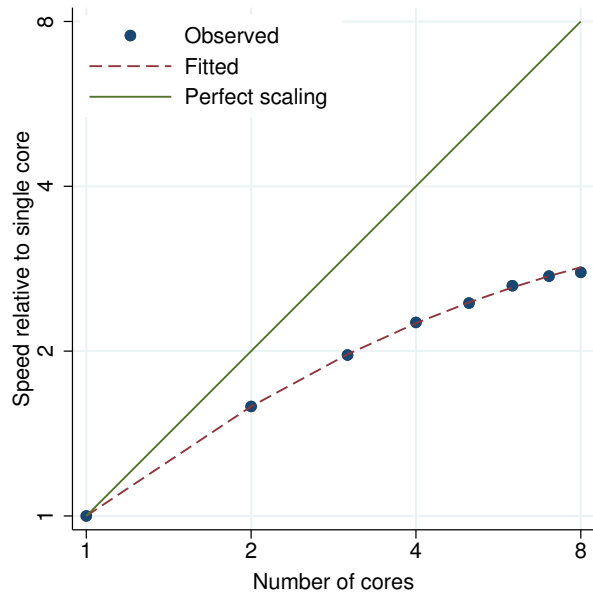


Figure 262. mepoisson performance plot.

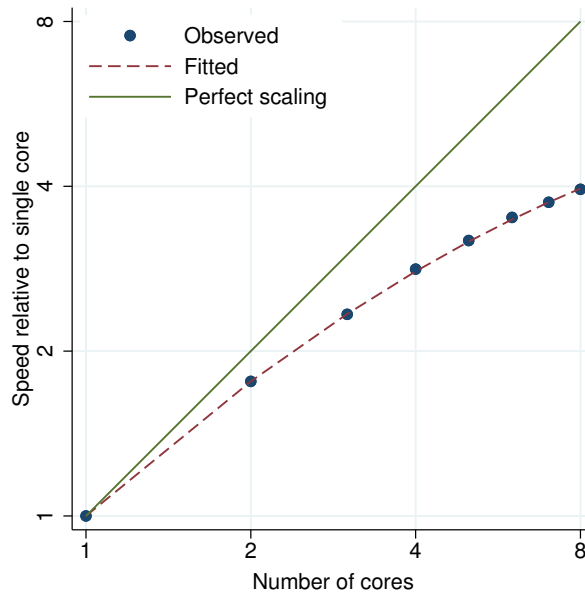


Figure 263. meprobit performance plot.

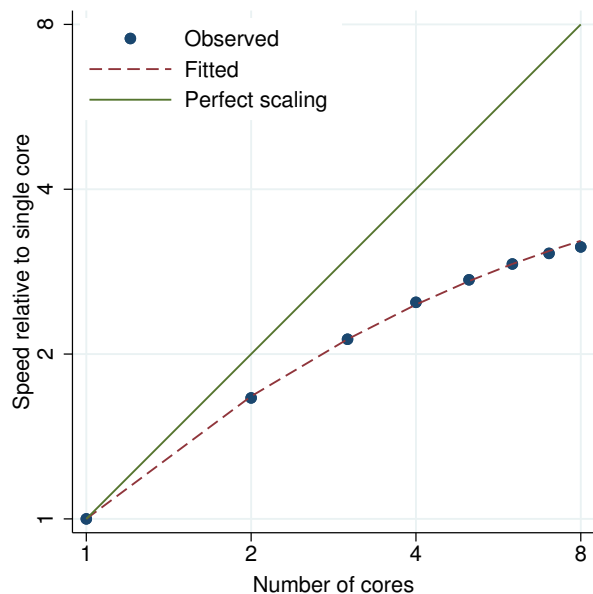


Figure 264. `mestreg, distribution(exp)` performance plot.

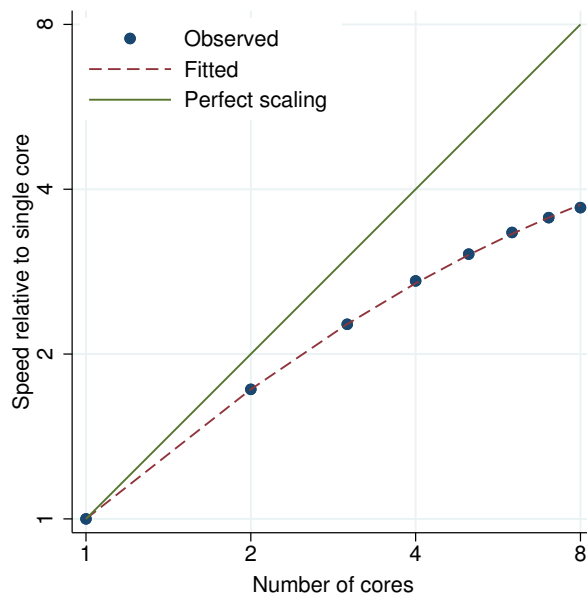


Figure 265. `mestreg, distribution(weibull)` performance plot.

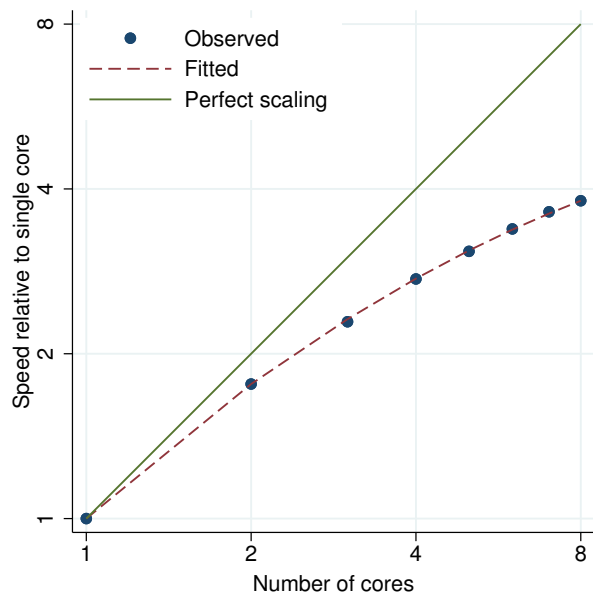


Figure 266. `metobit` performance plot.

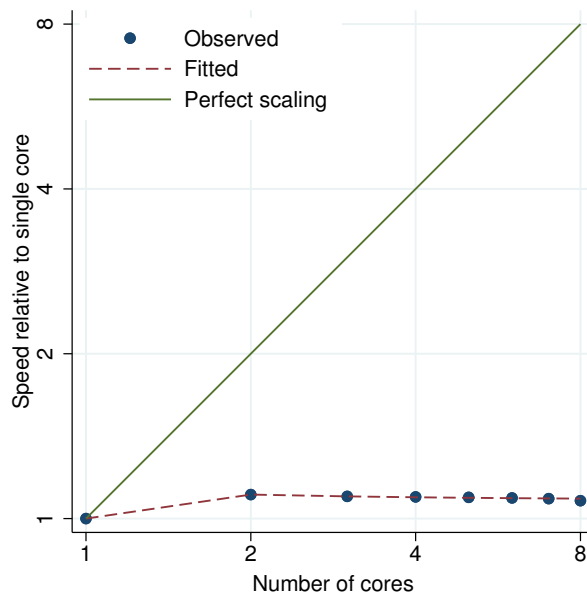


Figure 267. `mgarch` performance plot.

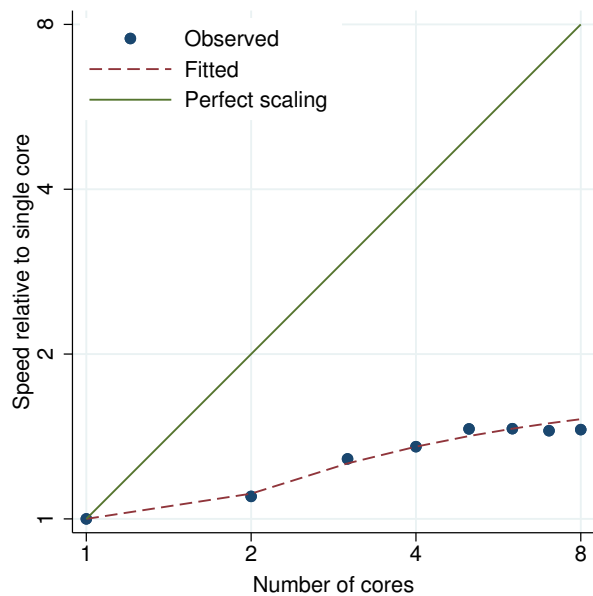


Figure 268. mhodds performance plot.

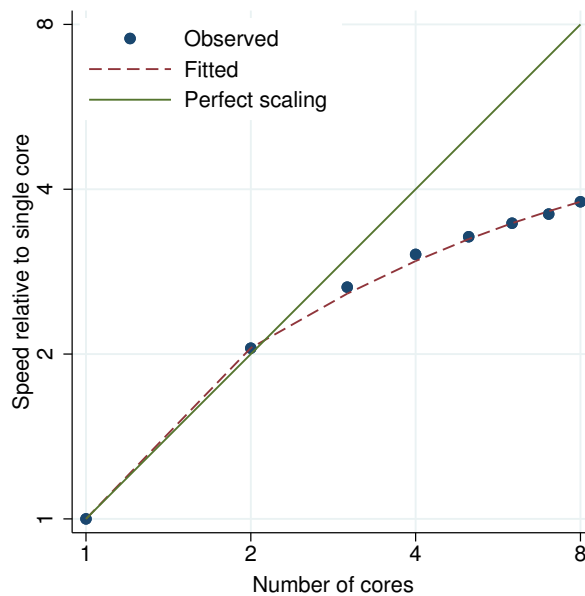


Figure 269. mhodds (adjusted) performance plot.

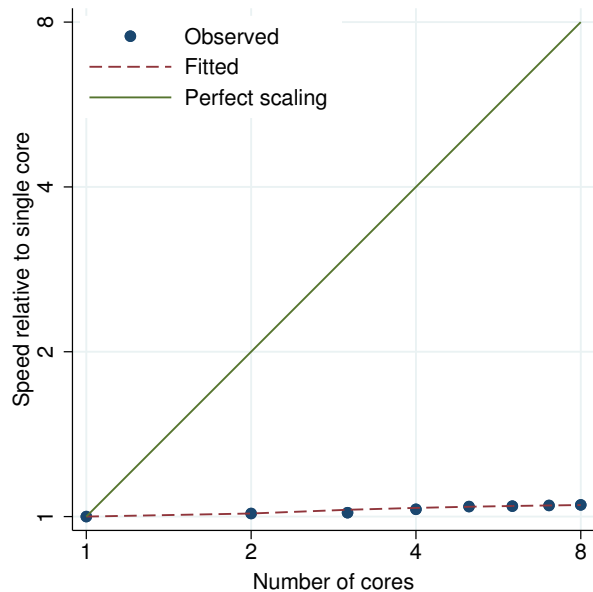


Figure 270. by: mhodds performance plot.

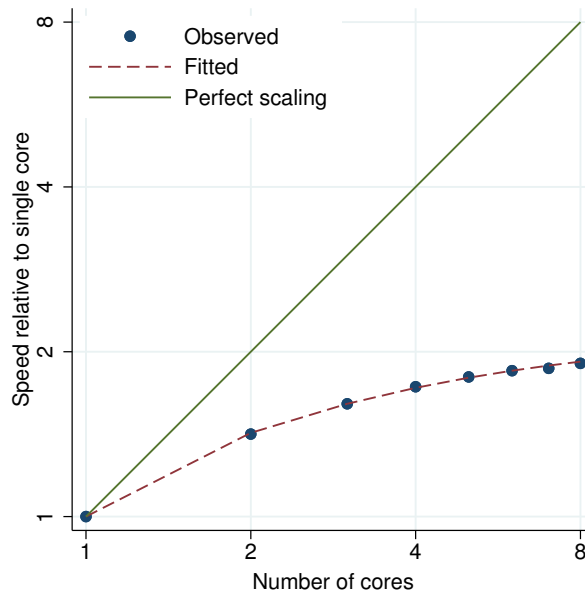


Figure 271. mhodds (trend) performance plot.

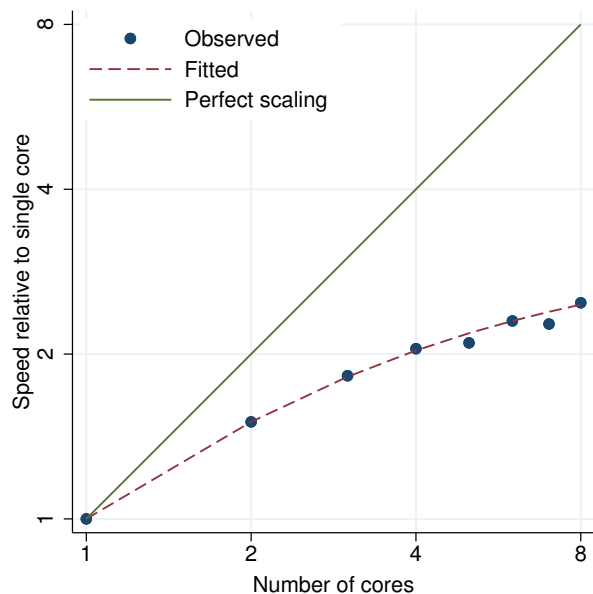


Figure 272. mi estimate: logit (flong) performance plot.

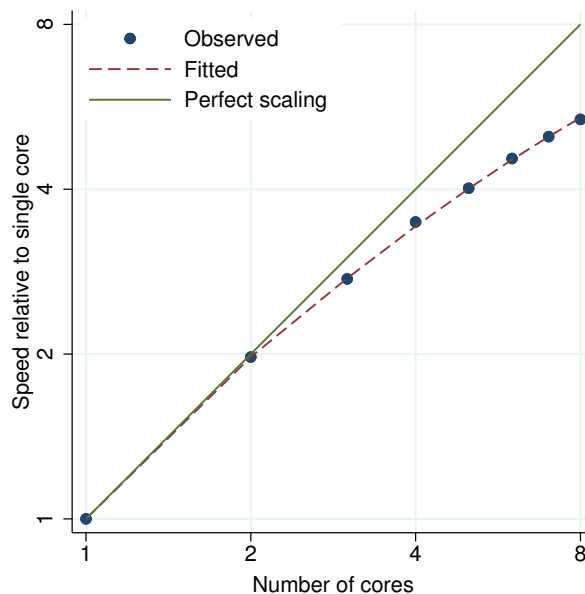


Figure 273. mi estimate: logit (flongsep) performance plot.

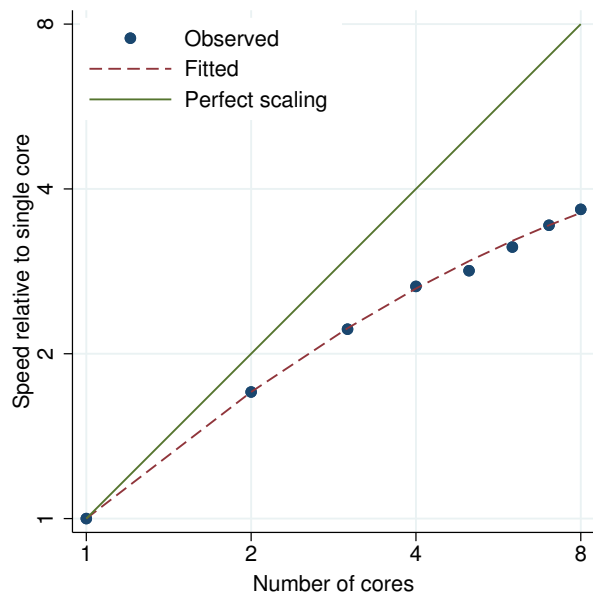


Figure 274. mi estimate: logit (mlong) performance plot.

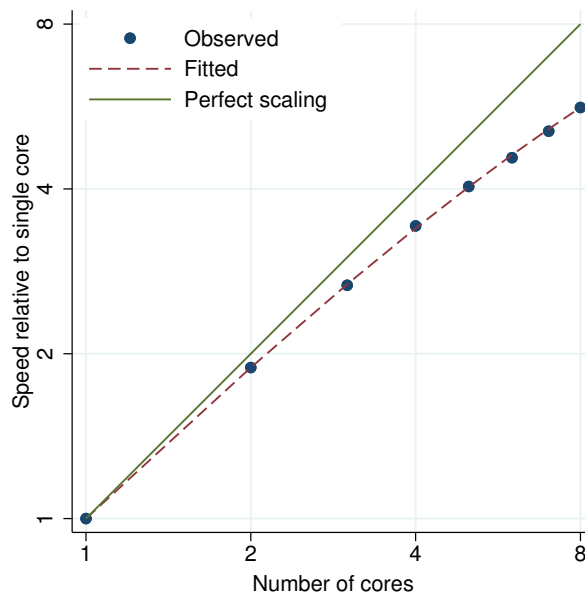


Figure 275. mi estimate: logit (wide) performance plot.

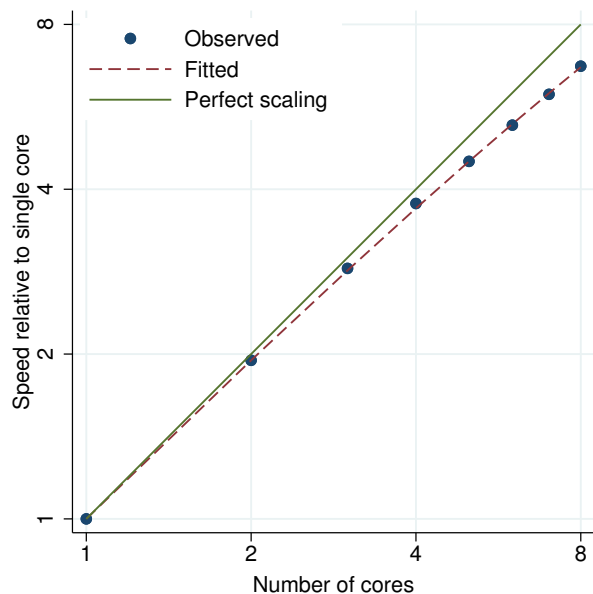


Figure 276. mi estimate: mlogit performance plot.

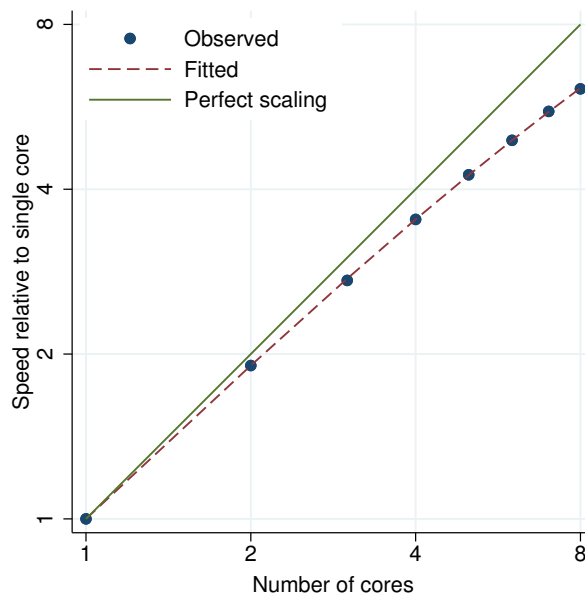


Figure 277. mi estimate: ologit performance plot.

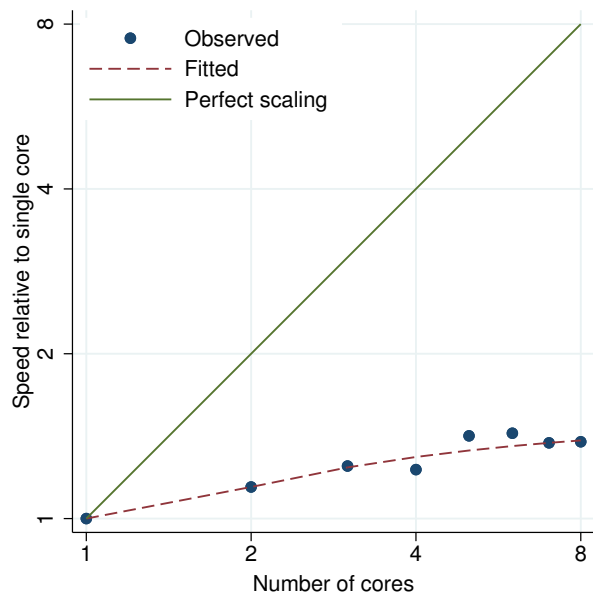


Figure 278. mi estimate: regress (flong) performance plot.

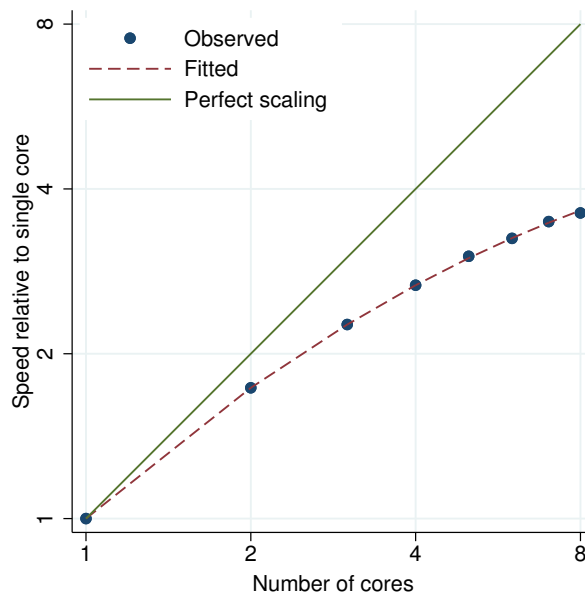


Figure 279. mi estimate: regress (flongsep) performance plot.

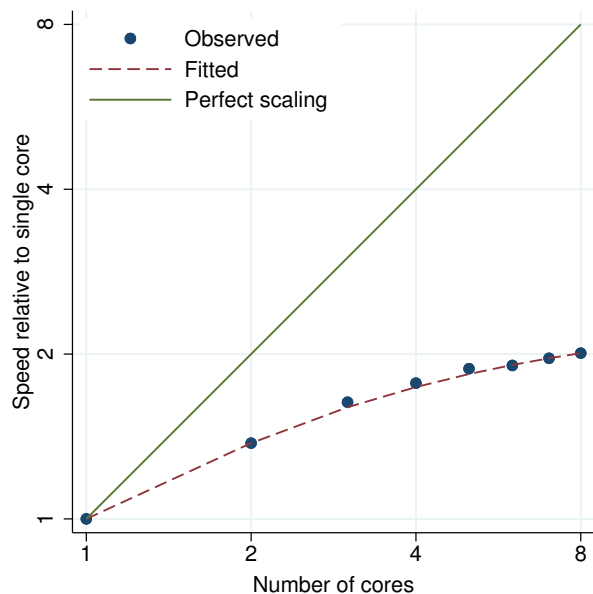


Figure 280. `mi estimate: regress (mlong)` performance plot.

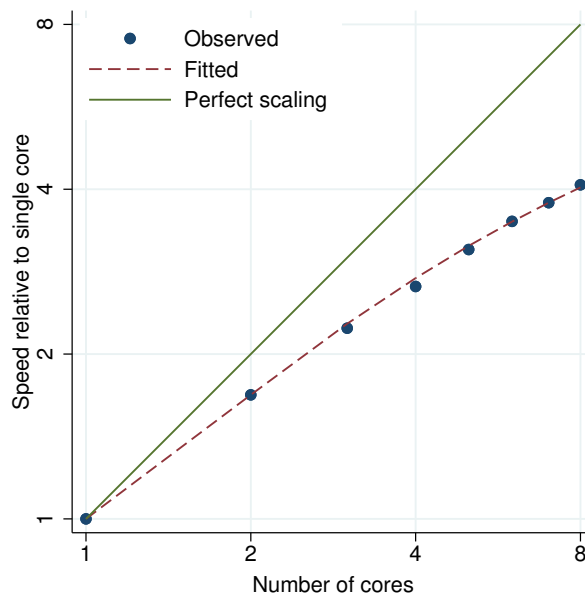


Figure 281. `mi estimate: regress (wide)` performance plot.

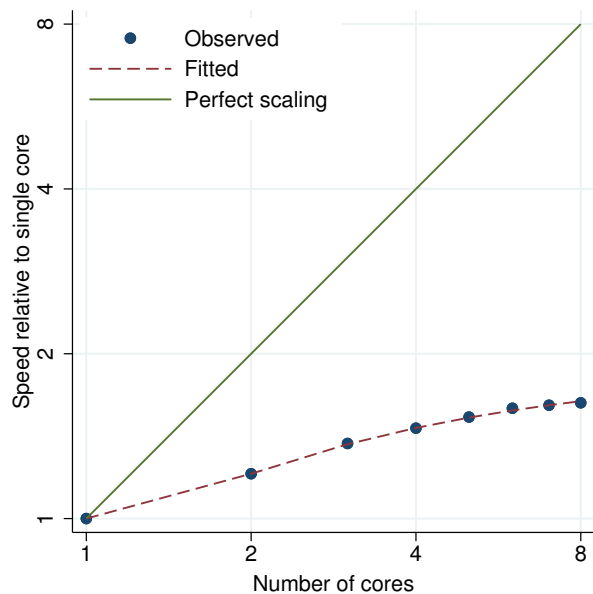


Figure 282. `mi impute chained (flong)` performance plot.

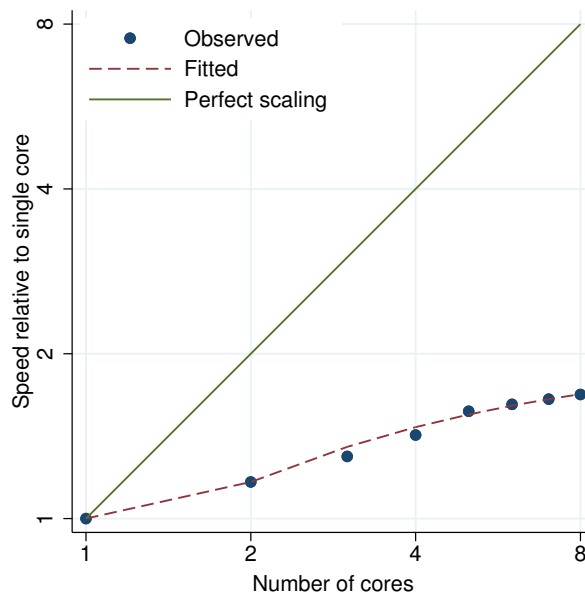


Figure 283. `mi impute chained (flongsep)` performance plot.



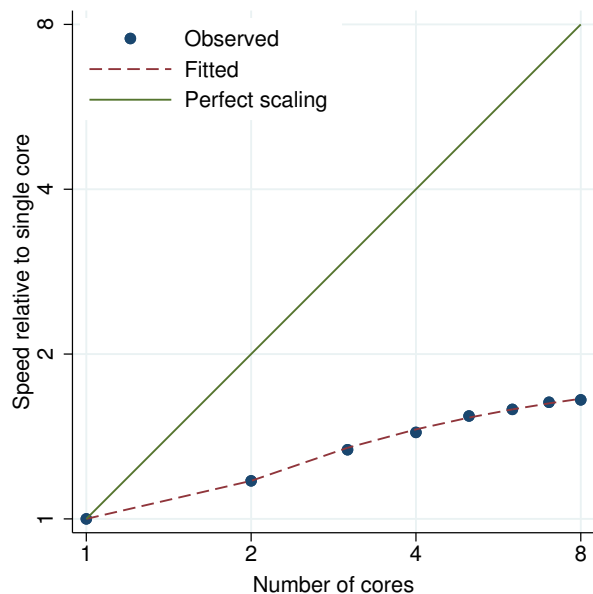


Figure 284. `mi impute chained (mlong)` performance plot.

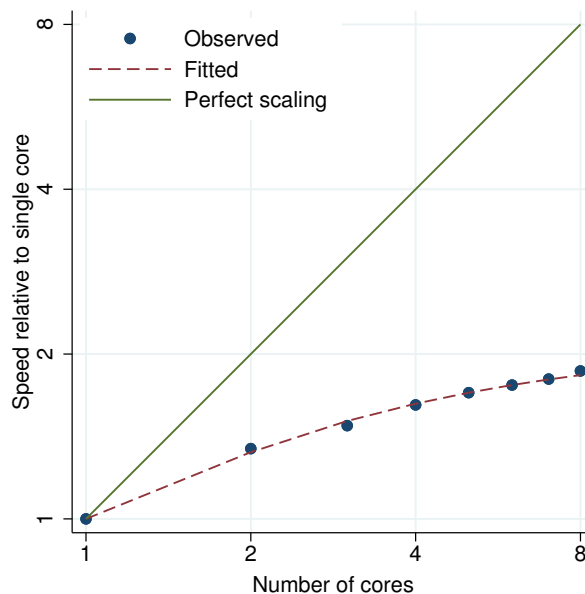


Figure 285. `mi impute chained (wide)` performance plot.

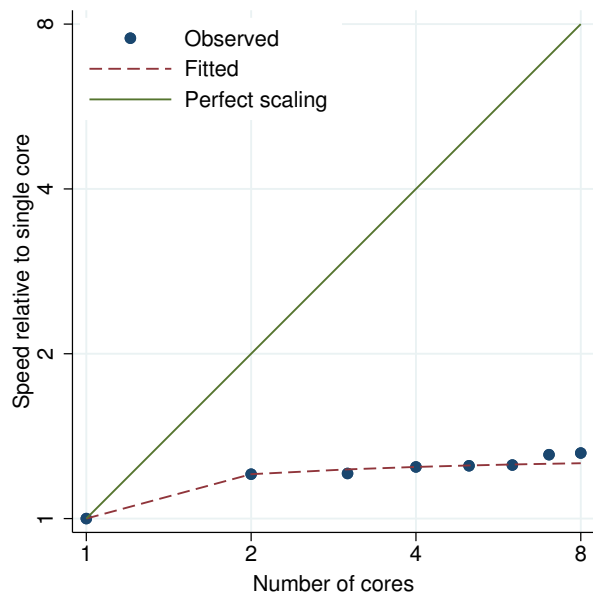


Figure 286. `mi impute logit (flong)` performance plot.

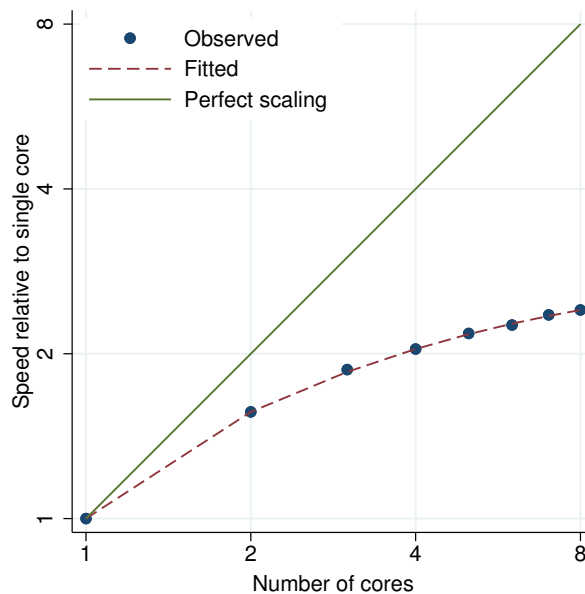


Figure 287. `mi impute logit (flongsep)` performance plot.

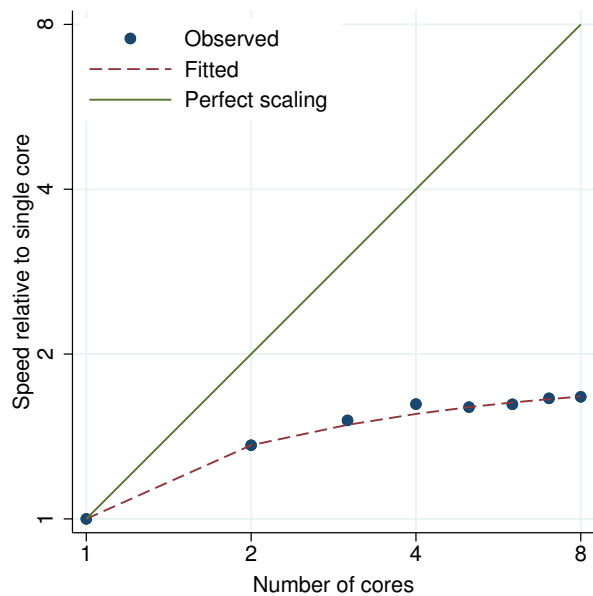


Figure 288. `mi impute logit (mlong)` performance plot.

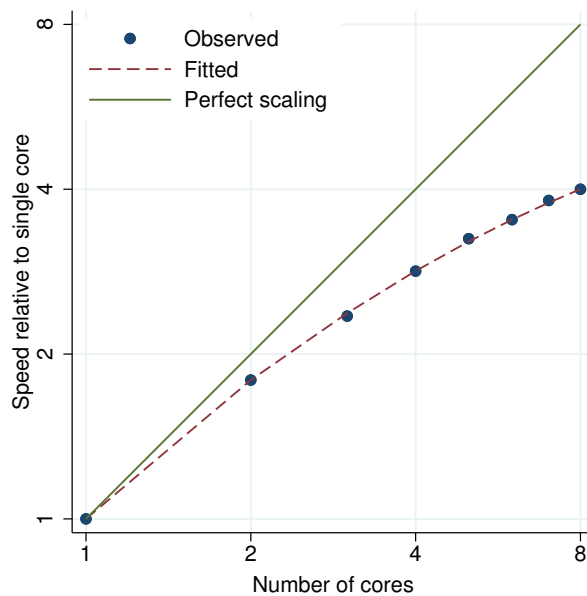


Figure 289. `mi impute logit (wide)` performance plot.

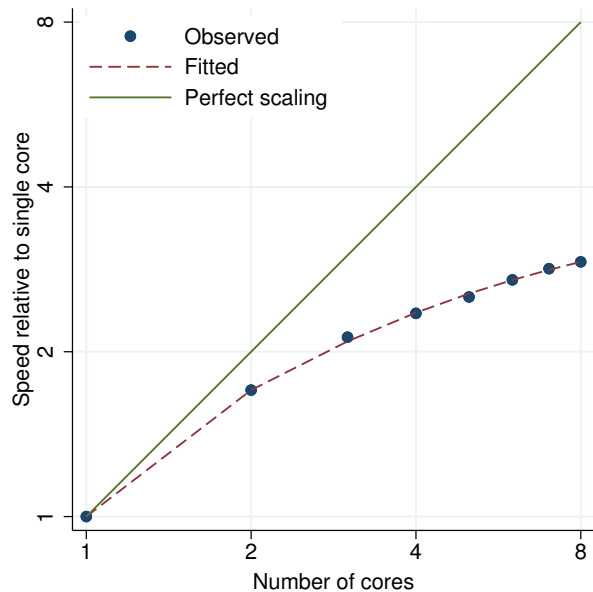


Figure 290. `mi impute mlogit` performance plot.

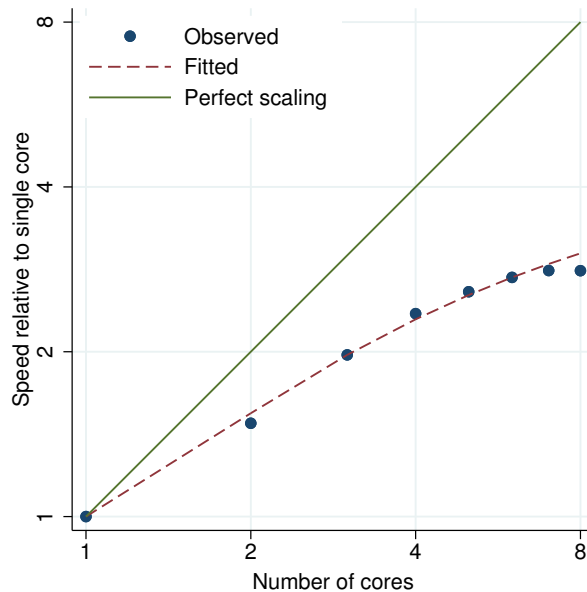
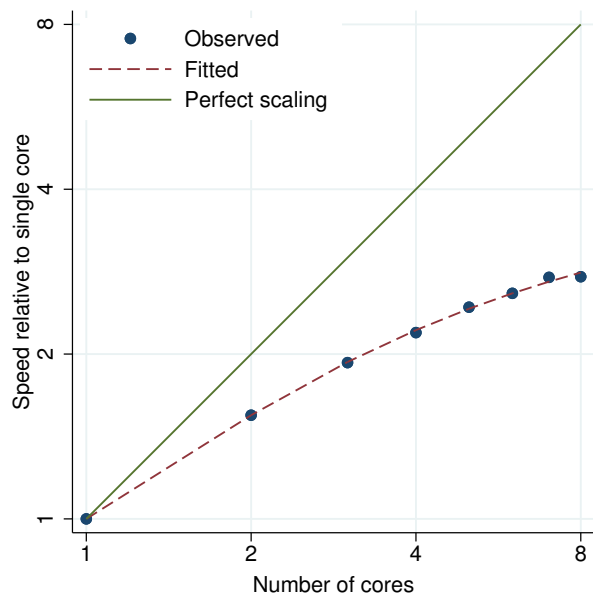
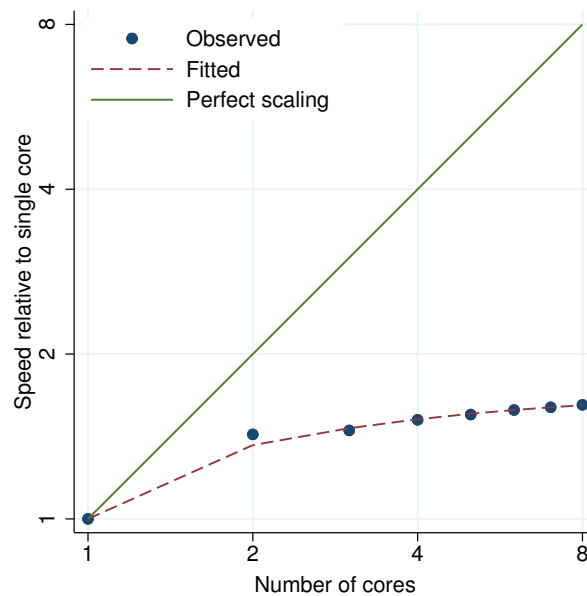
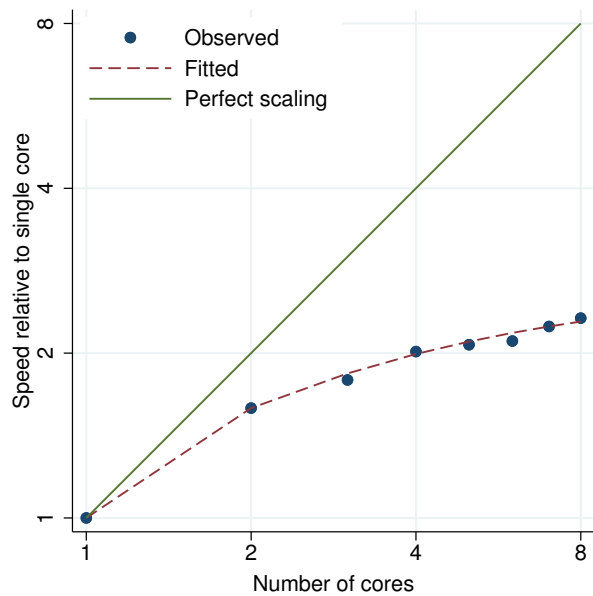
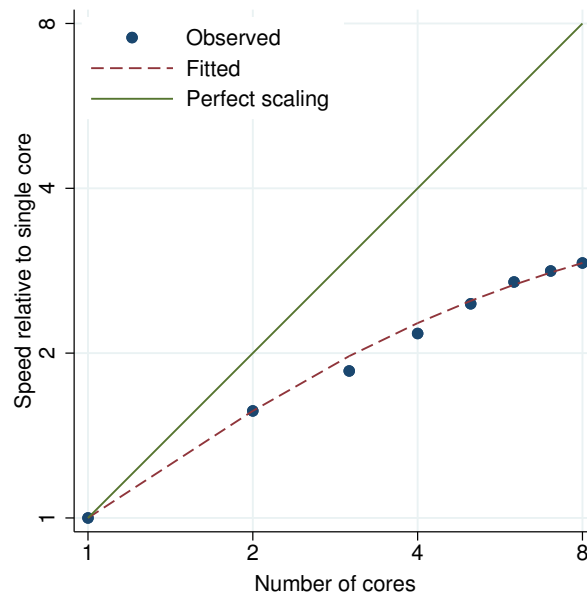


Figure 291. `mi impute mono pmm` performance plot.

Figure 292. `mi impute mono regress` performance plot.Figure 293. `mi impute mvn` performance plot.Figure 294. `mi impute ologit` performance plot.Figure 295. `mi impute pmm` performance plot.

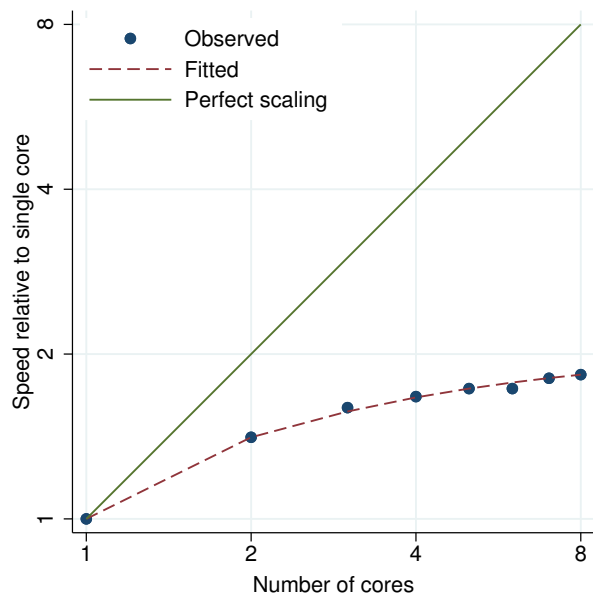


Figure 296. `mi impute regress` performance plot.

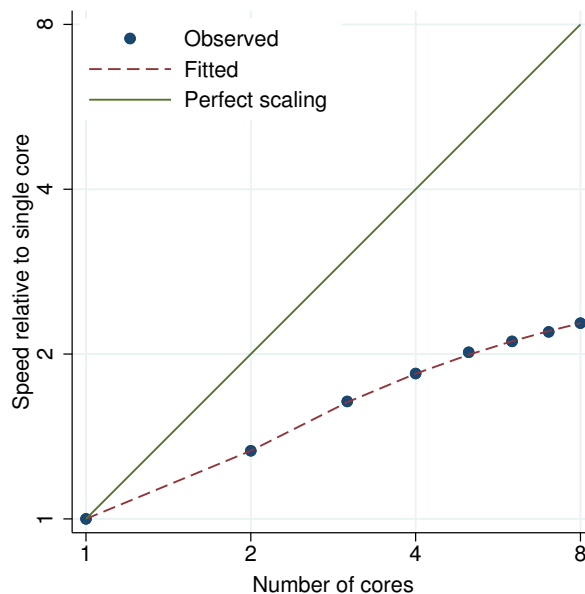


Figure 297. `misstable nested` performance plot.

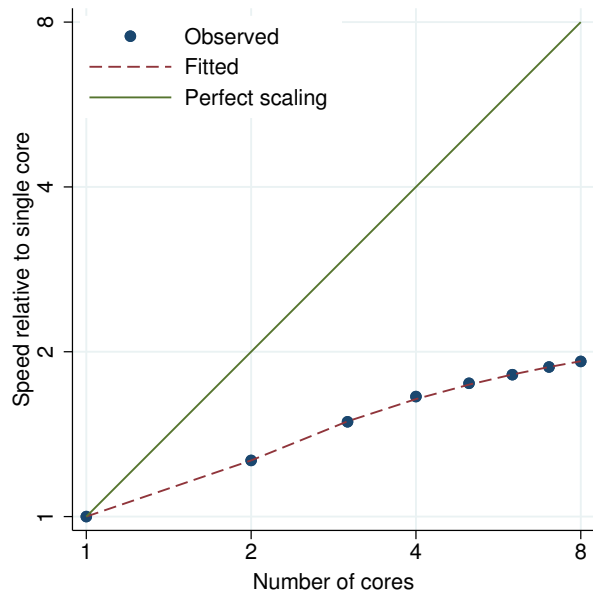


Figure 298. `misstable patterns` performance plot.

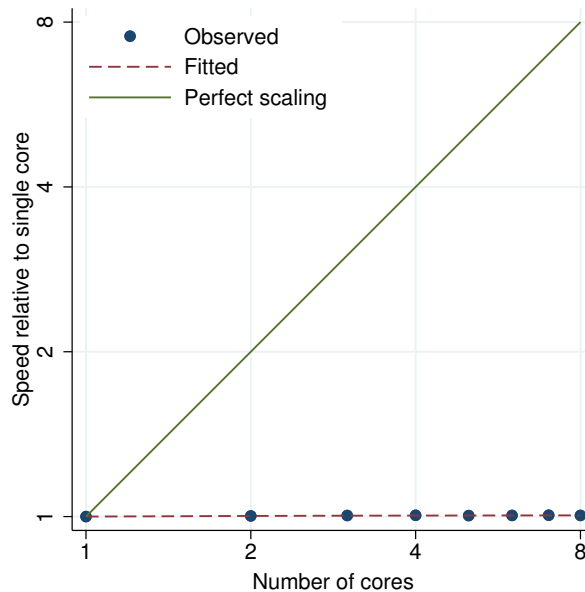
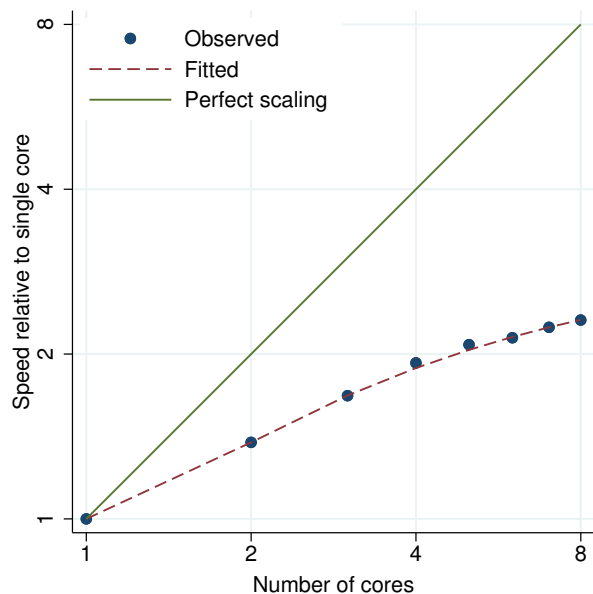
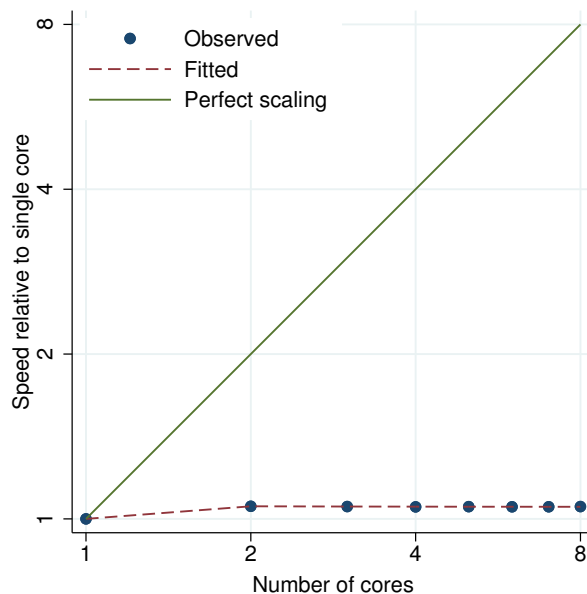
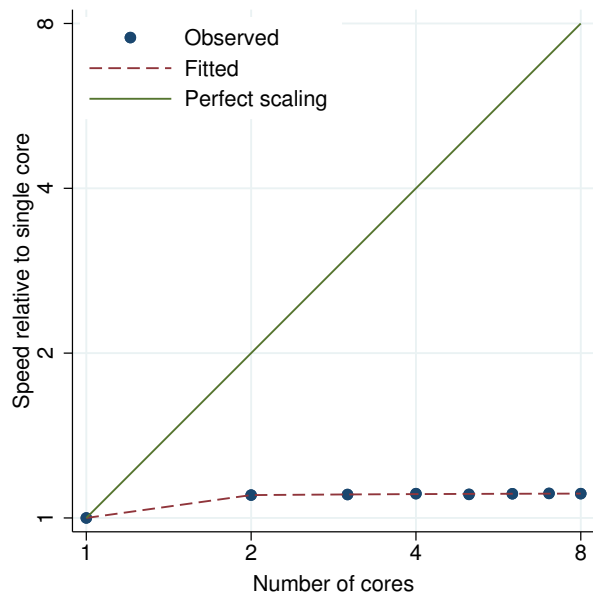
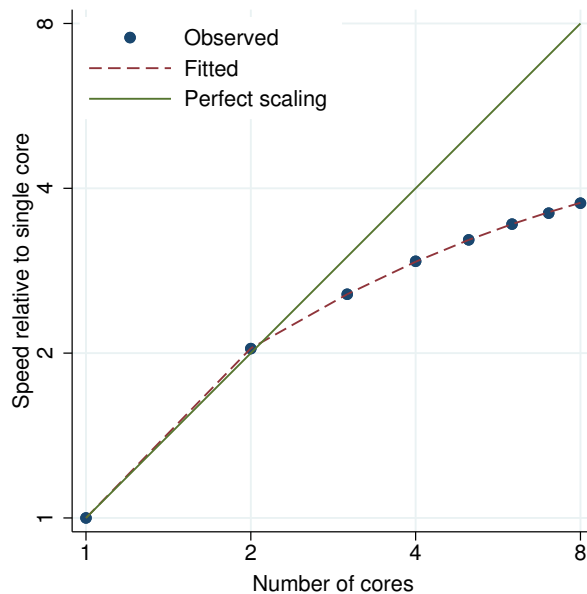


Figure 299. `misstable summarize` performance plot.

Figure 300. `misstable tree` performance plot.Figure 301. `mixed` performance plot.Figure 302. `mixed` (crossed effects) performance plot.Figure 303. `mkspine` performance plot.

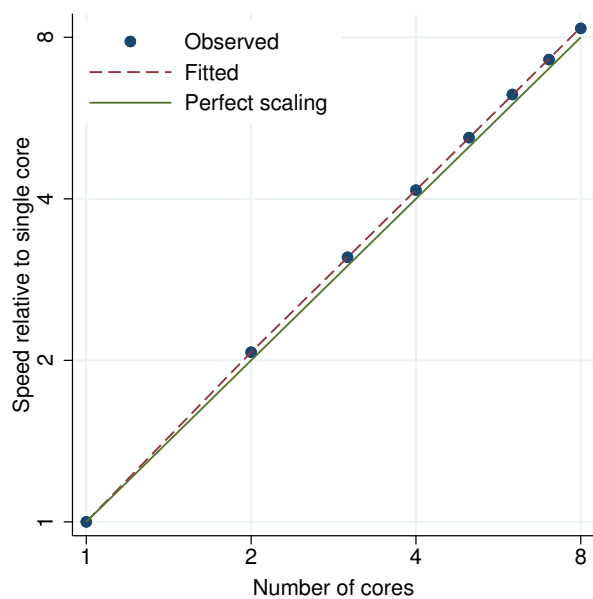


Figure 304. mlevel performance plot.

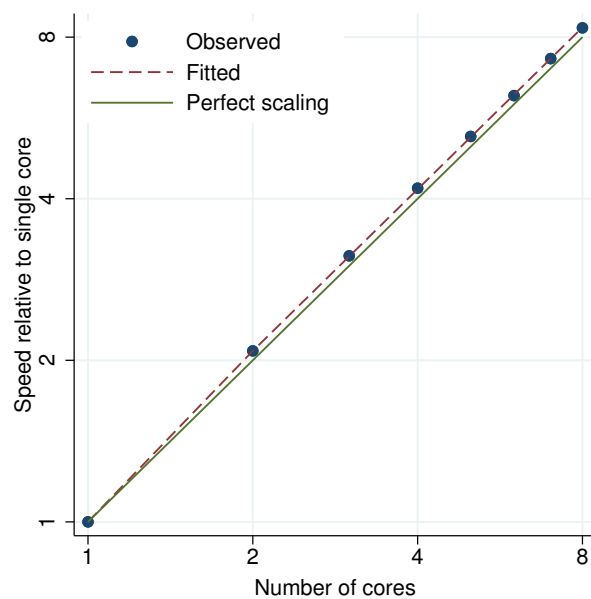


Figure 305. mlevel, nocons performance plot.

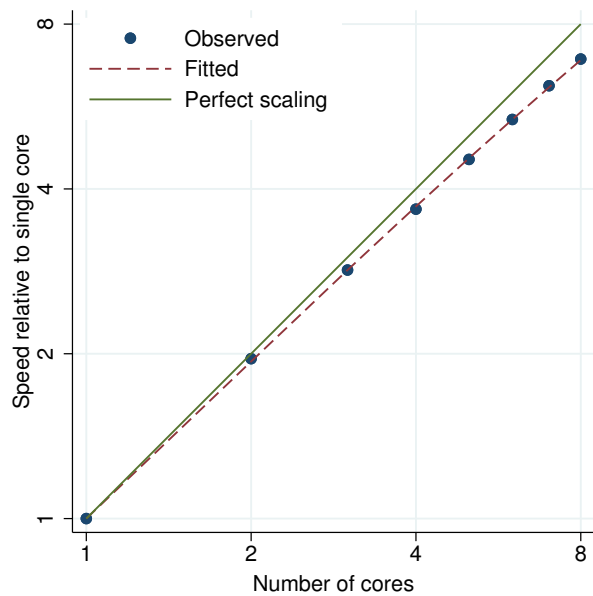


Figure 306. mlmatbysum performance plot.

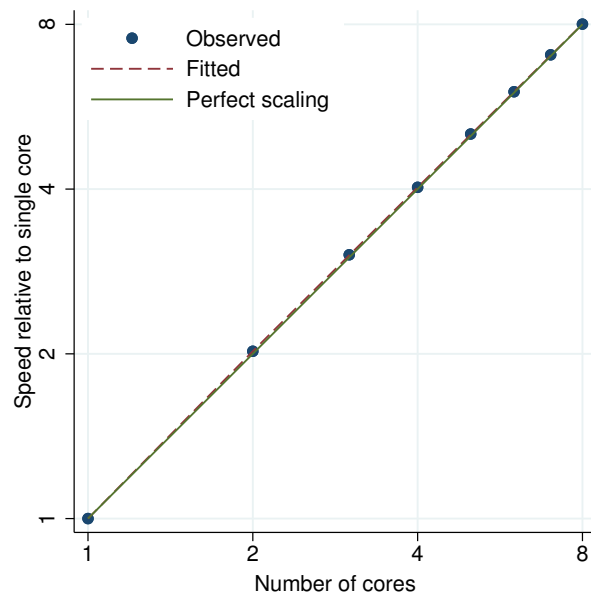


Figure 307. mlmatsum performance plot.

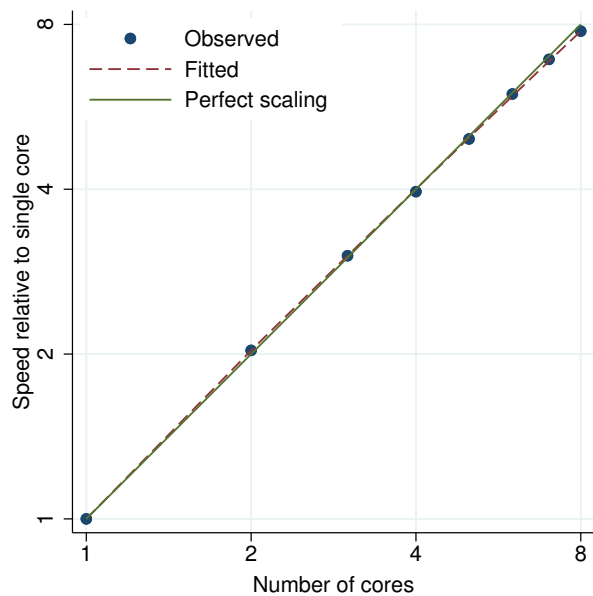


Figure 308. mlogit performance plot.

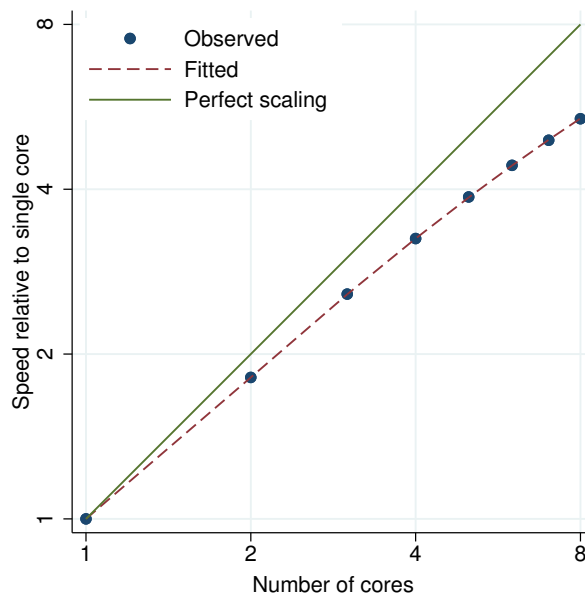


Figure 309. mlsum performance plot.

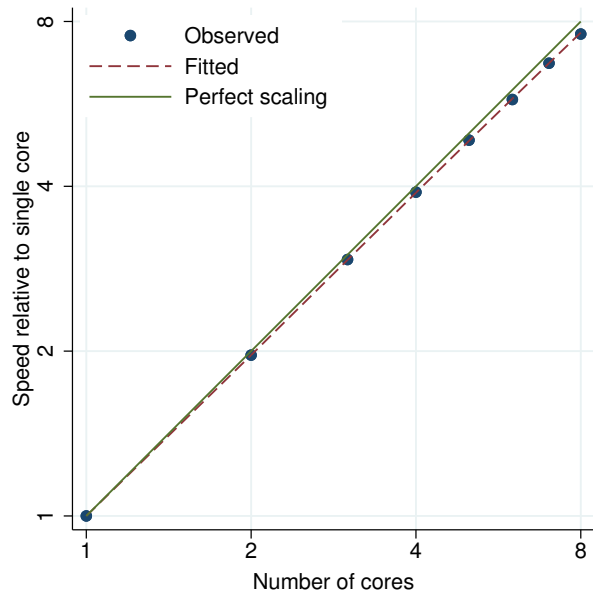


Figure 310. mlvecsum performance plot.

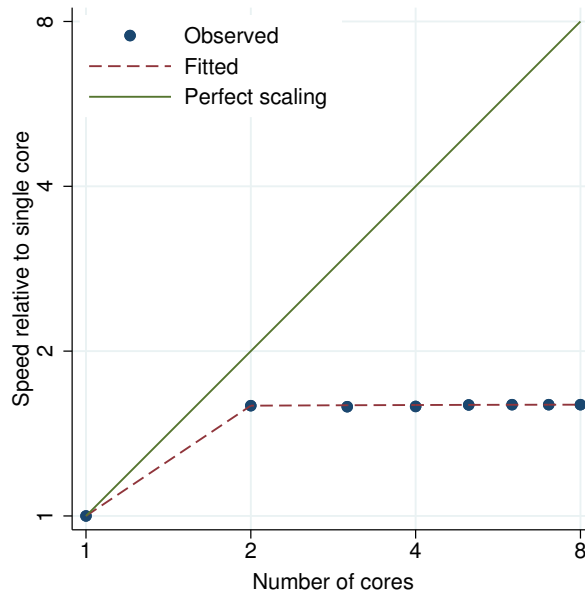
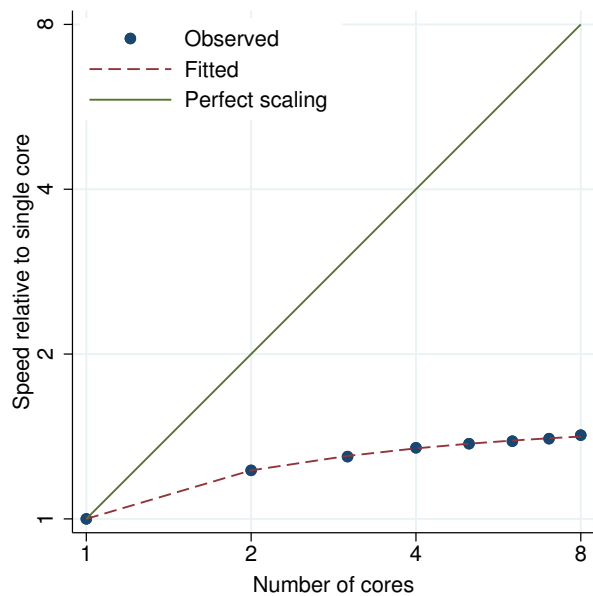
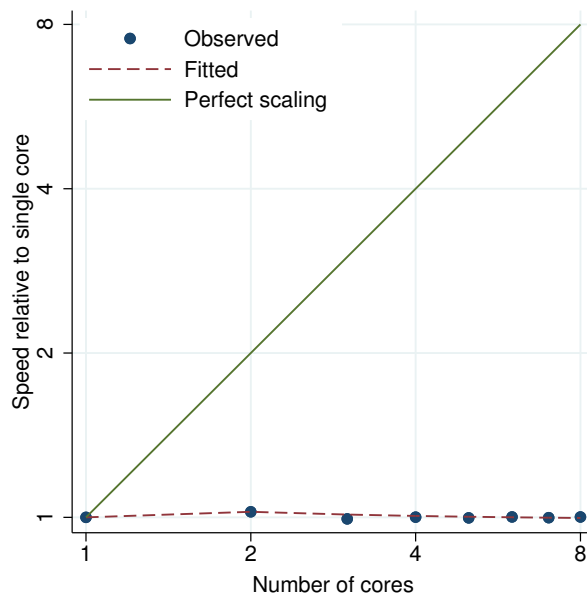
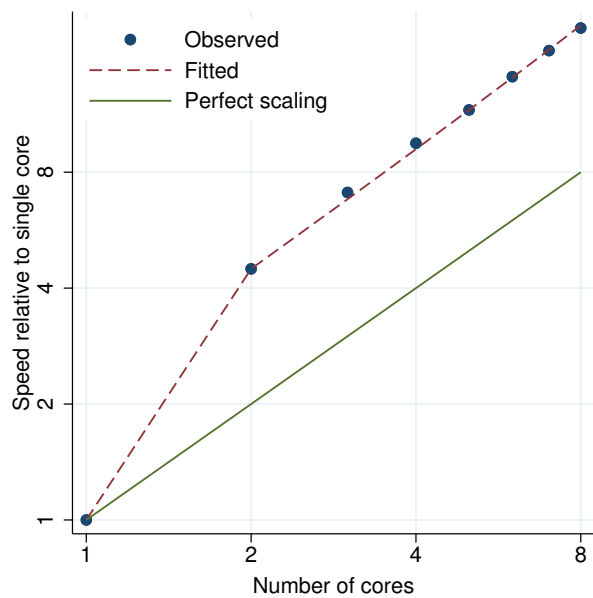
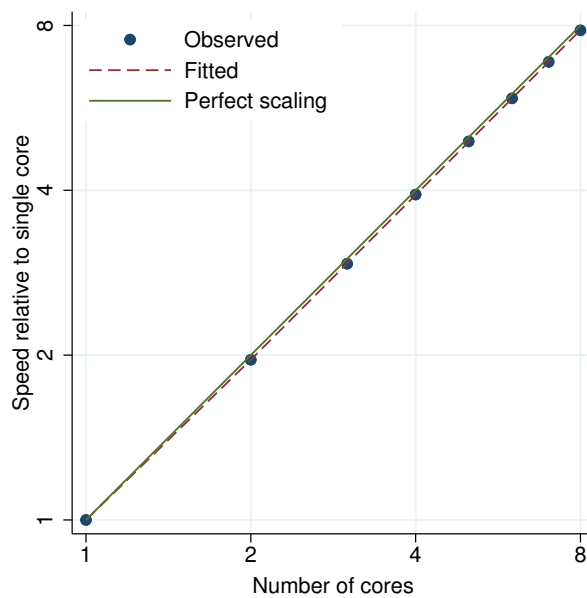


Figure 311. mprobit performance plot.

Figure 312. `mswitch ar` performance plot.Figure 313. `mswitch dr` performance plot.Figure 314. `mvdecode` performance plot.Figure 315. `mvencode` performance plot.



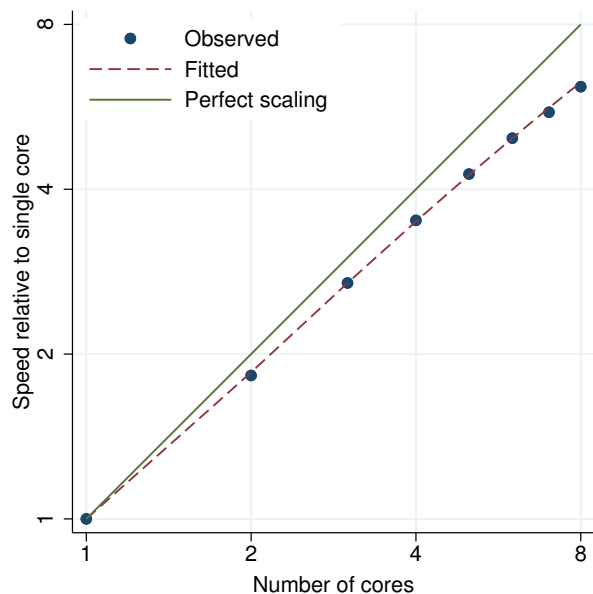


Figure 316. mvreg performance plot.

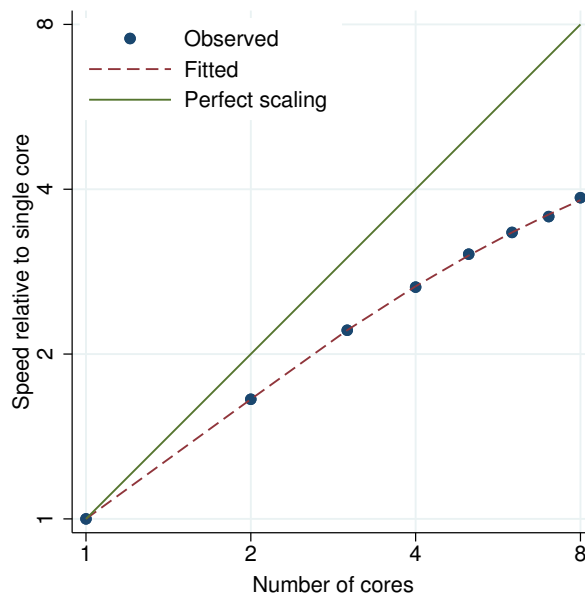


Figure 317. mvtest correlations performance plot.

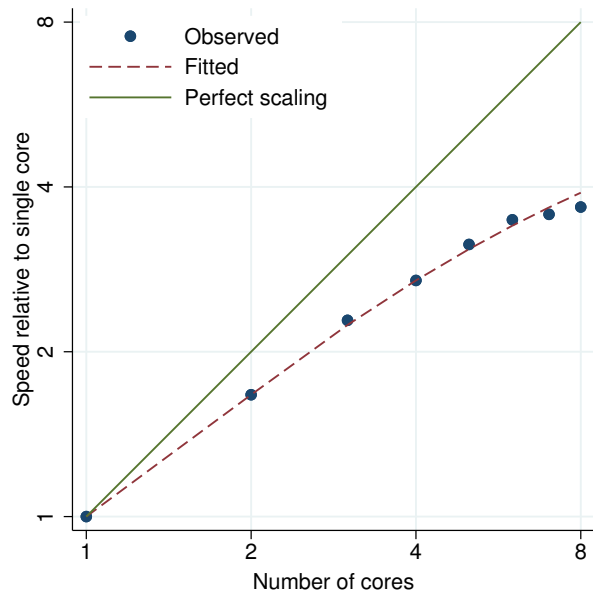


Figure 318. mvtest covariances performance plot.

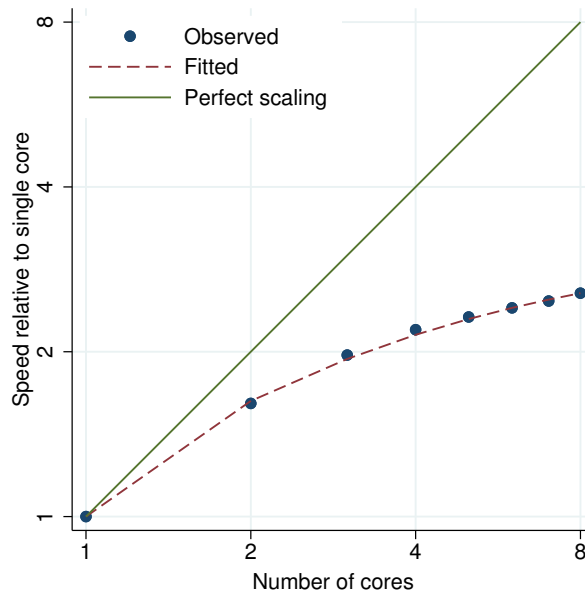


Figure 319. mvtest means, heterogeneous performance plot.

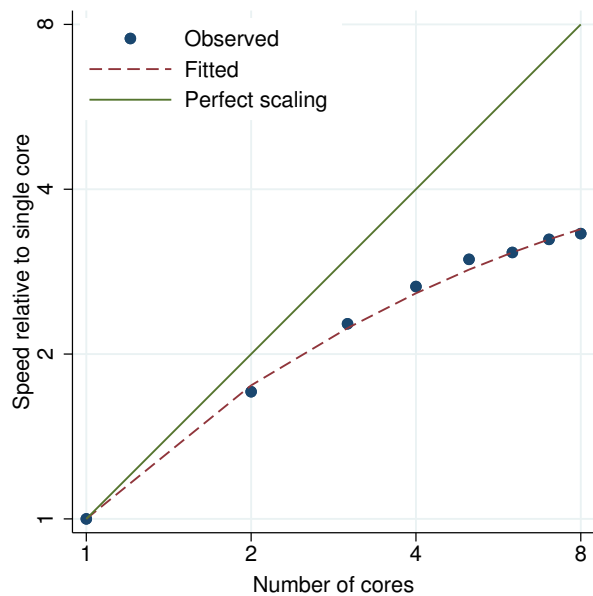


Figure 320. `mvtest` means, homogeneous performance plot.

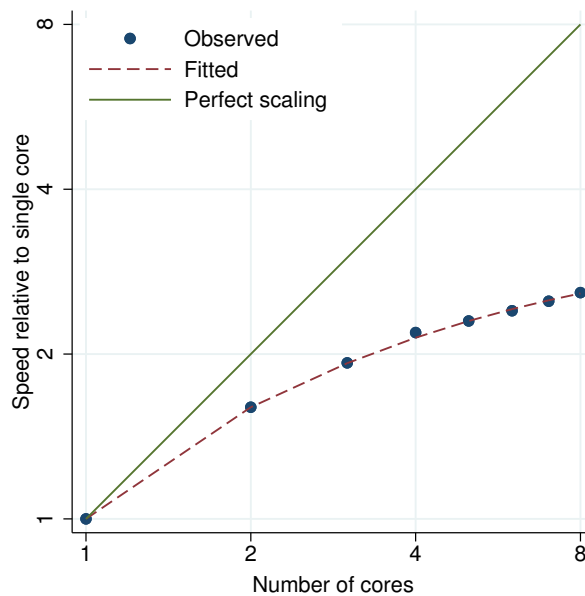


Figure 321. `mvtest` means, lr performance plot.

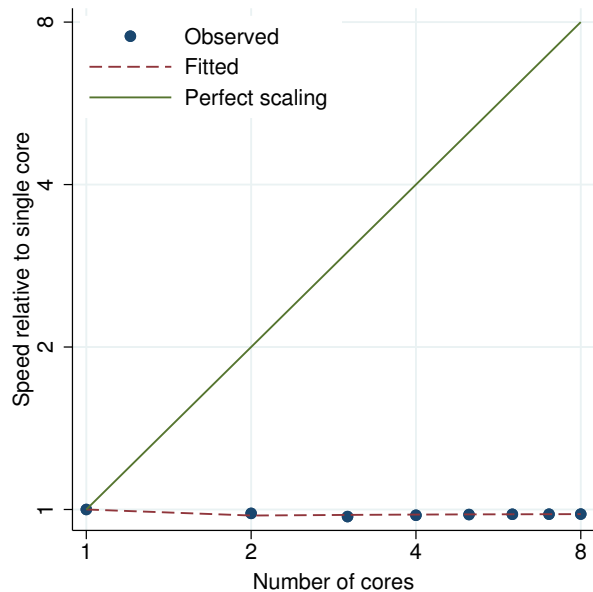


Figure 322. `mvtest` normality performance plot.

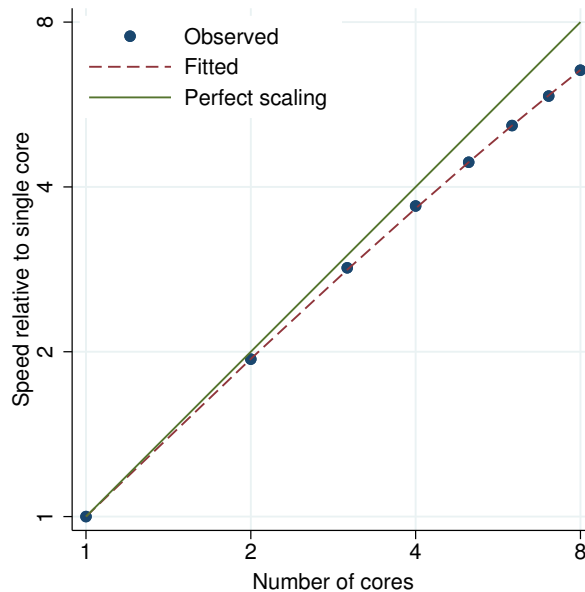


Figure 323. `nbreg` performance plot.

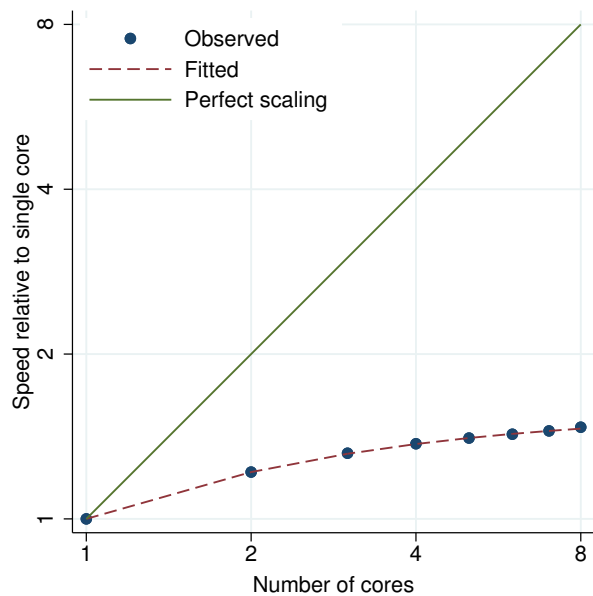


Figure 324. newey performance plot.

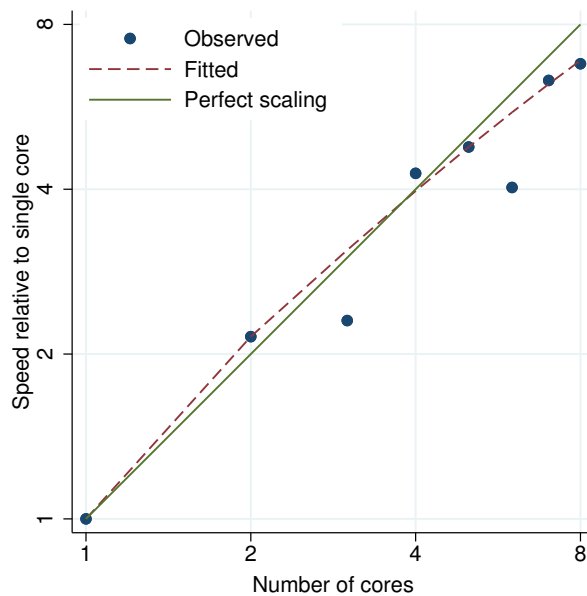


Figure 325. nl performance plot.

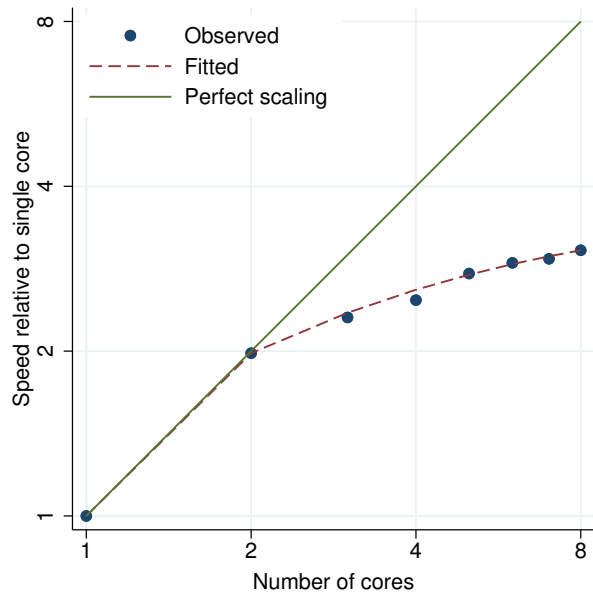


Figure 326. nlogit performance plot.

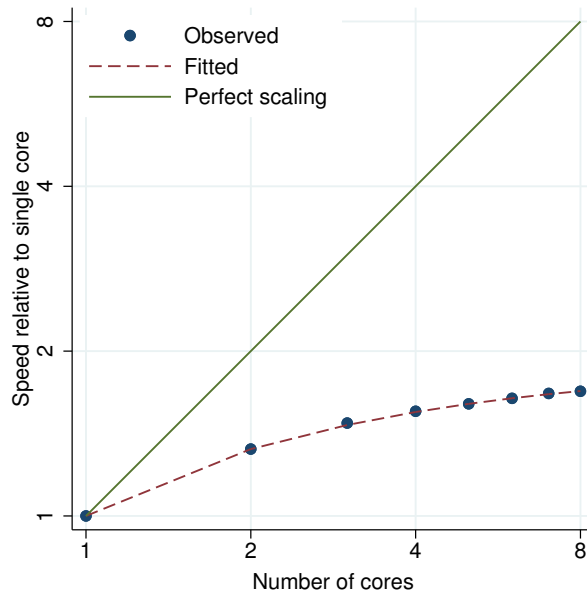
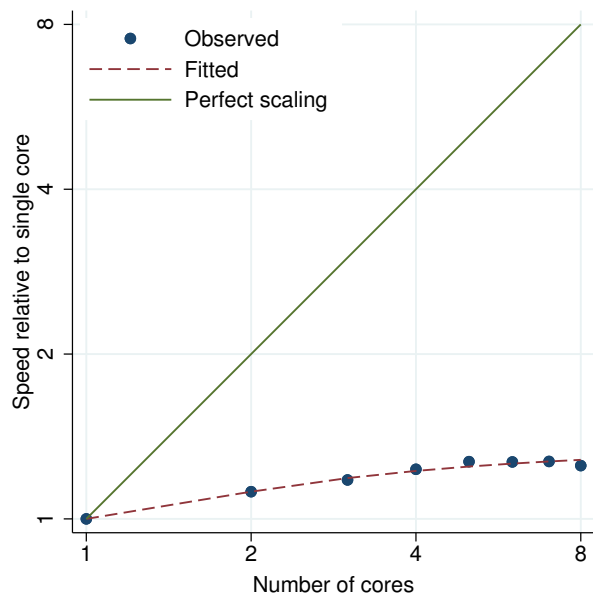
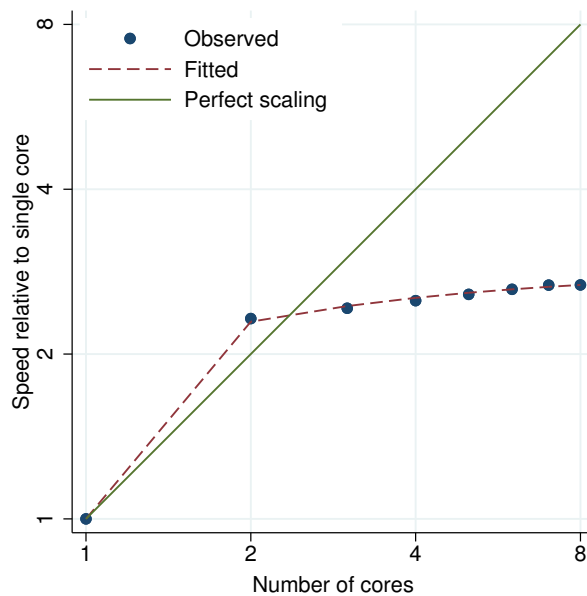
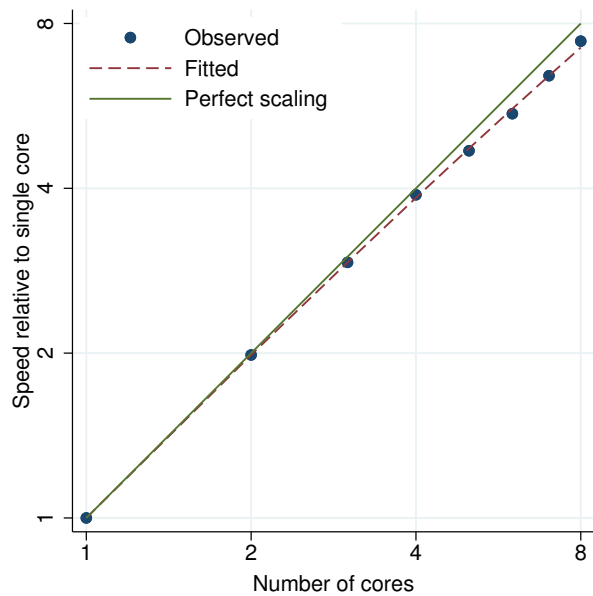
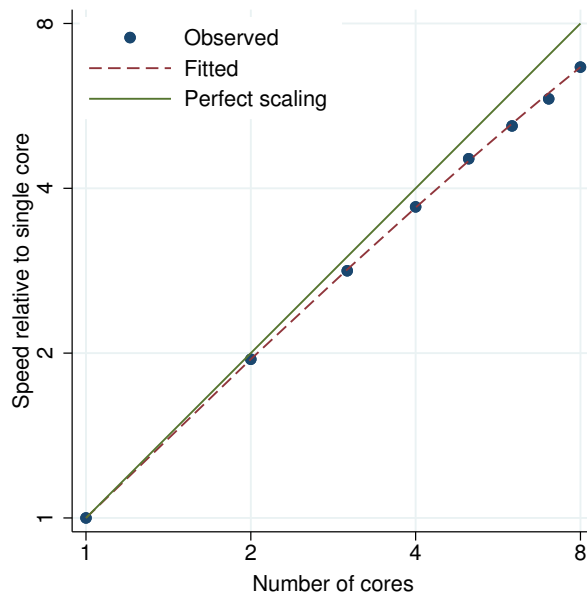


Figure 327. nlstur performance plot.

Figure 328. `npregress` kernel performance plot.Figure 329. `nptrend` performance plot.Figure 330. `ologit` performance plot.Figure 331. `ologit`, `vce(cluster)` performance plot.

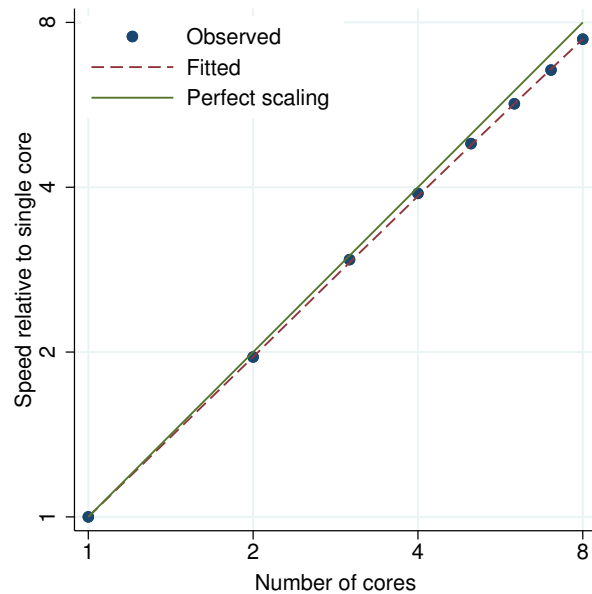


Figure 332. `ologit, vce(robust)` performance plot.

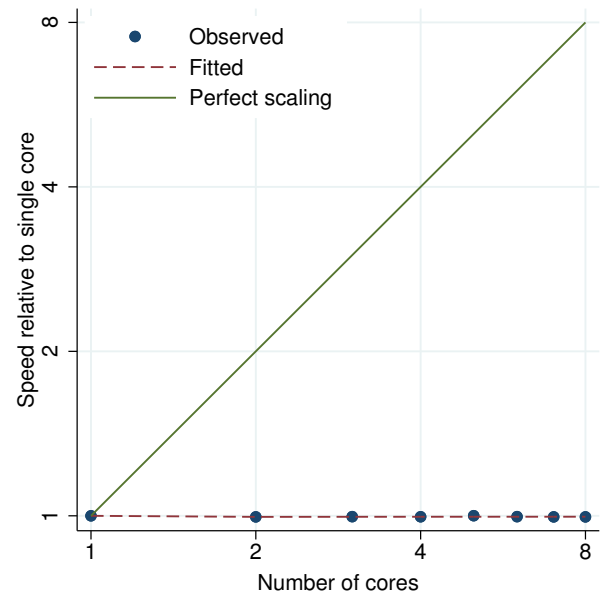


Figure 333. `oneway` performance plot.

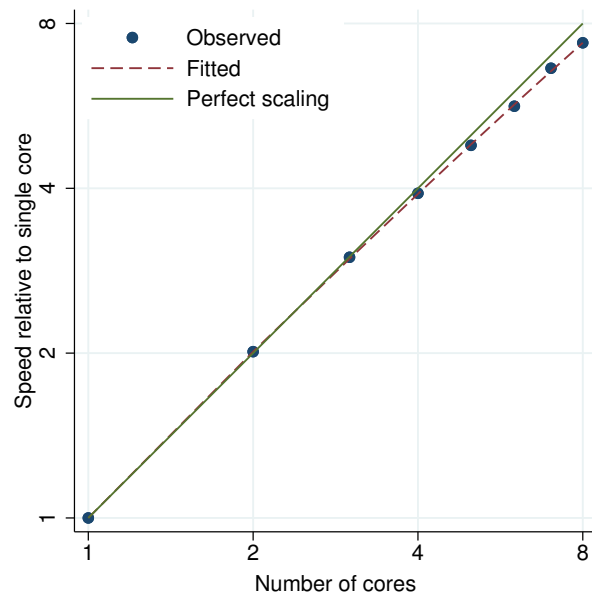


Figure 334. `oprobit` performance plot.

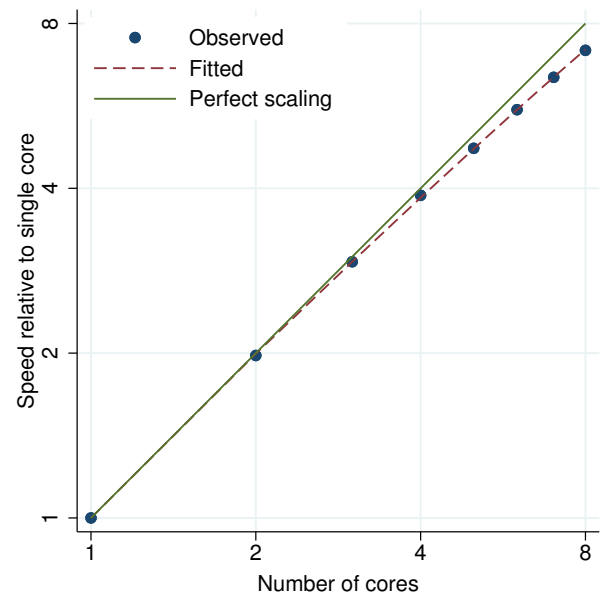


Figure 335. `oprobit, vce(cluster)` performance plot.

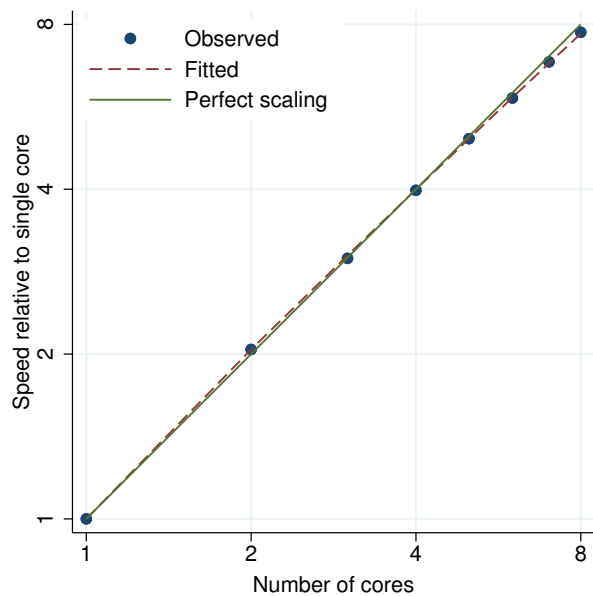


Figure 336. `oprobit, vce(robust)` performance plot.

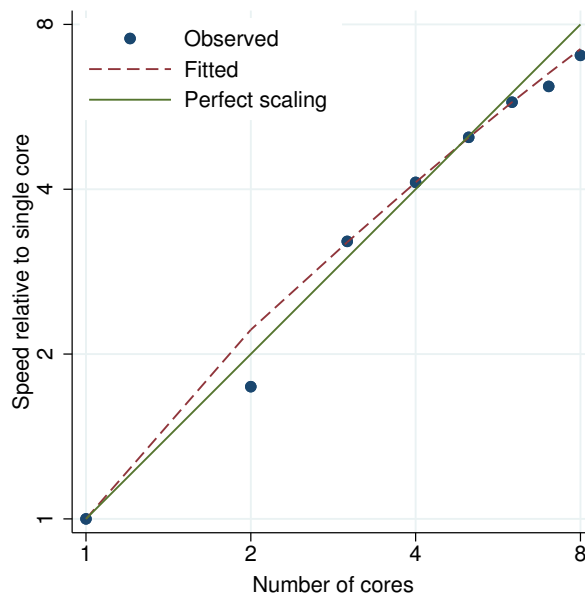


Figure 337. `orthog` performance plot.

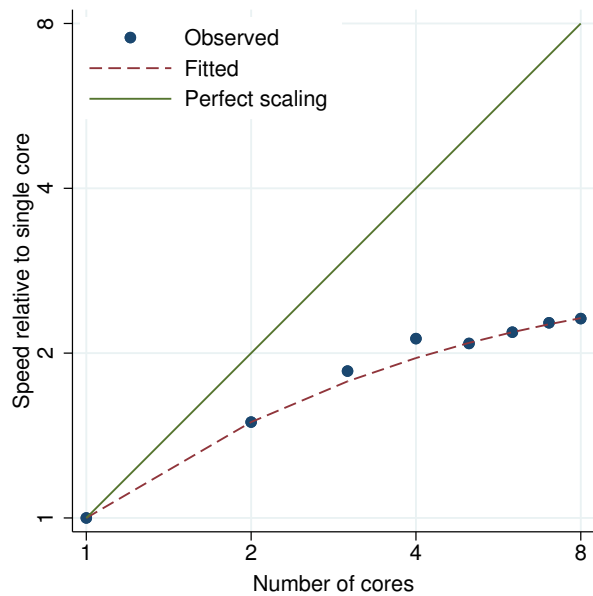


Figure 338. `pca` performance plot.

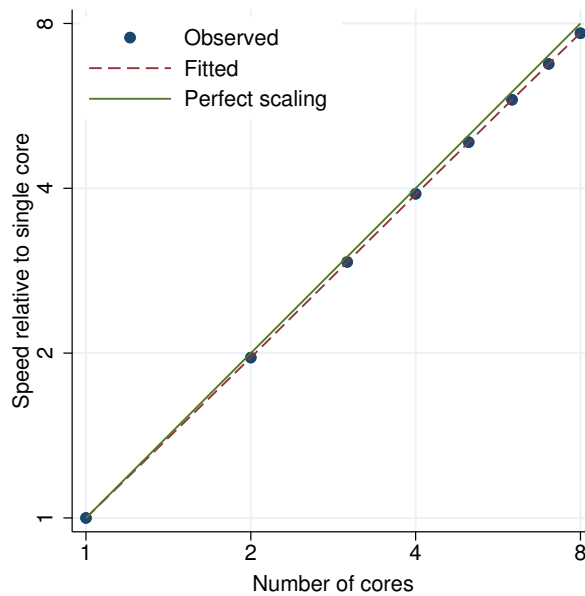
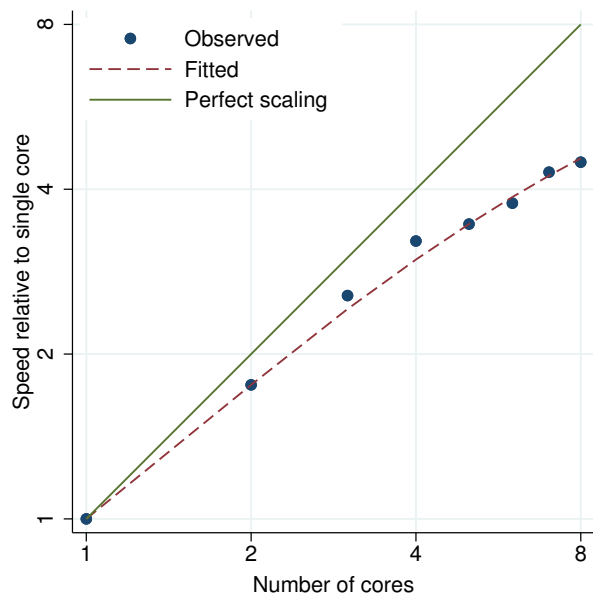
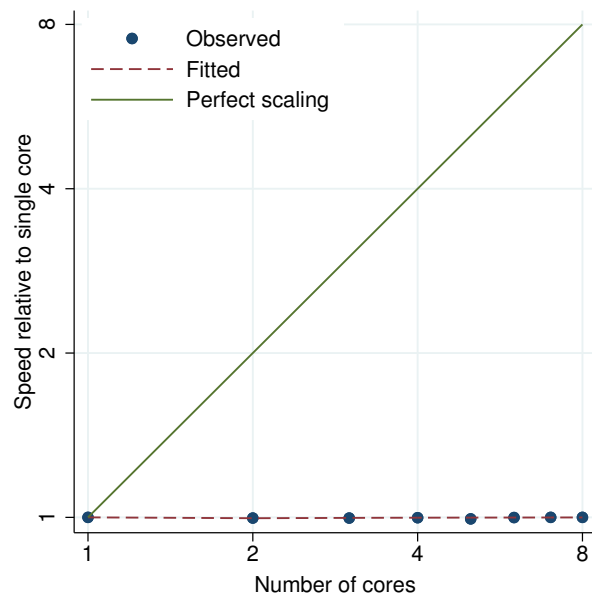
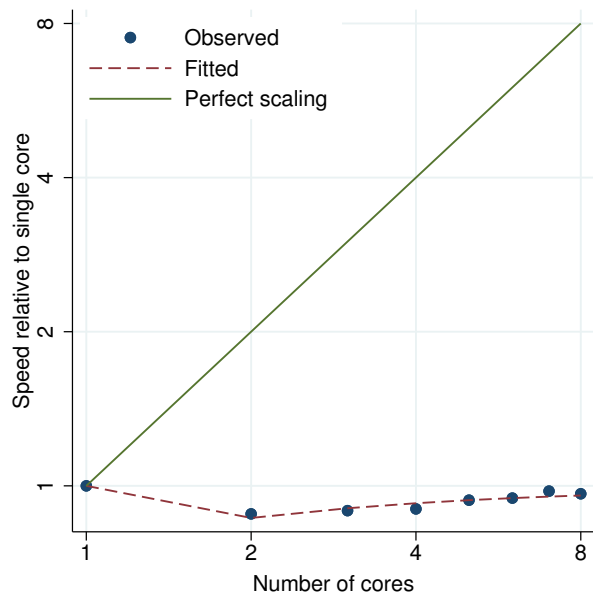
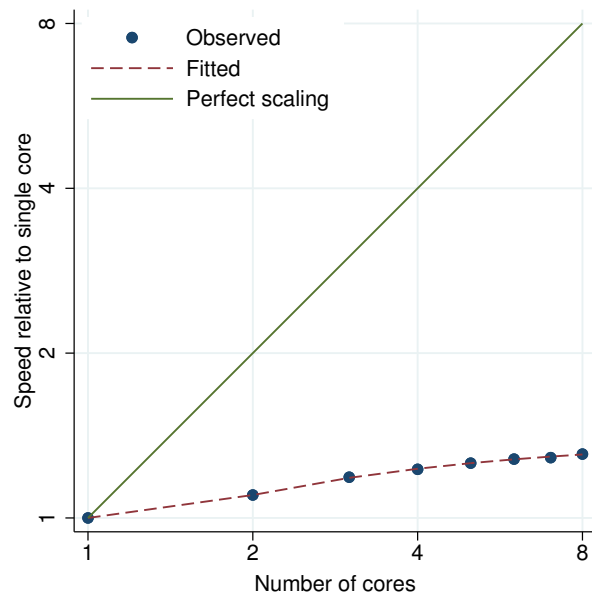
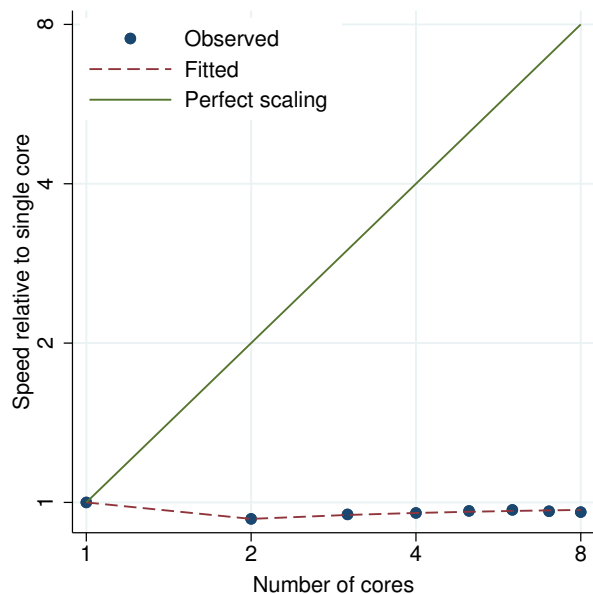
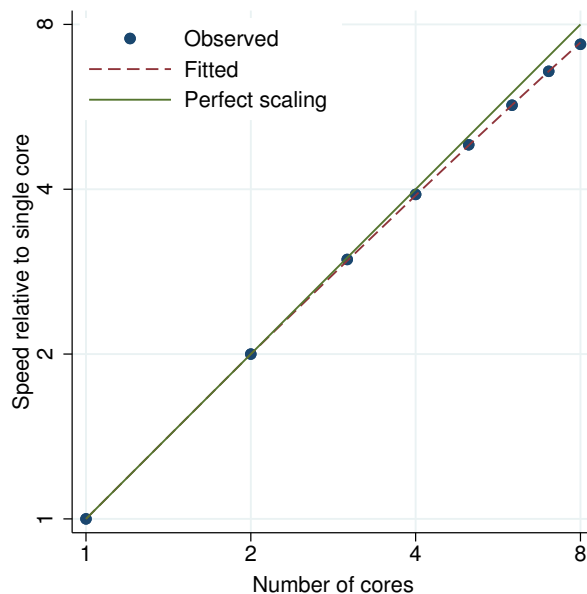
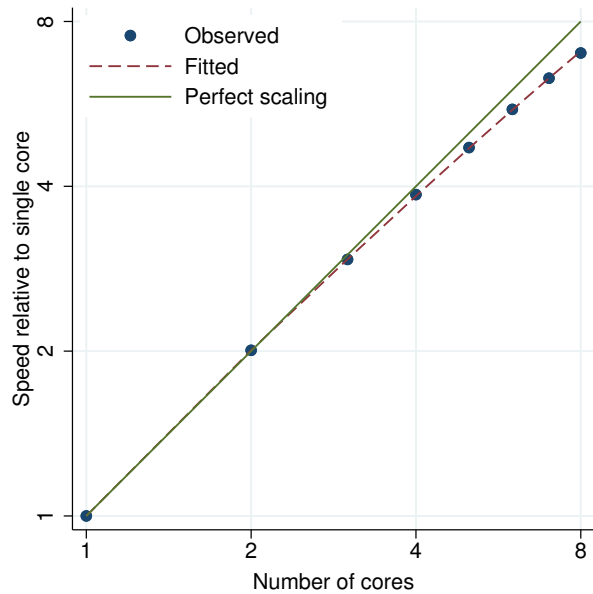
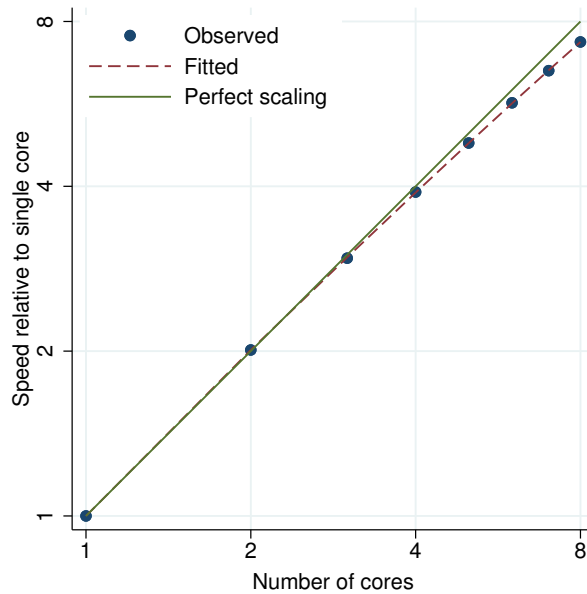
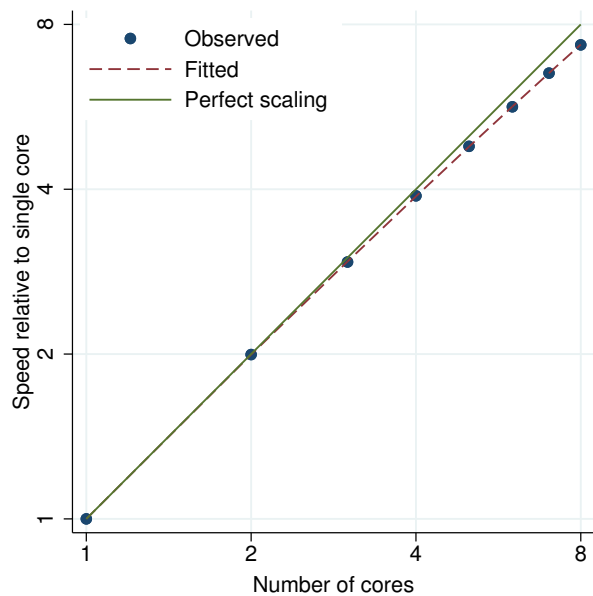
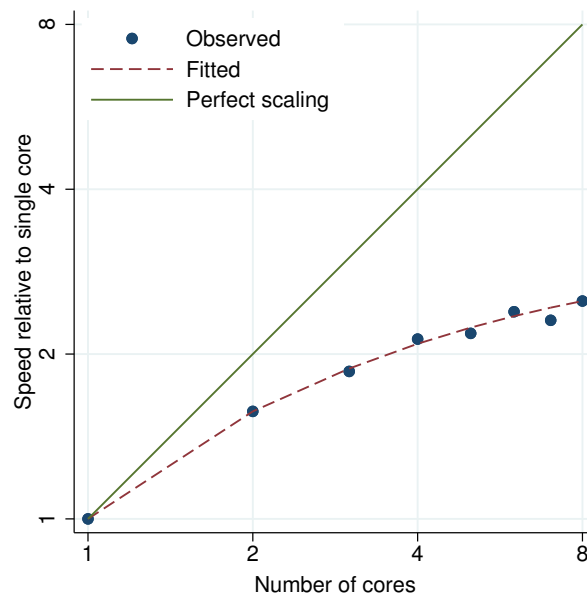
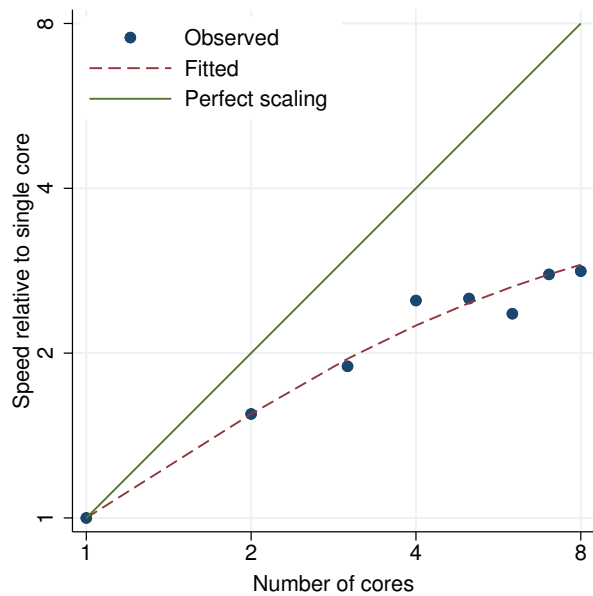
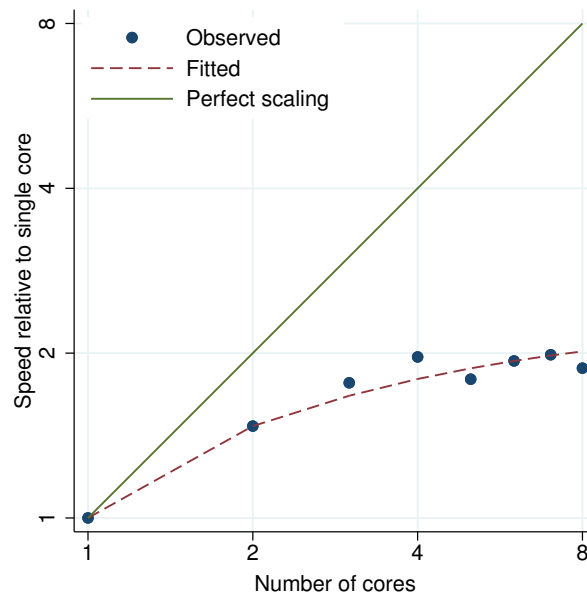


Figure 339. `pcorr` performance plot.

Figure 340. `pctile` performance plot.Figure 341. `pergram` performance plot.Figure 342. `pkcollapse` performance plot.Figure 343. `pkexamine` performance plot.

Figure 344. `pksumm` performance plot.Figure 345. `poisson` performance plot.Figure 346. `poisson, vce(cluster)` performance plot.Figure 347. `poisson, exposure()` performance plot.



Figure 348. `poisson, vce(robust)` performance plot.Figure 349. `pologit` performance plot.Figure 350. `popoisson` performance plot.Figure 351. `poregress` performance plot.

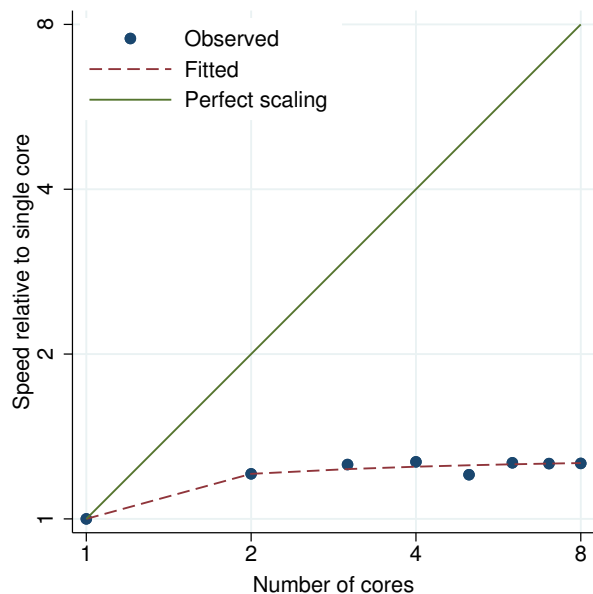


Figure 352. pperron performance plot.

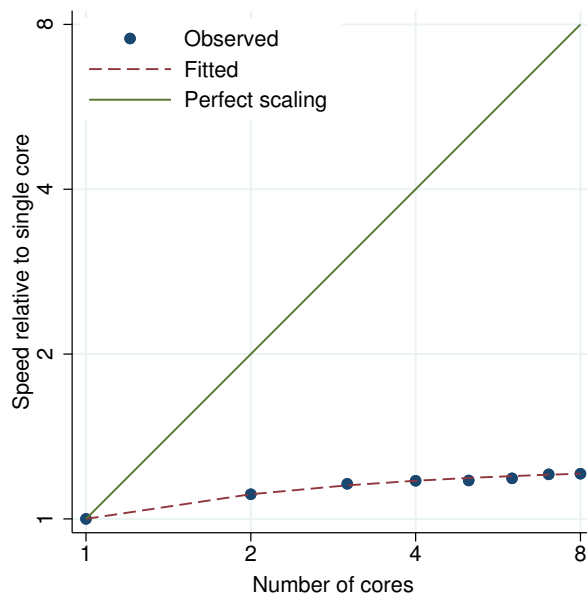


Figure 353. prais performance plot.

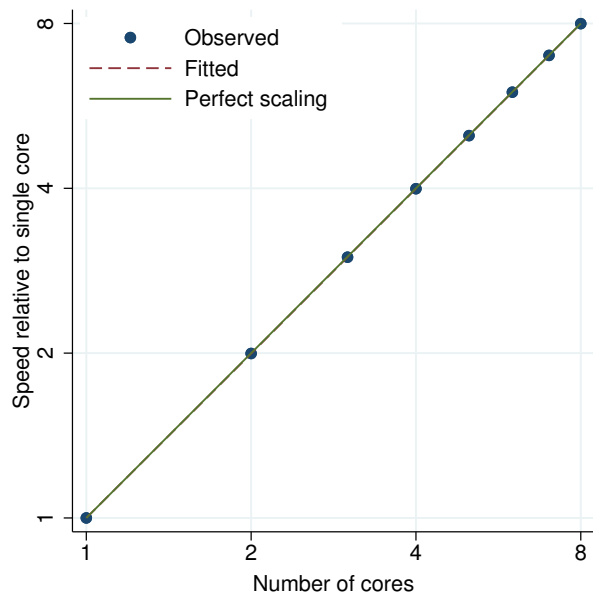


Figure 354. predict, cooksd performance plot.

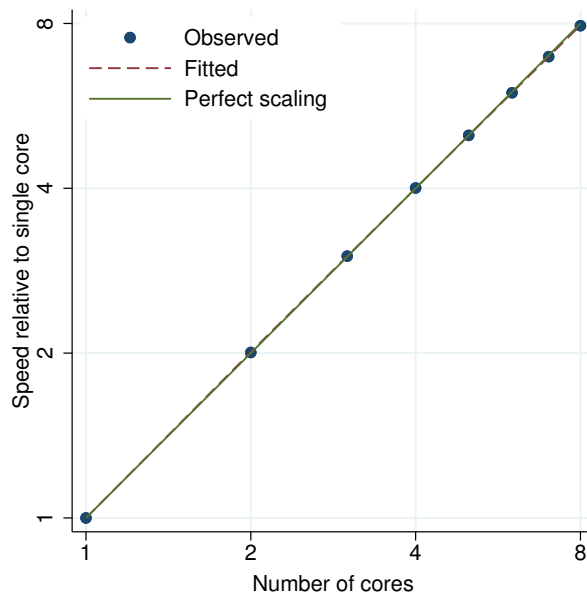
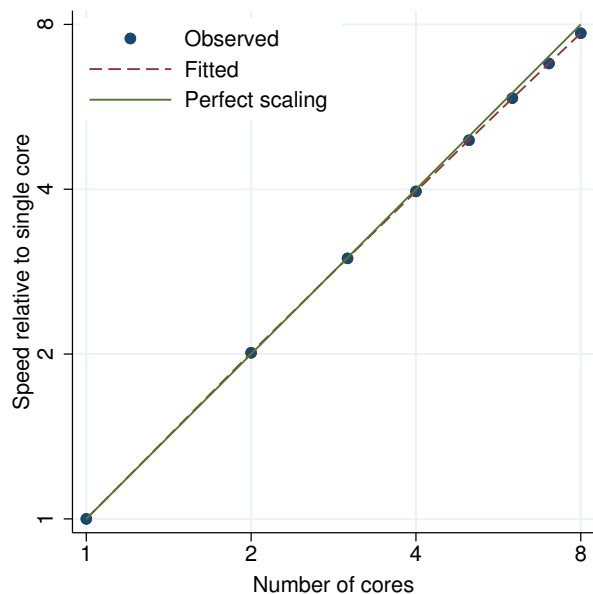
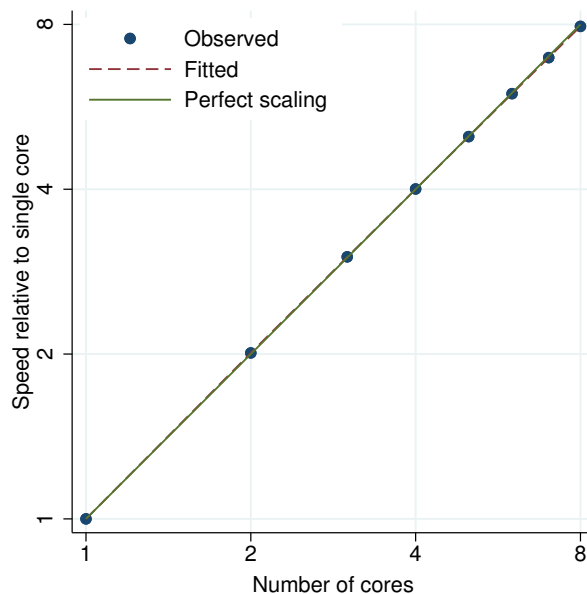
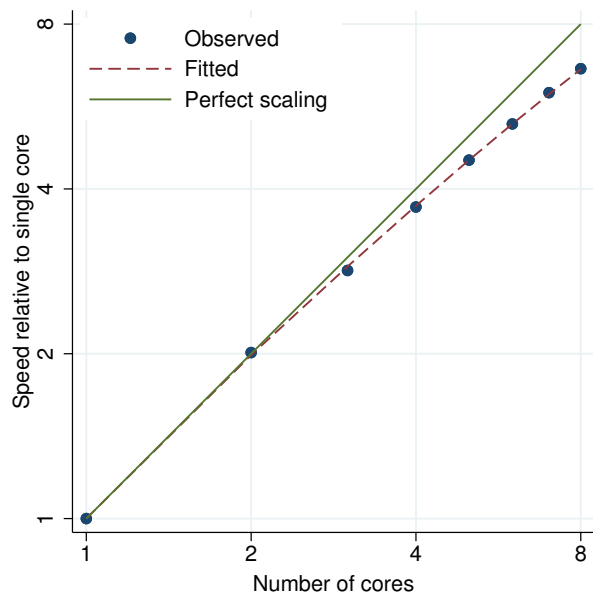
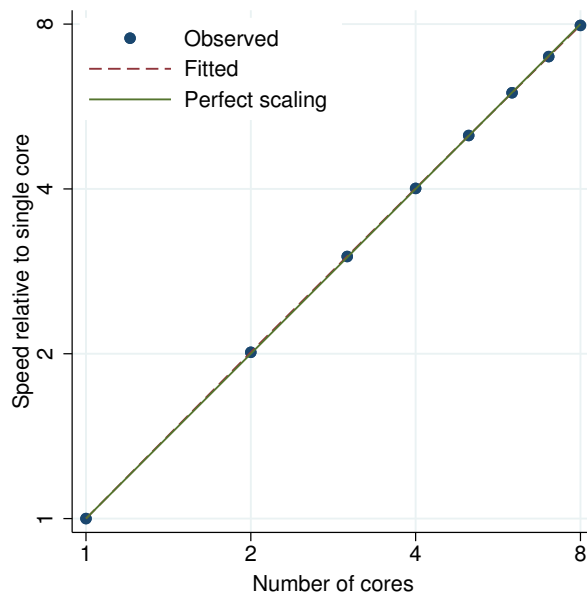


Figure 355. predict, covratio performance plot.

Figure 356. `predict, dfbeta` performance plot.Figure 357. `predict, dfits` performance plot.Figure 358. `predict, e` performance plot.Figure 359. `predict, leverage` performance plot.

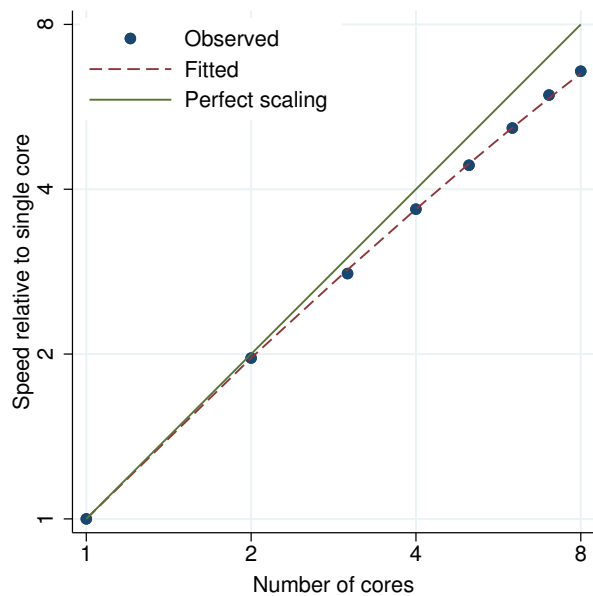


Figure 360. predict, pr performance plot.

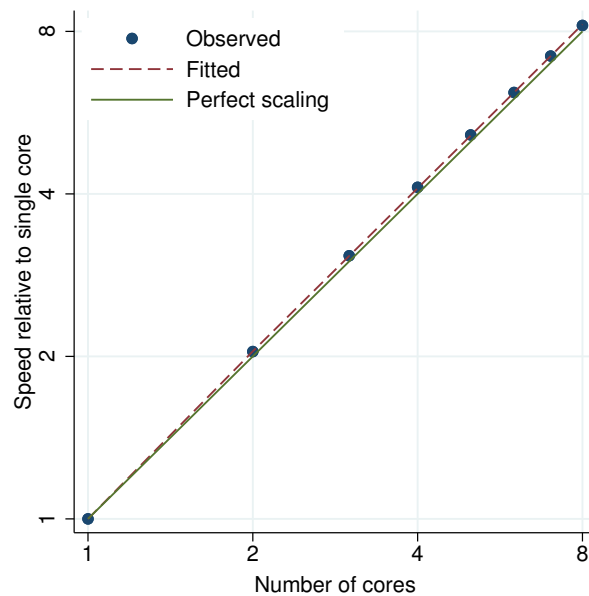


Figure 361. predict, residuals performance plot.

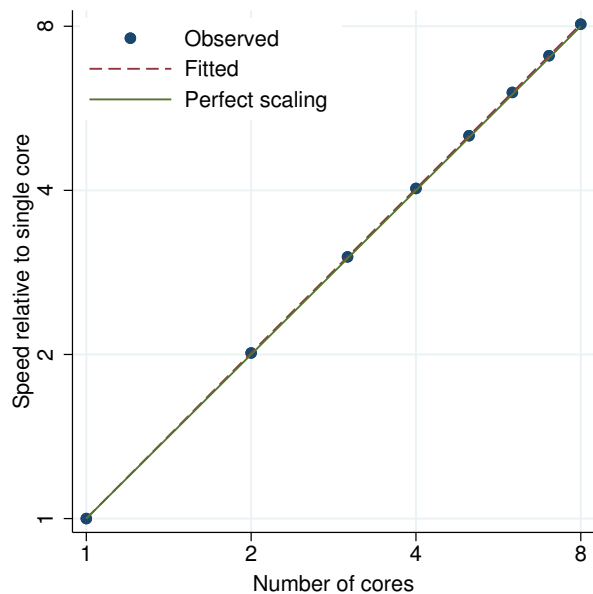


Figure 362. predict, rstandard performance plot.

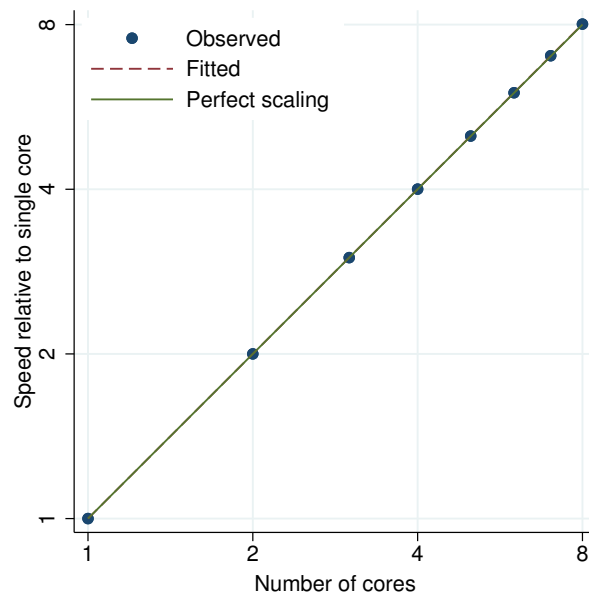


Figure 363. predict, rstudent performance plot.

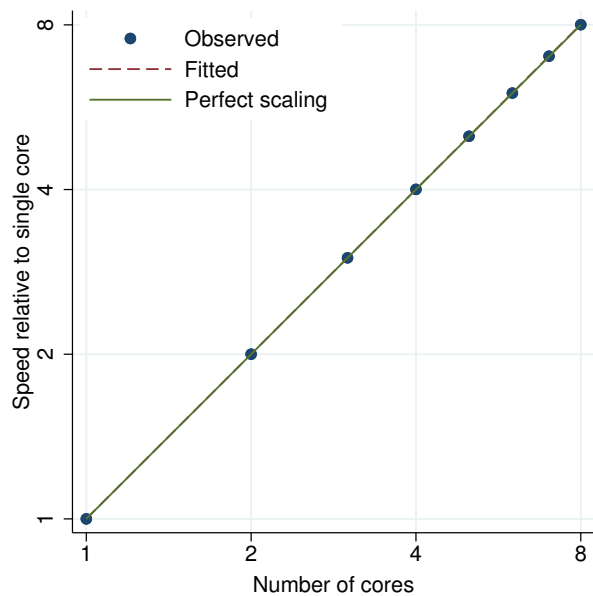


Figure 364. predict, stdf performance plot.

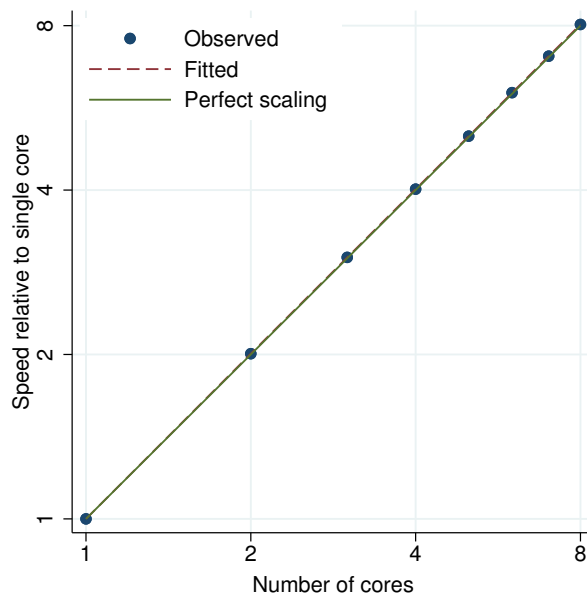


Figure 365. predict, stdp performance plot.

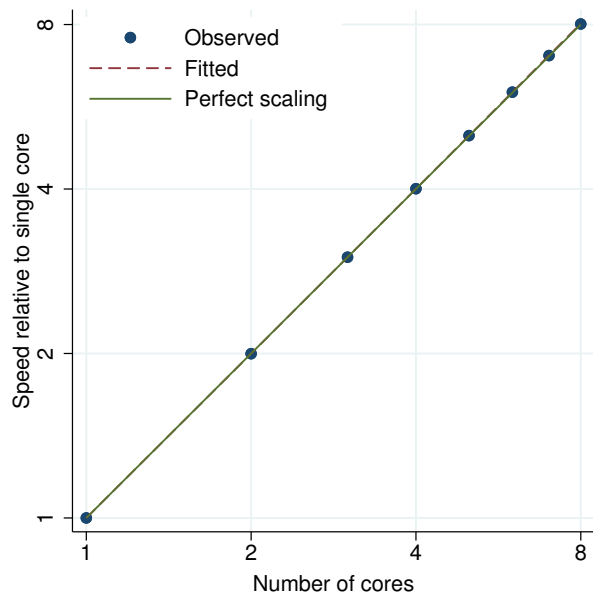


Figure 366. predict, stdr performance plot.

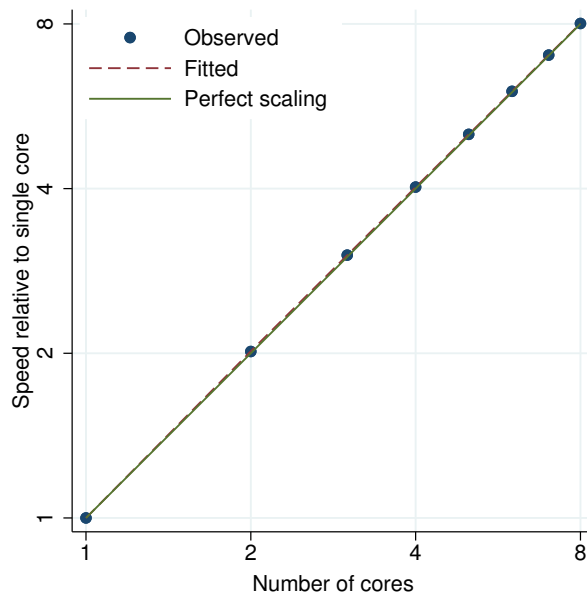
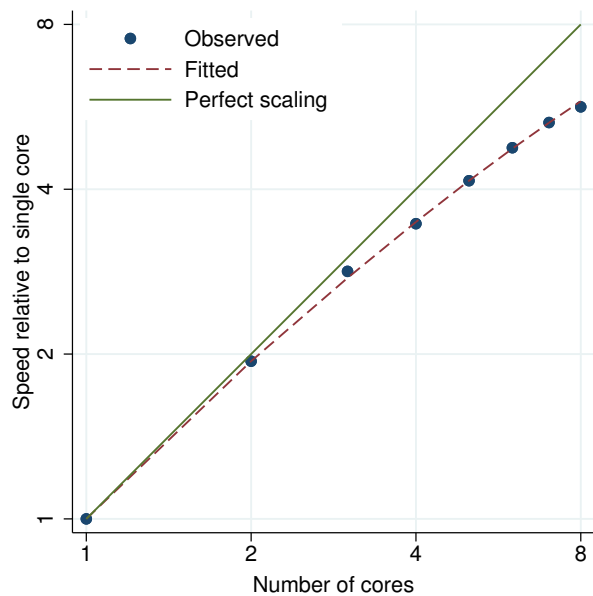
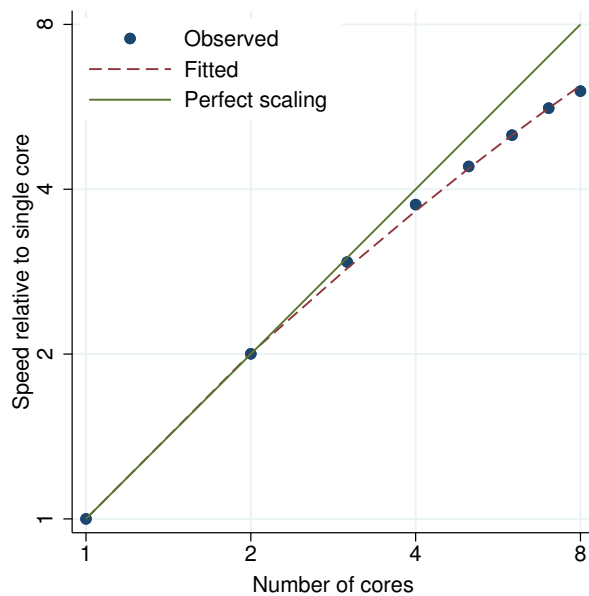
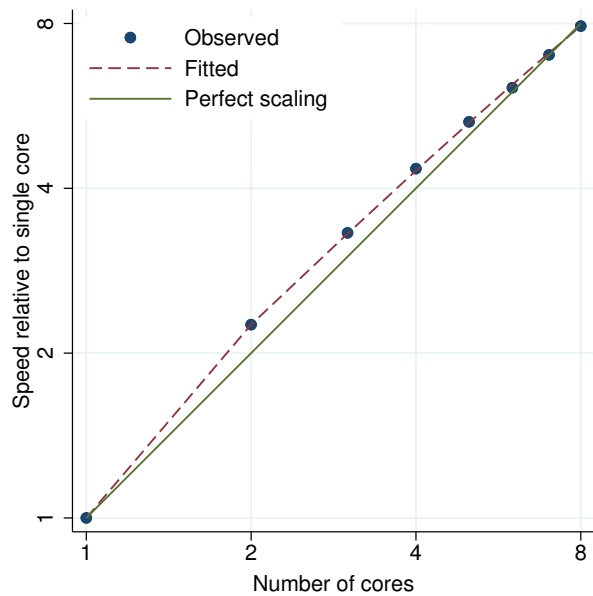
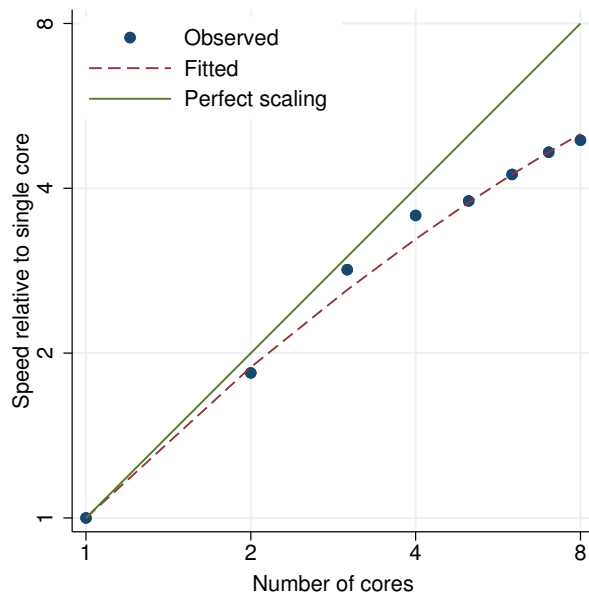


Figure 367. predict, welsch performance plot.

Figure 368. `predict, ystar` performance plot.Figure 369. `predictnl` performance plot.Figure 370. `probit` performance plot.Figure 371. `procrustes` performance plot.

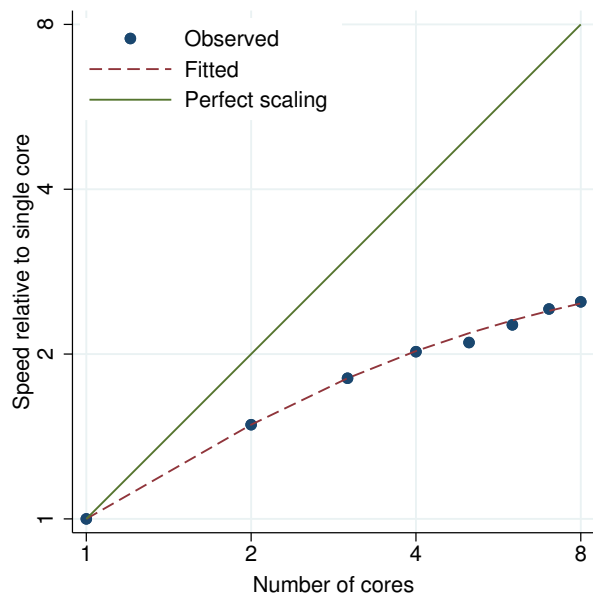


Figure 372. proportion performance plot.

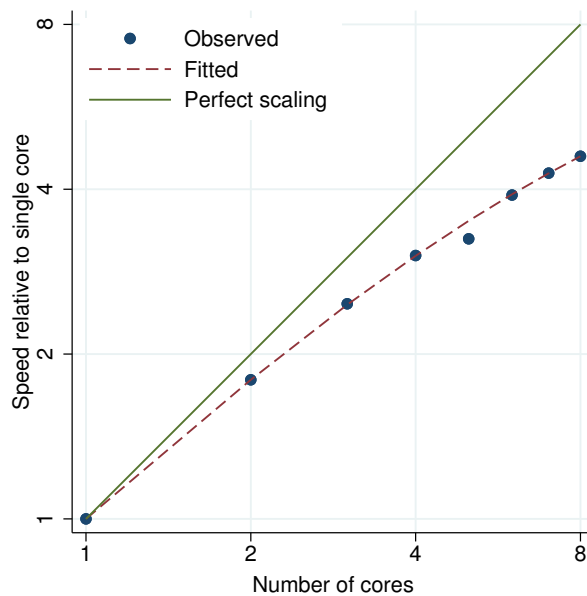


Figure 373. prtest1 performance plot.

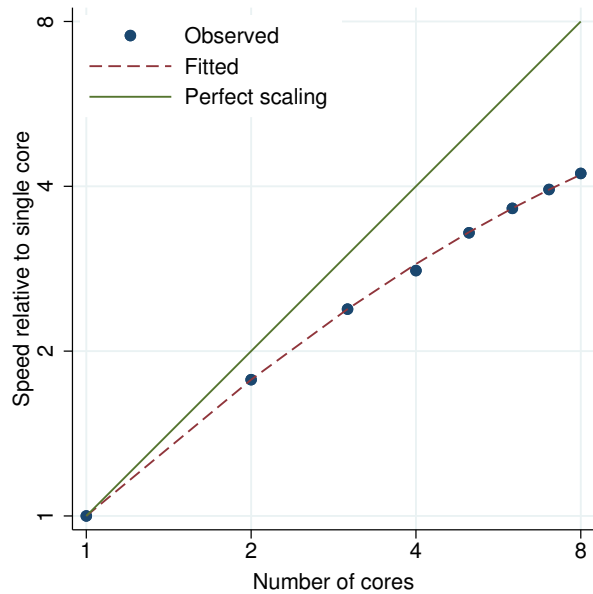


Figure 374. prtest2 performance plot.

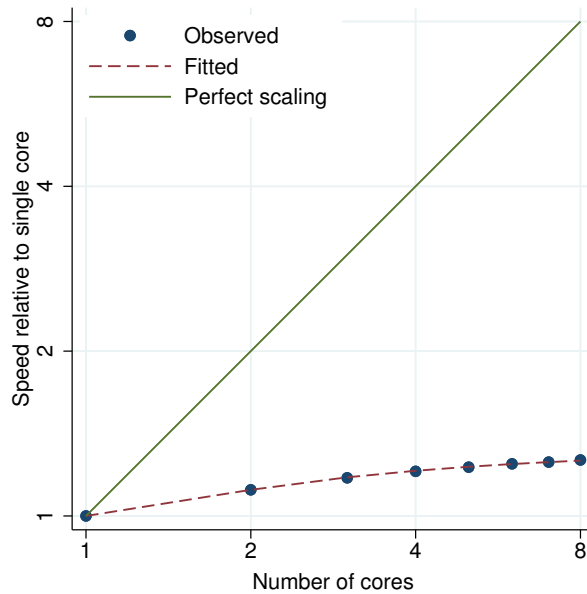
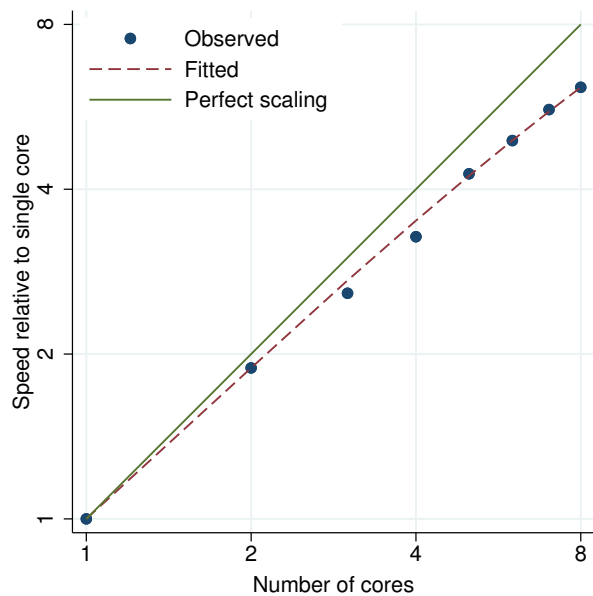
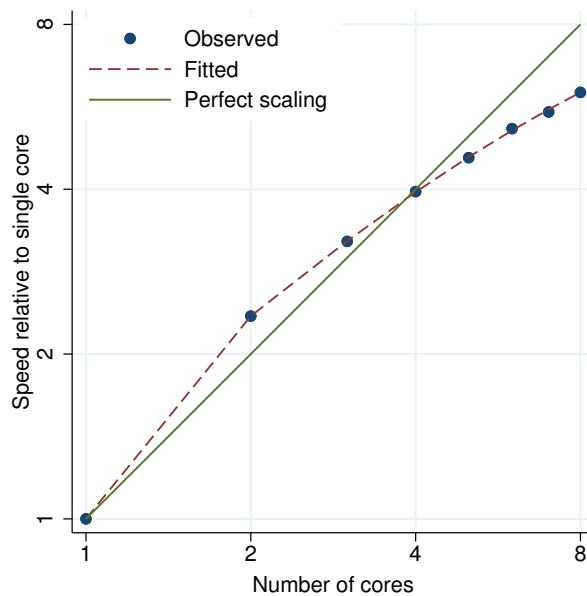
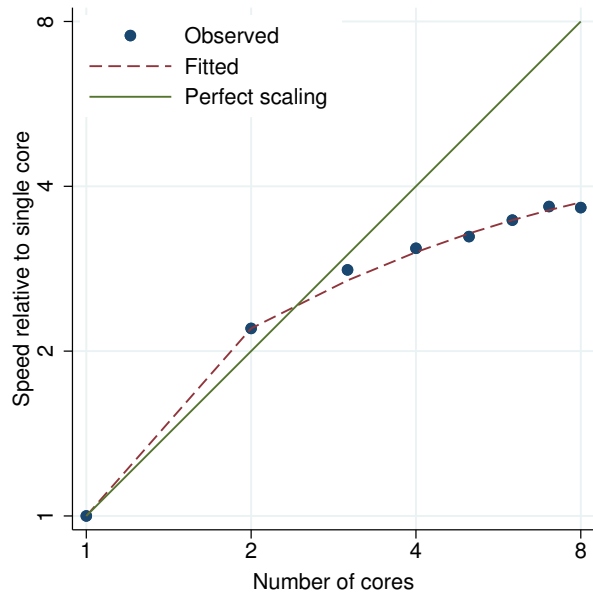
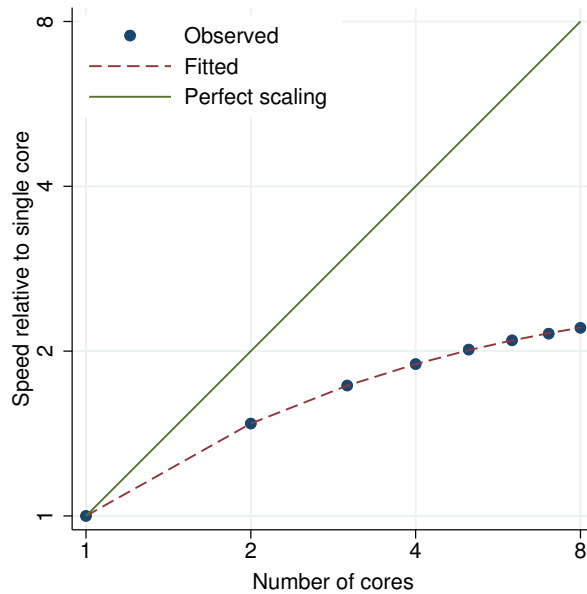


Figure 375. prtest, by() performance plot.

Figure 376. `pwcorr` performance plot.Figure 377. `qreg` performance plot.Figure 378. `ranksum` performance plot.Figure 379. `ratio` performance plot.



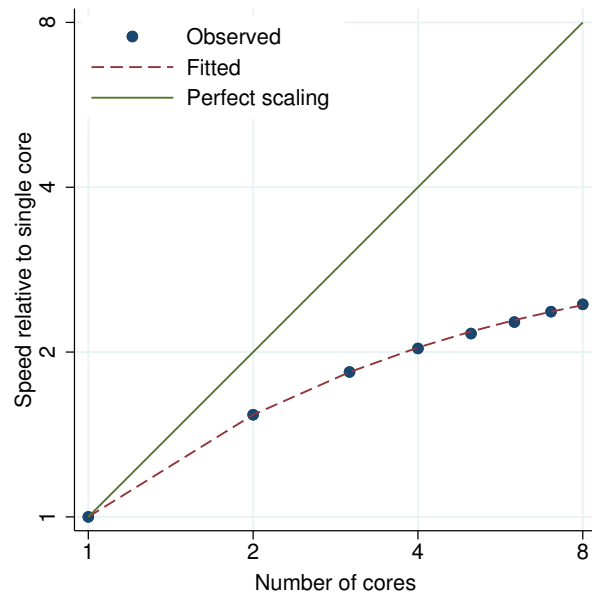


Figure 380. `ratio (exp1) (exp2)` performance plot.

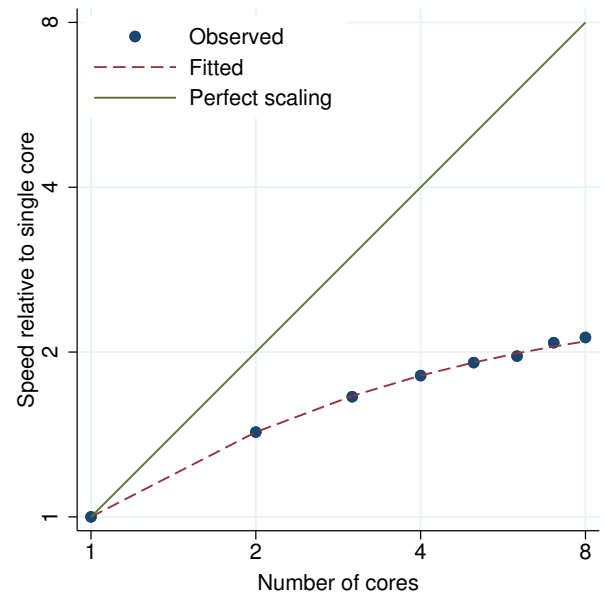


Figure 381. `recode` performance plot.

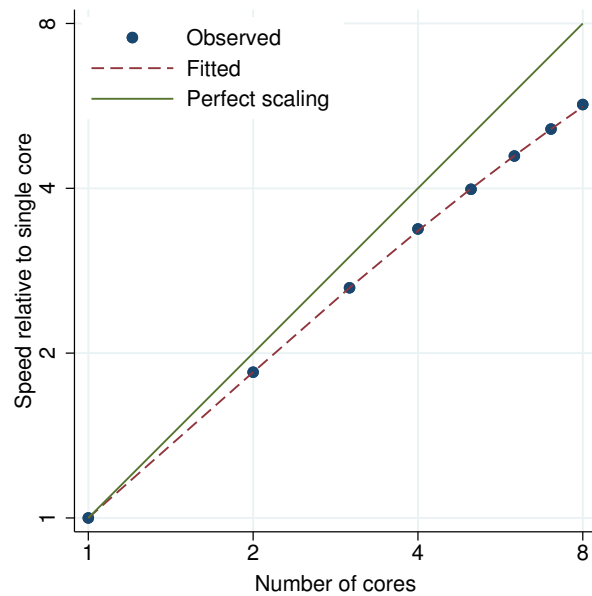


Figure 382. `reg3` performance plot.

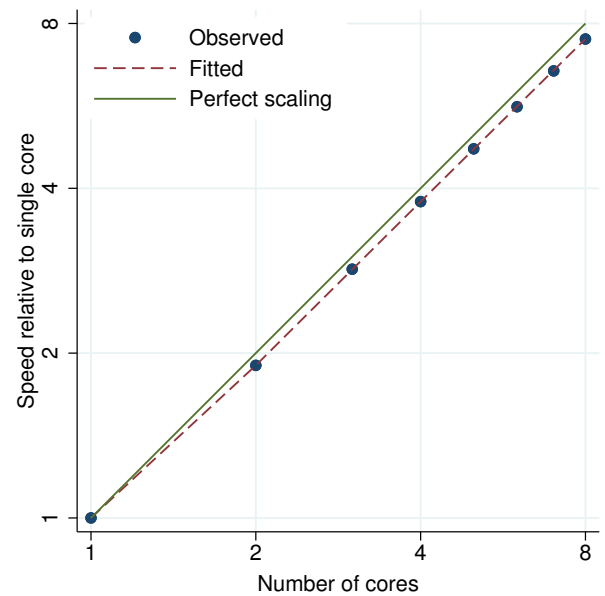


Figure 383. `regress` performance plot.

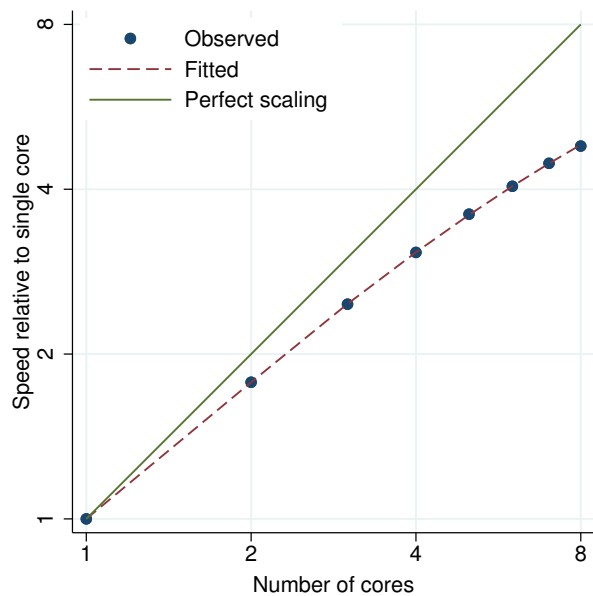


Figure 384. `regress, vce(cluster)` performance plot.

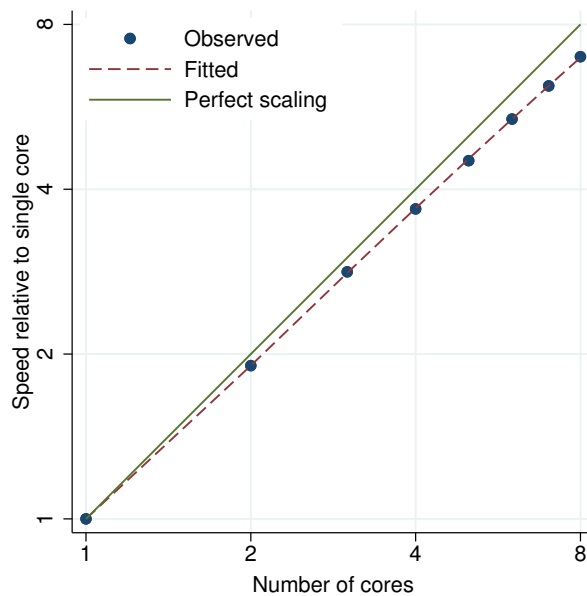


Figure 385. `regress, vce(robust)` performance plot.

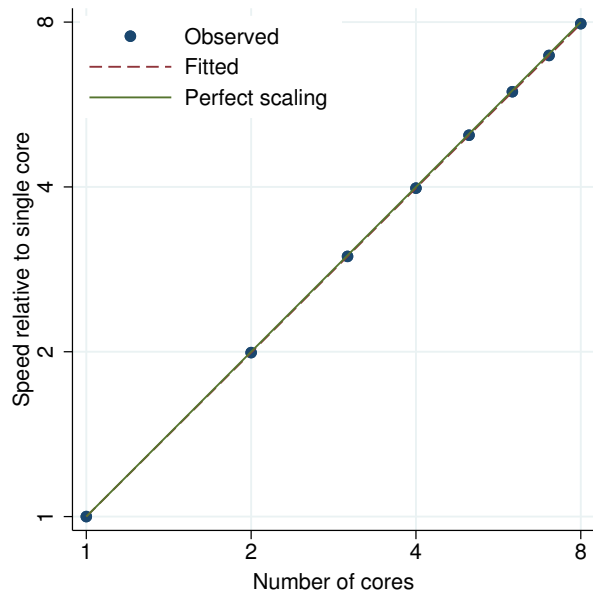


Figure 386. `replace` performance plot.

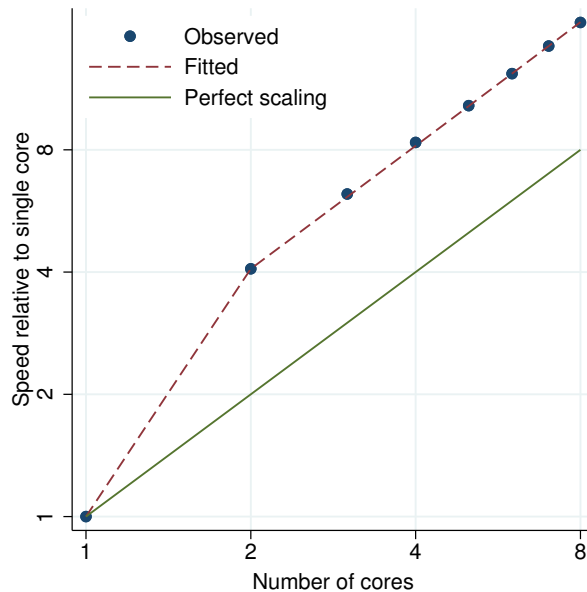


Figure 387. `replace (small expressions)` performance plot.

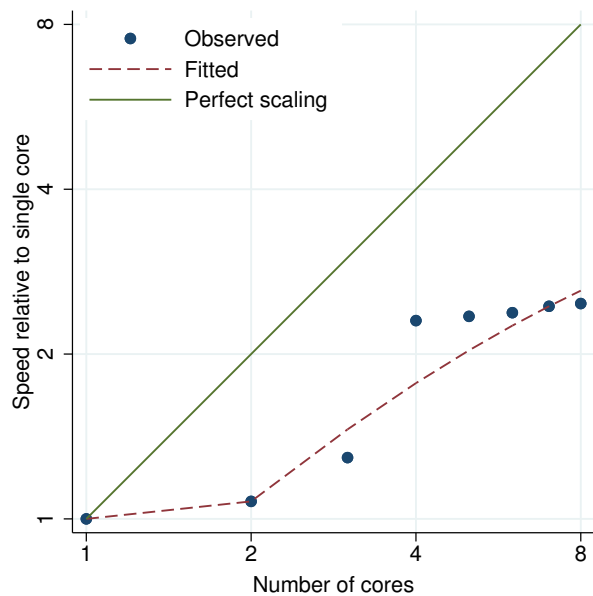


Figure 388. reshape long performance plot.

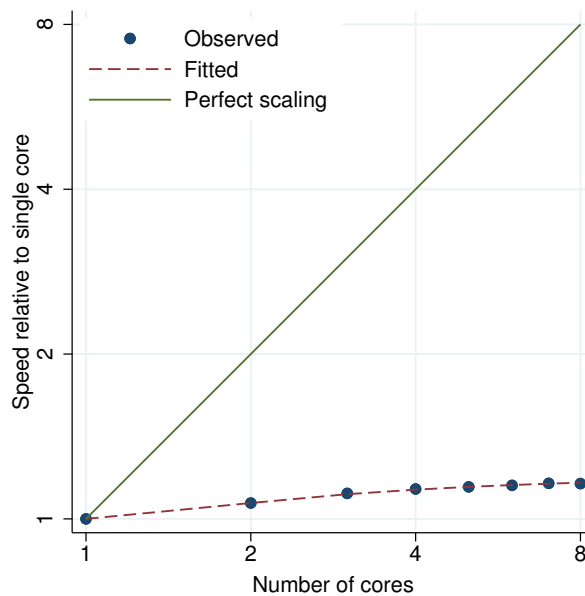


Figure 389. reshape wide performance plot.

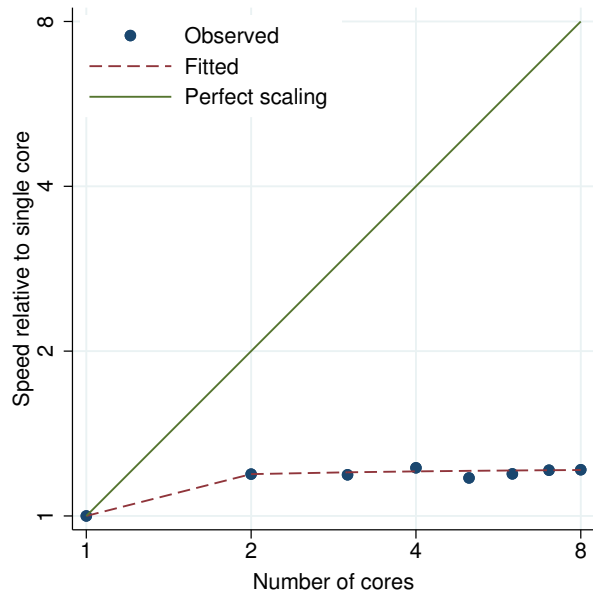


Figure 390. robvar performance plot.

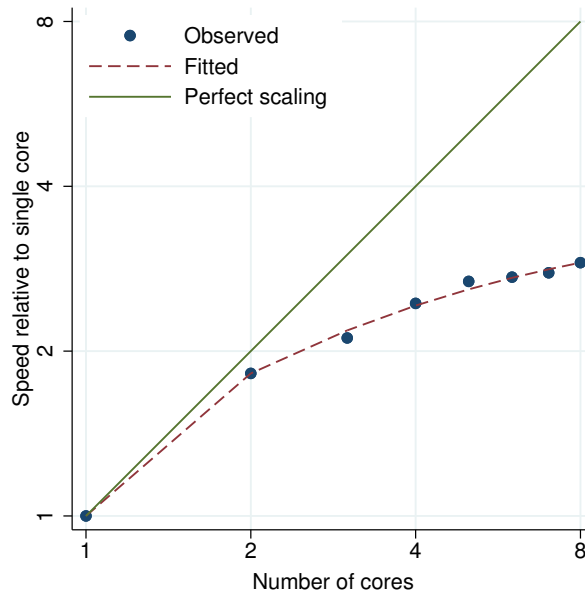


Figure 391. rocfit performance plot.

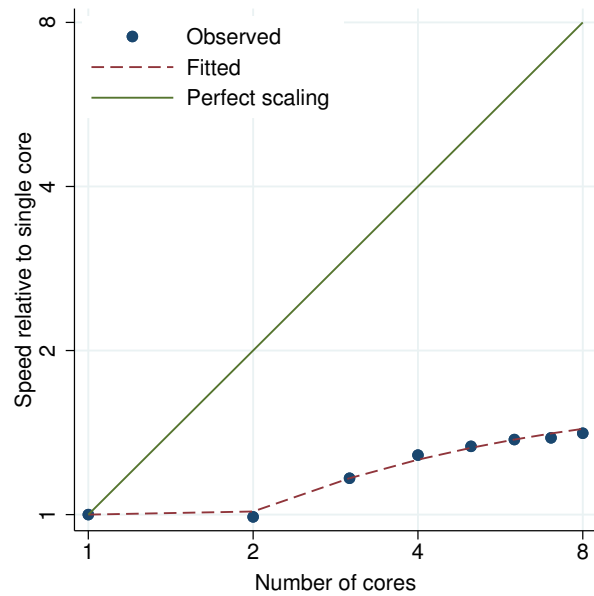


Figure 392. roctab performance plot.

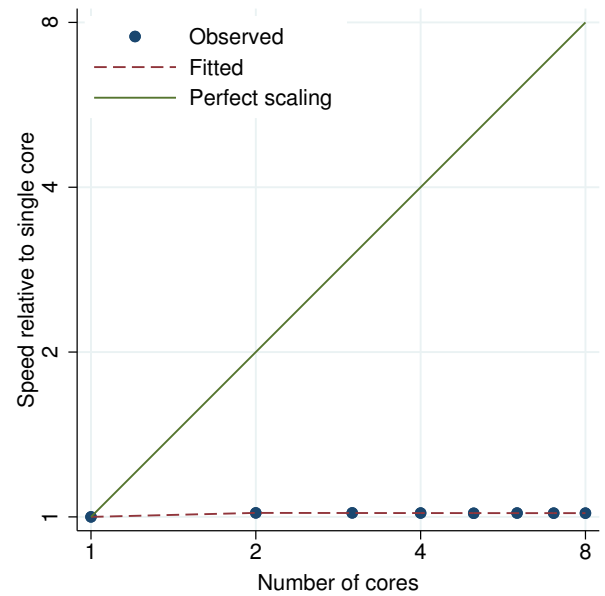


Figure 393. rotate performance plot.

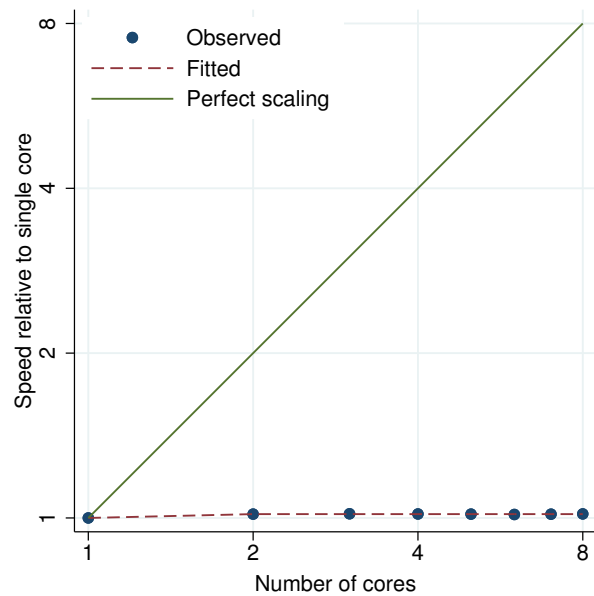


Figure 394. rotatemat performance plot.

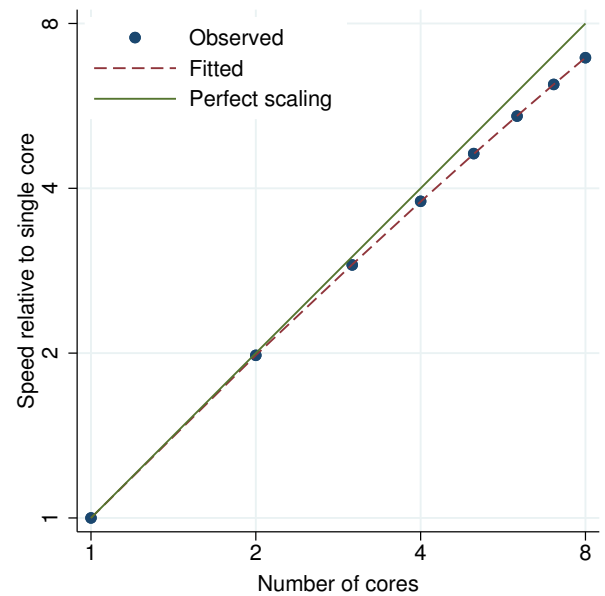


Figure 395. rreg performance plot.

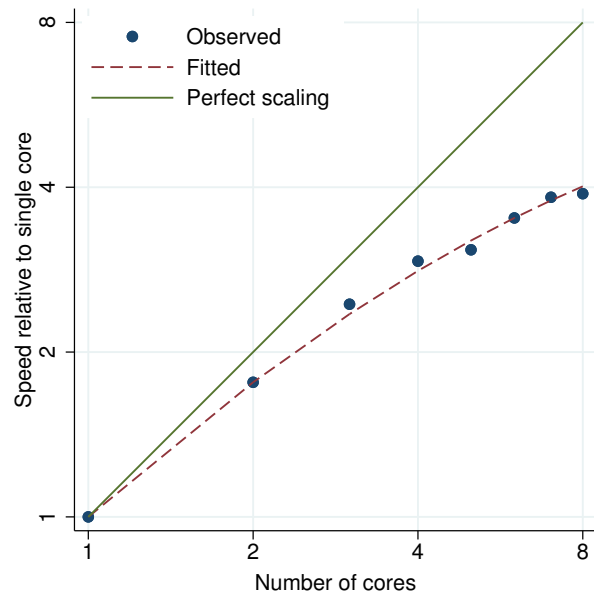


Figure 396. runtest performance plot.

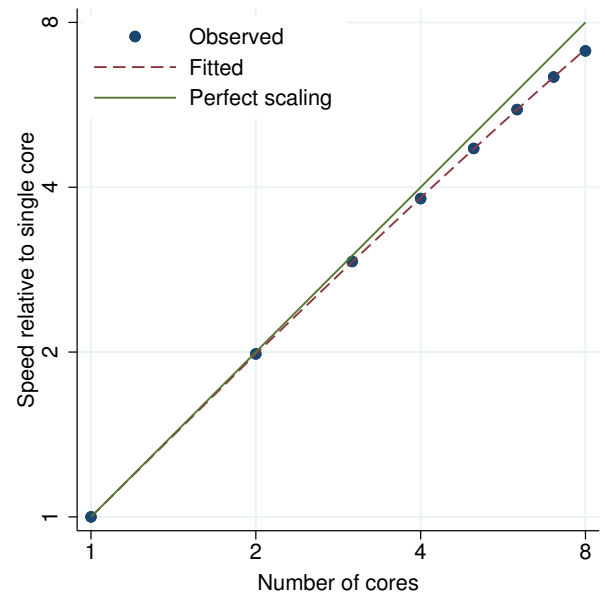


Figure 397. scobit performance plot.

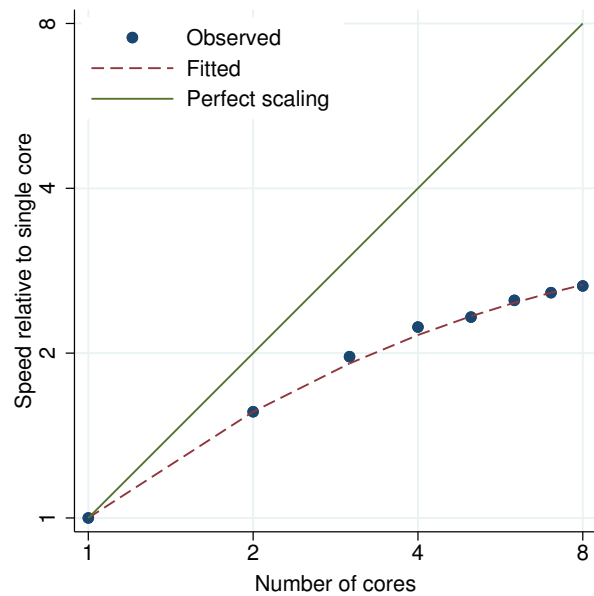


Figure 398. scoreplot performance plot.

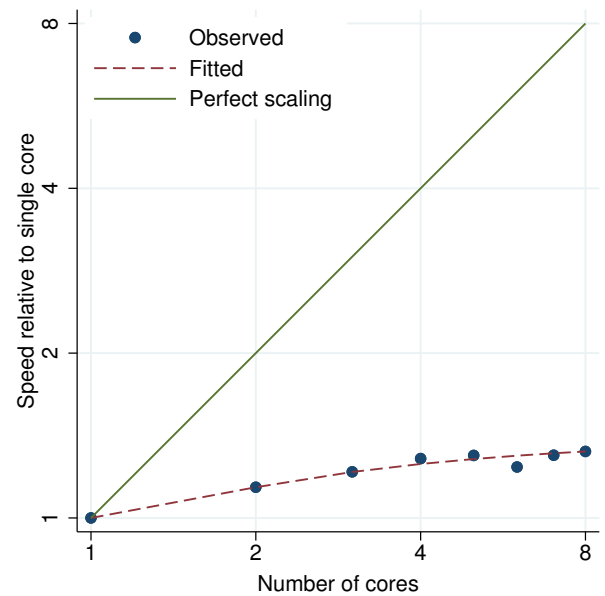
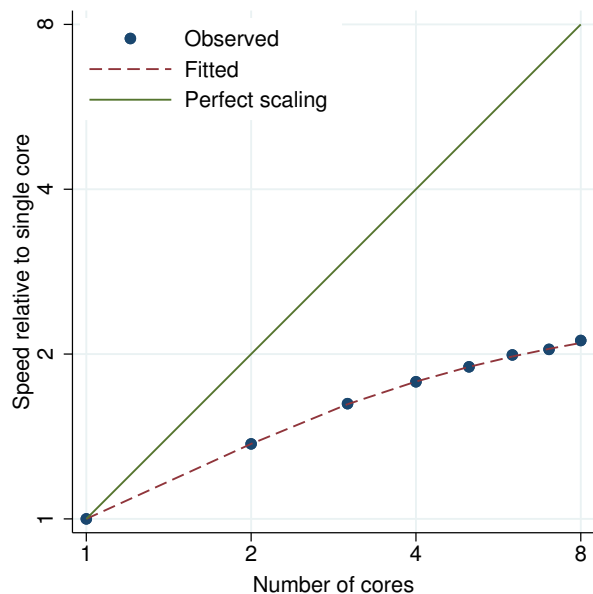
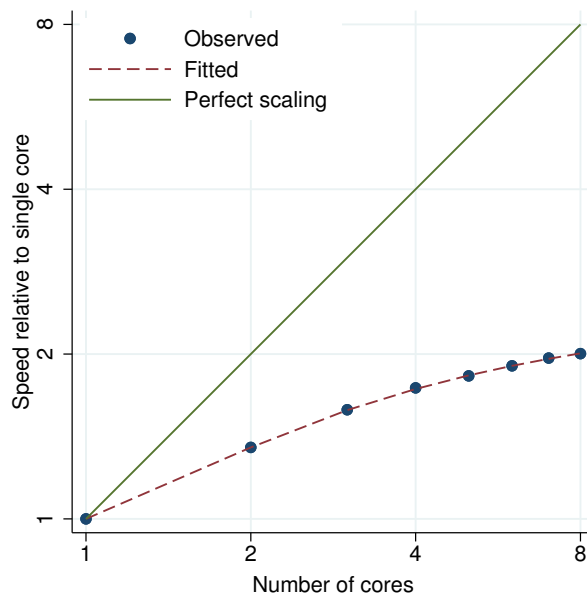
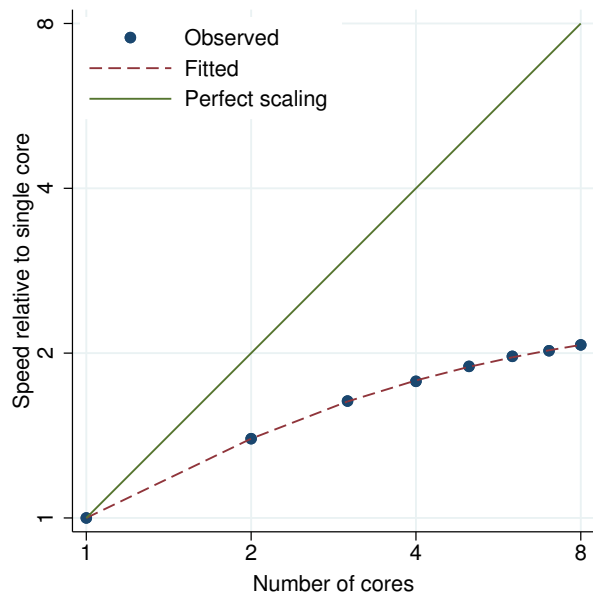
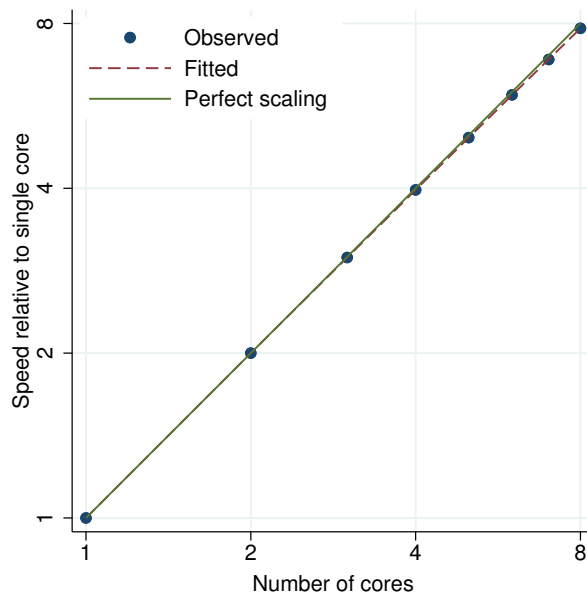


Figure 399. screeplot performance plot.

Figure 400. `sdtest1` performance plot.Figure 401. `sdtest2` performance plot.Figure 402. `sdtest, by()` performance plot.Figure 403. `sem, method(adf)` (CFA) performance plot.

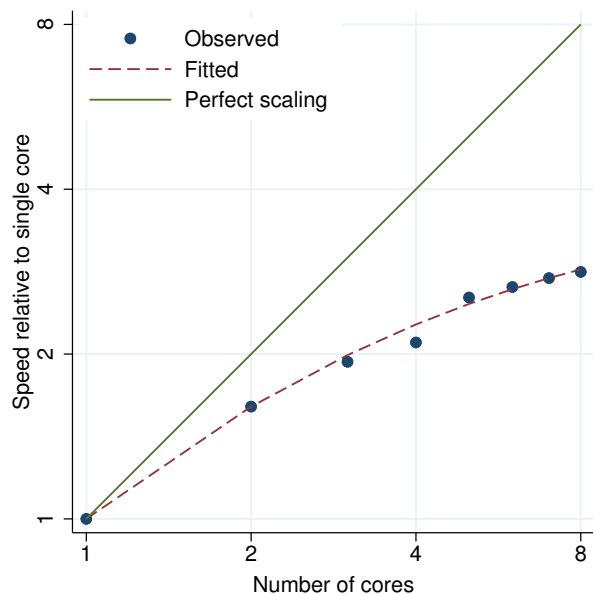


Figure 404. `sem, method(ml)` (CFA) performance plot.

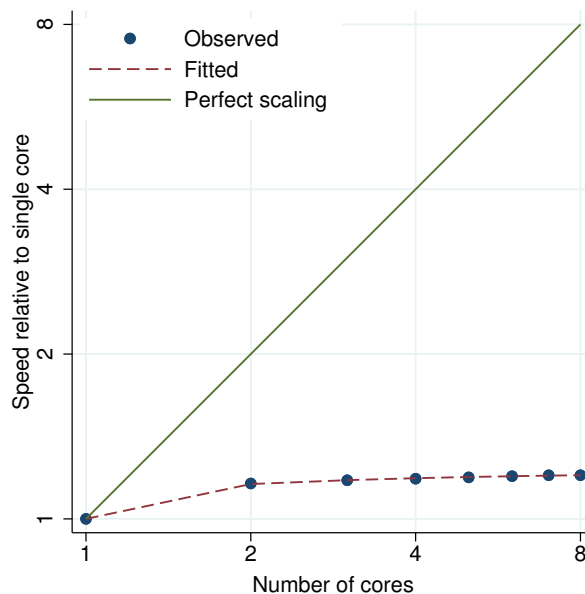


Figure 405. `sem, method(mlmv)` (CFA) performance plot.

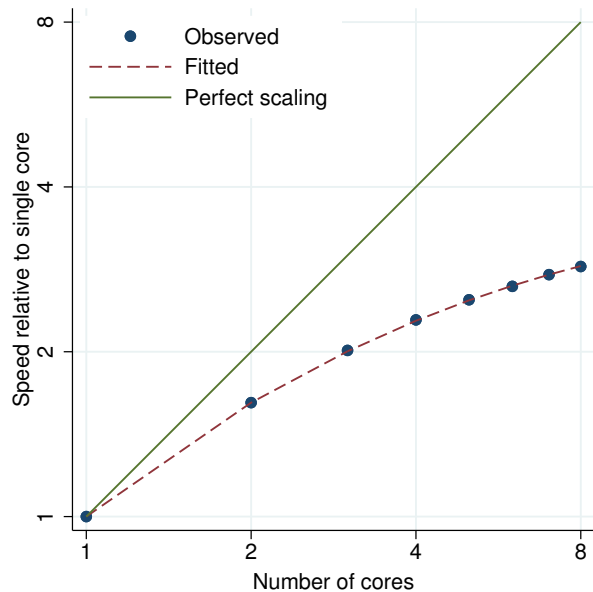


Figure 406. `sem` (SEM latent) performance plot.

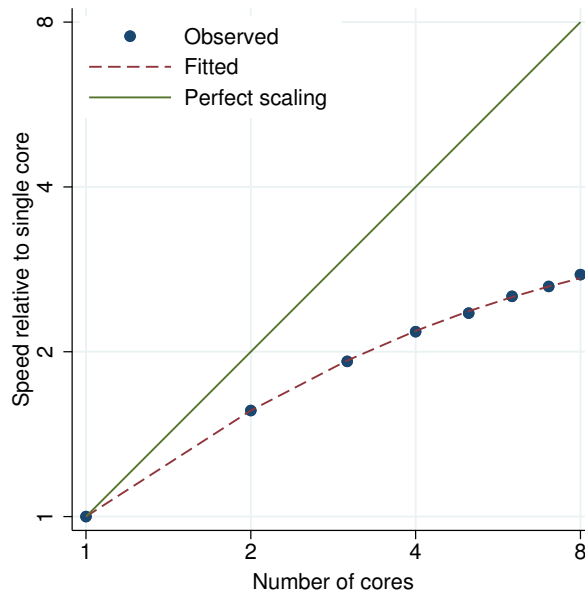
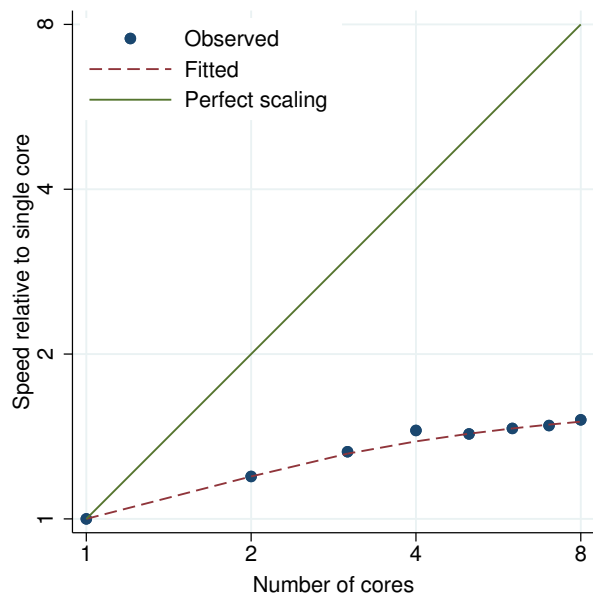
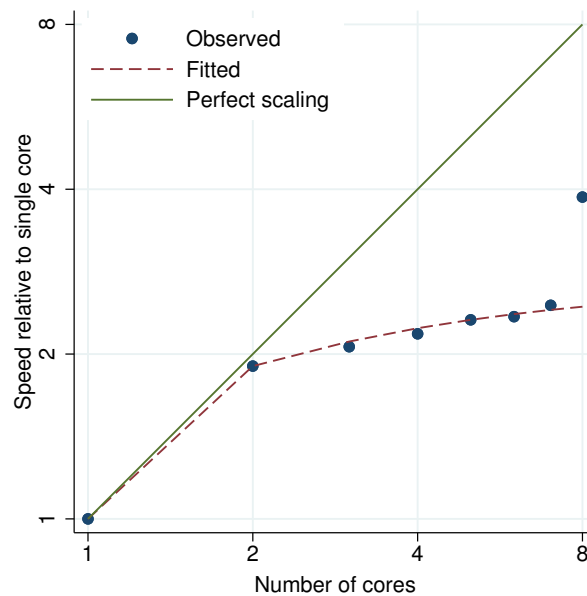
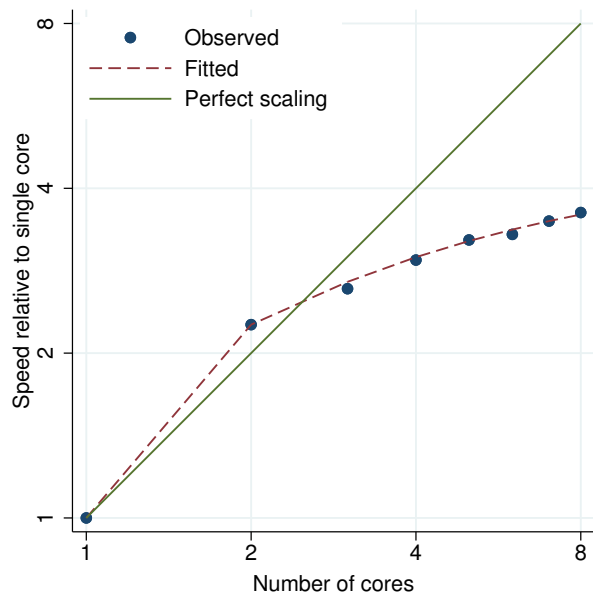
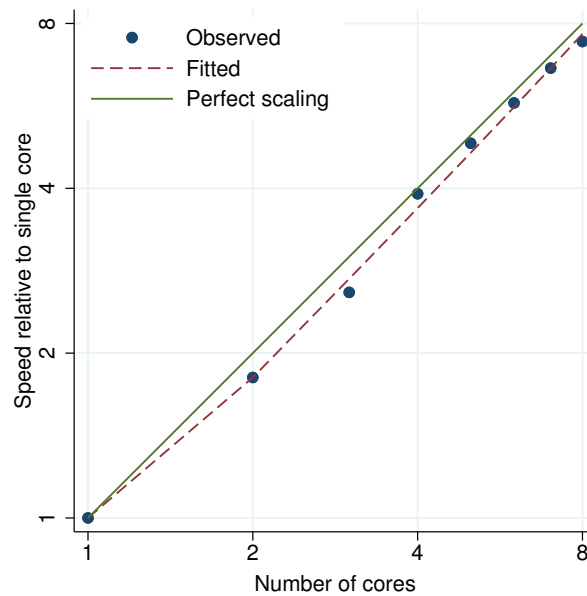
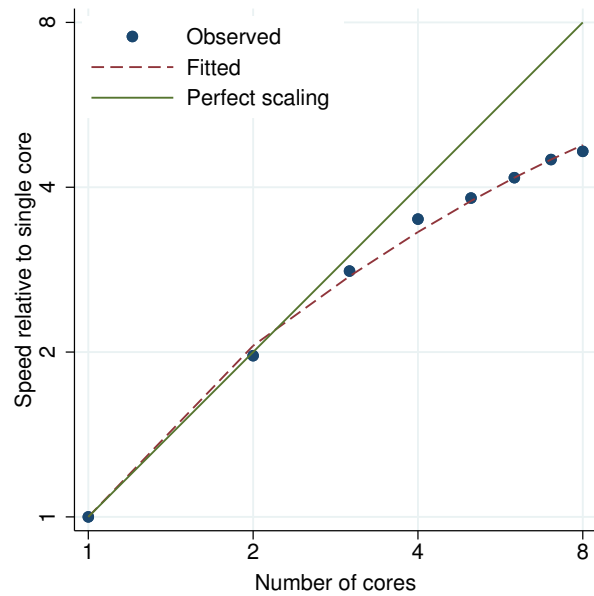
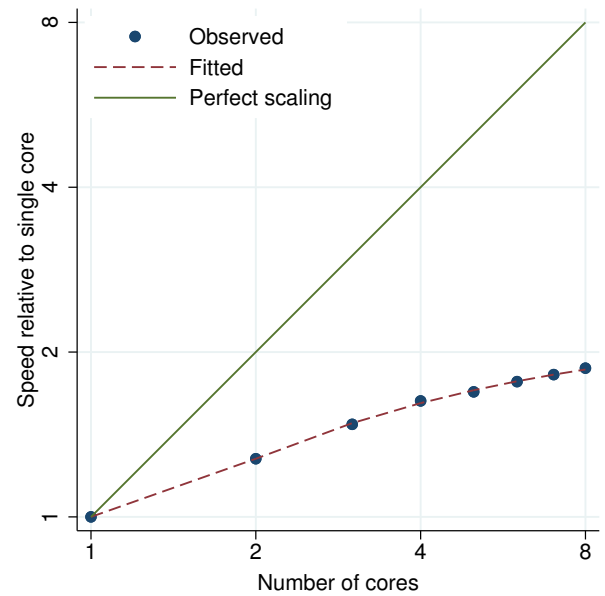
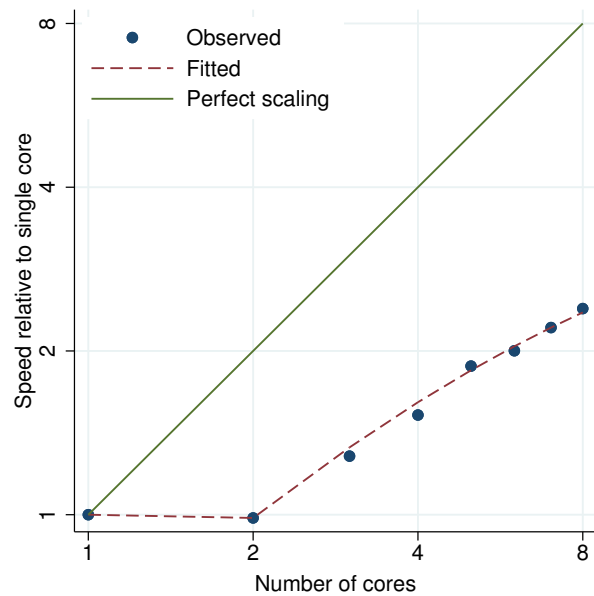
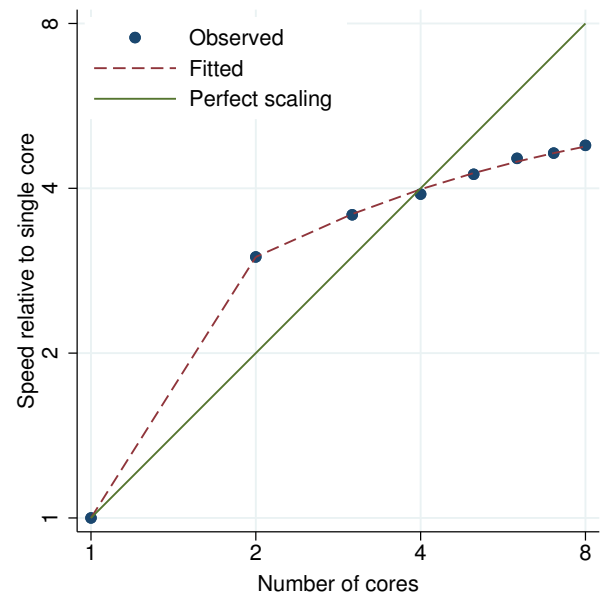
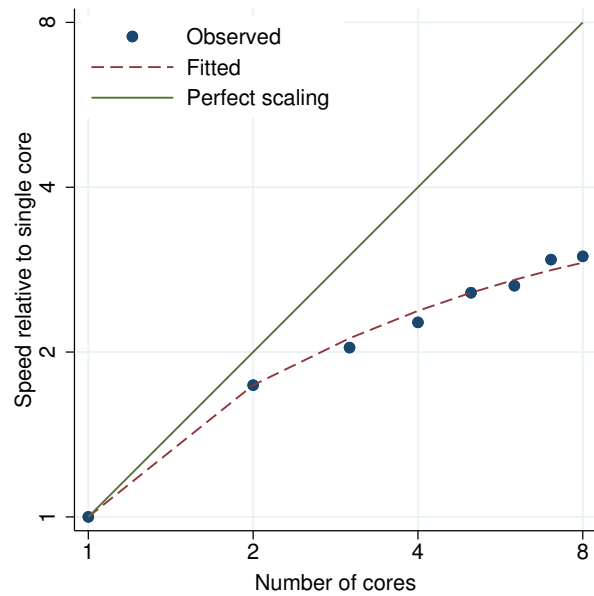
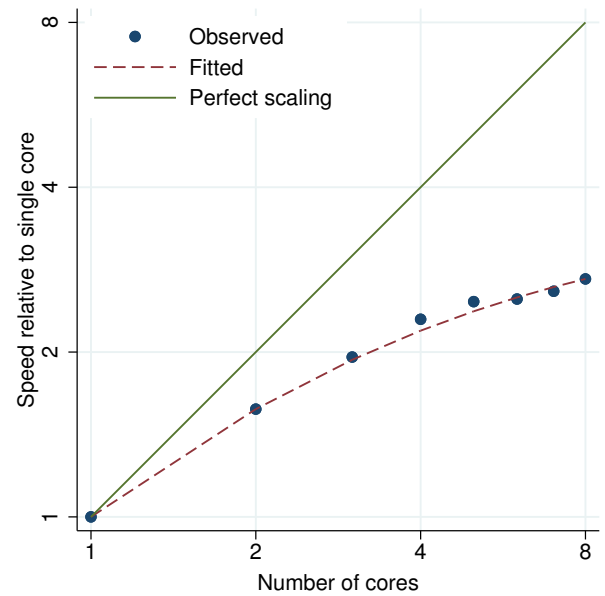
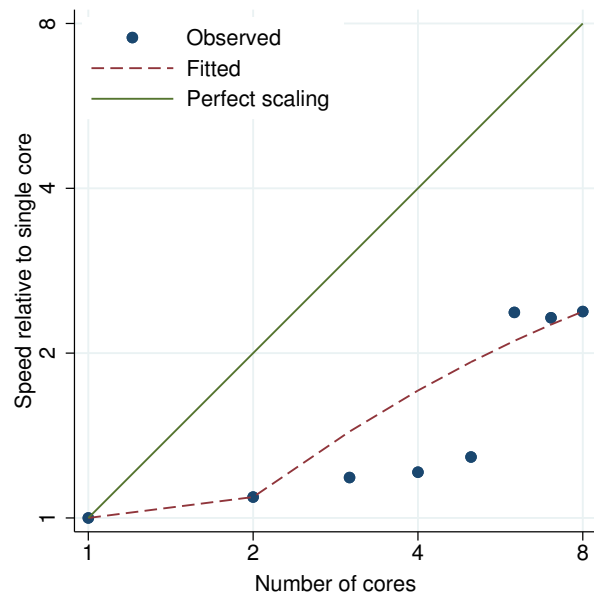
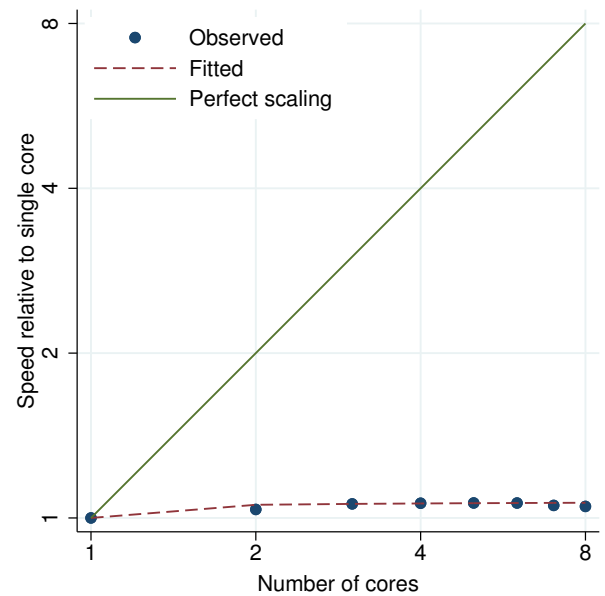


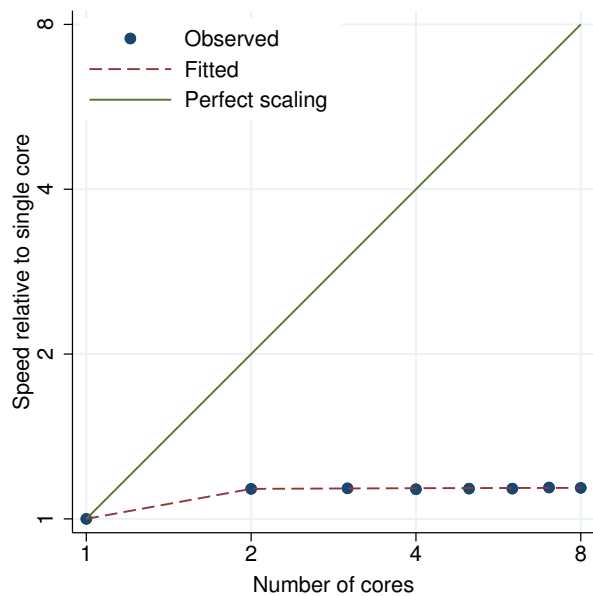
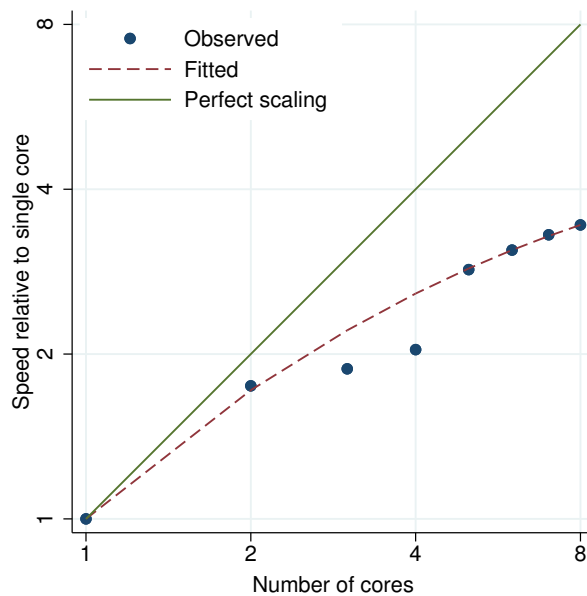
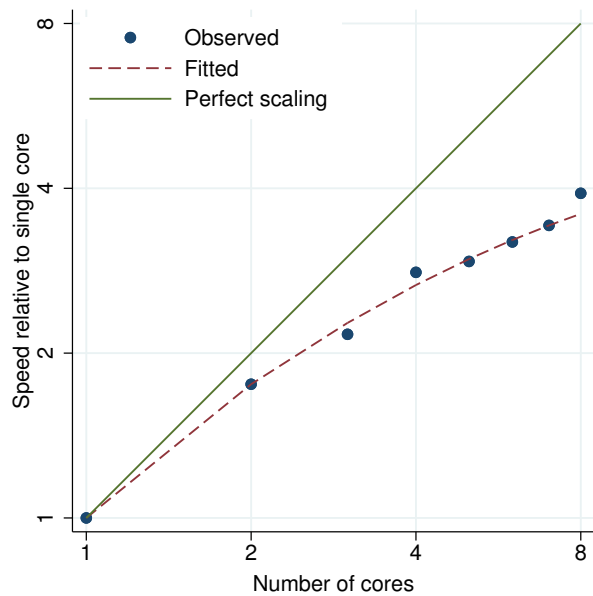
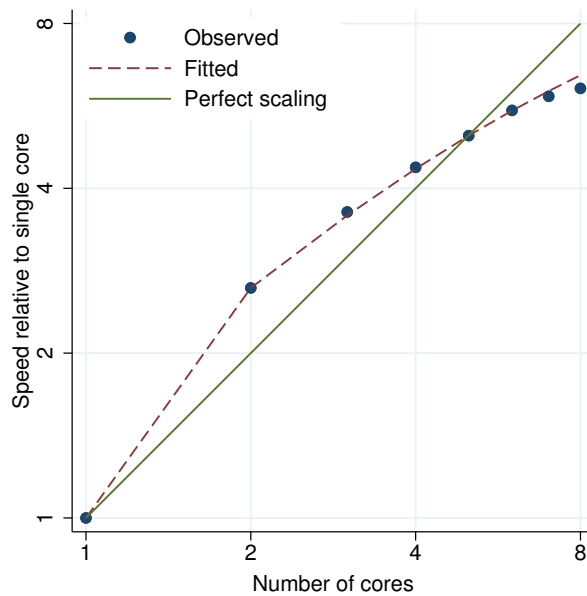
Figure 407. `sem` (SEM observed) performance plot.

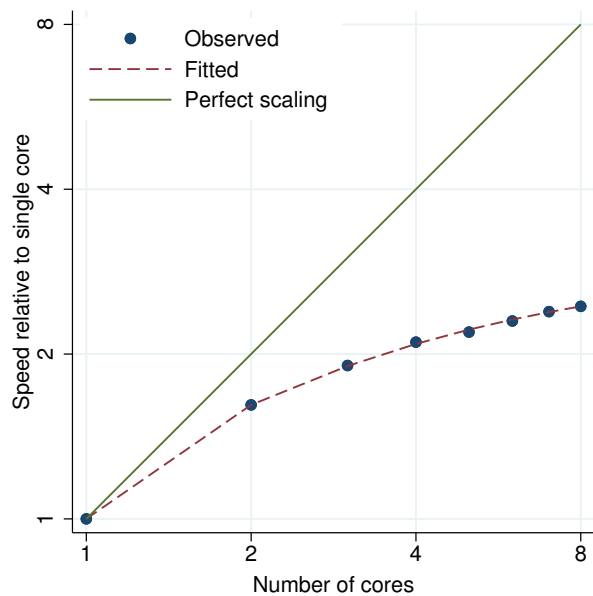
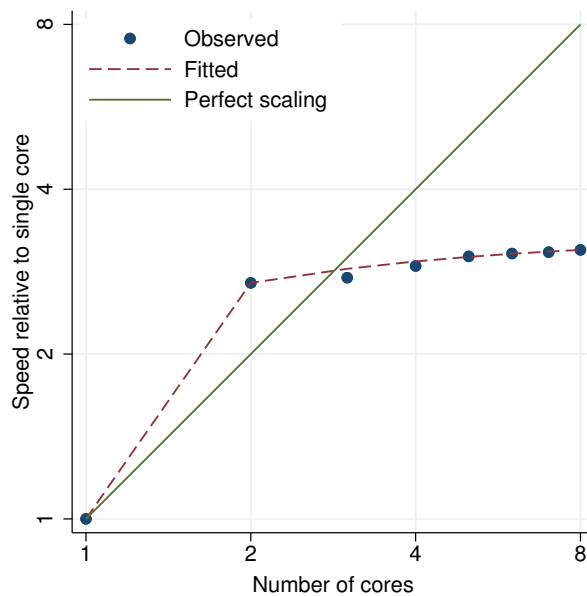
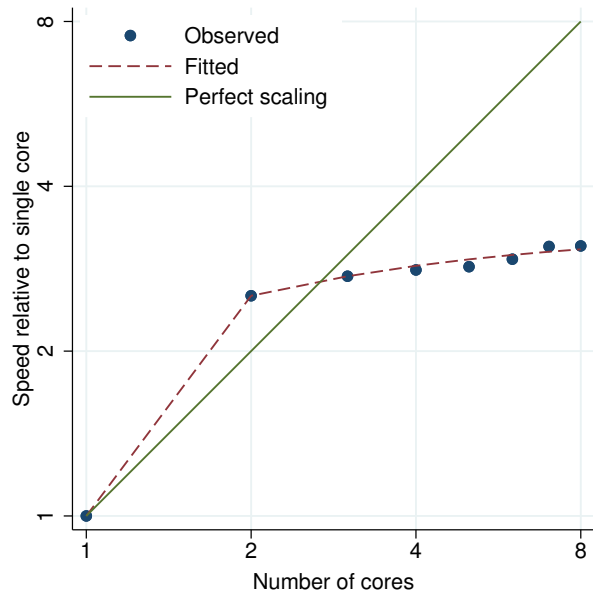
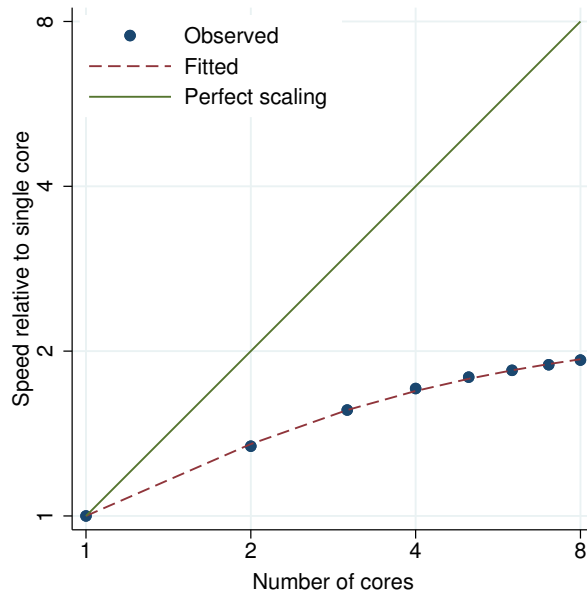
Figure 408. `separate` performance plot.Figure 409. `sfrancia` performance plot.Figure 410. `signrank` performance plot.Figure 411. `signtest` performance plot.

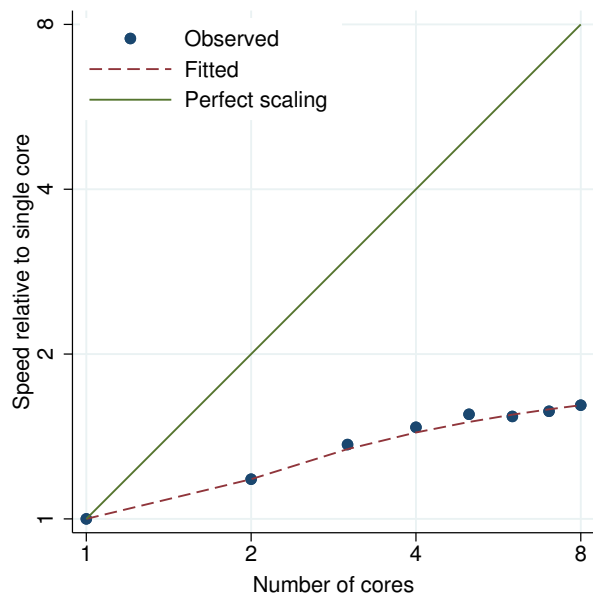
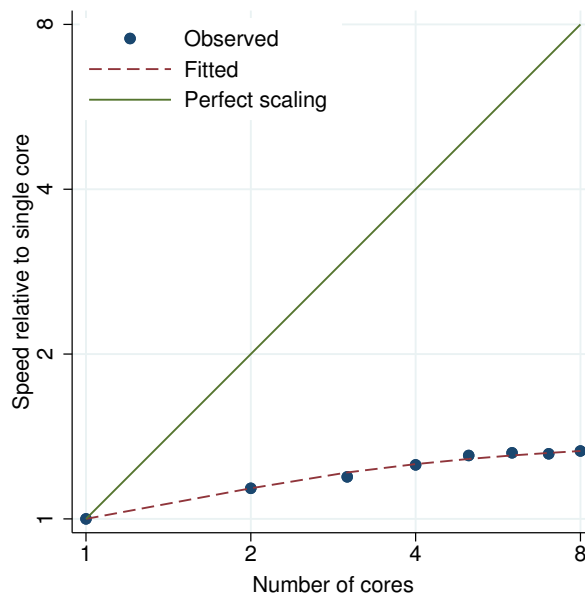
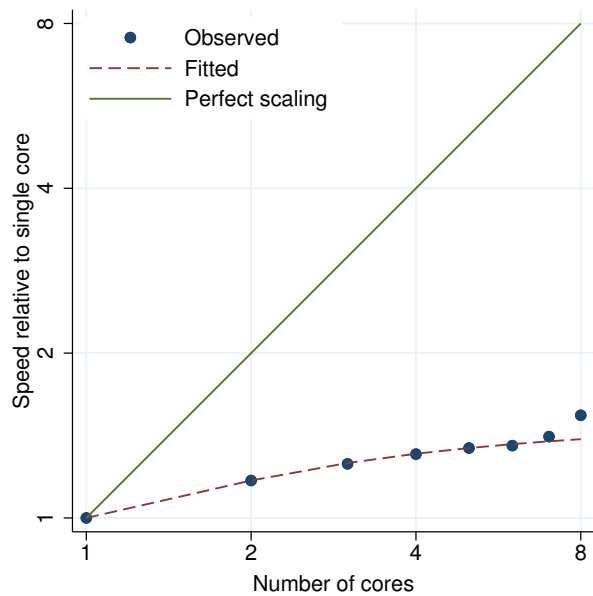
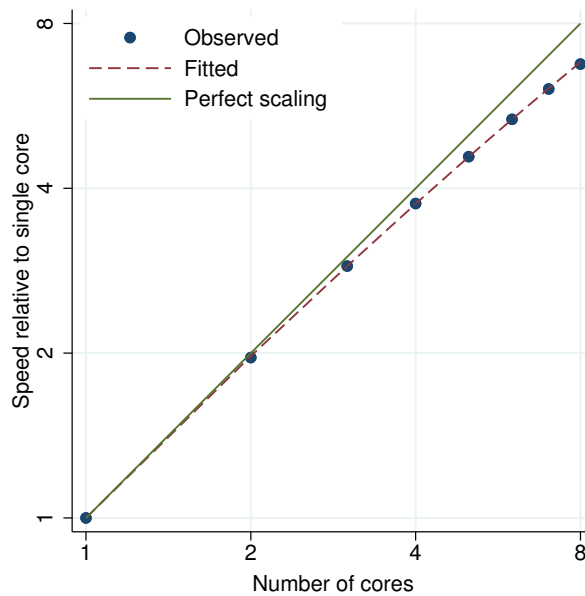


Figure 412. `sktest` performance plot.Figure 413. `slogit` performance plot.Figure 414. `sort` performance plot.Figure 415. `spearman` performance plot.

Figure 416. `sspace` performance plot.Figure 417. `stack` performance plot.Figure 418. `stci` performance plot.Figure 419. `stcox` performance plot.

Figure 420. `stcrreg` performance plot.Figure 421. `stgen` performance plot.Figure 422. `stintreg`, `d(exponential)` performance plot.Figure 423. `stintreg`, `d(weibull)` performance plot.

Figure 424. `stir` performance plot.Figure 425. `stmc` performance plot.Figure 426. `by: stmc` performance plot.Figure 427. `stmh` performance plot.

Figure 428. `by: stmh` performance plot.Figure 429. `stptime` performance plot.Figure 430. `strate` performance plot.Figure 431. `streg, distribution(exponential)` performance plot.

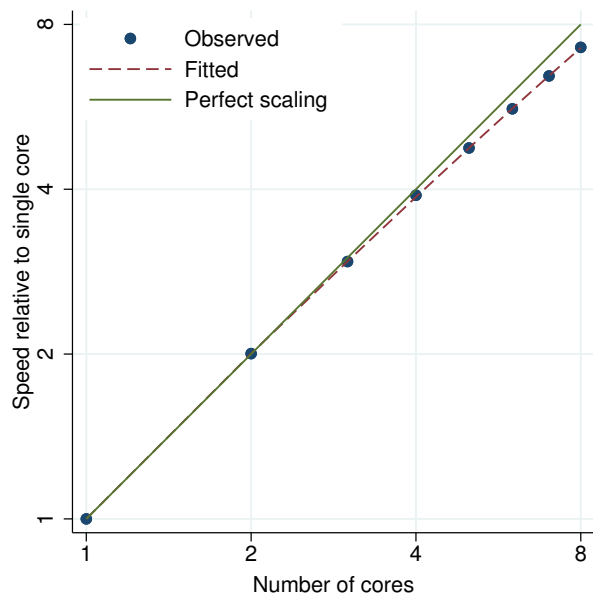


Figure 432. `streg, dist(exp) vce(cluster)` performance plot.

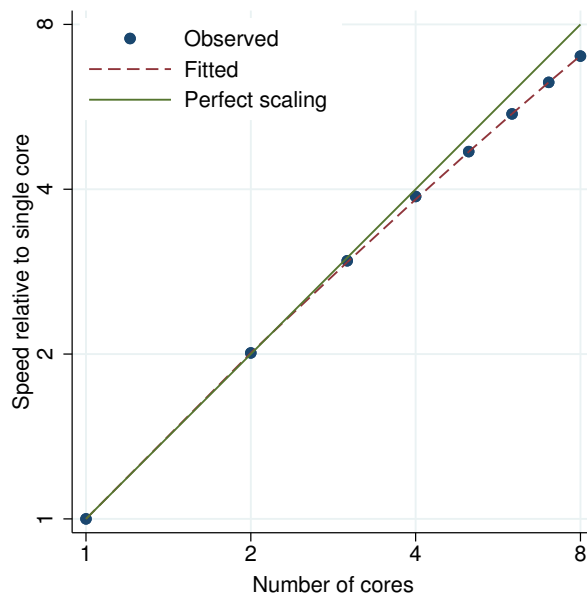


Figure 433. `streg, dist(exp) frailty()` performance plot.

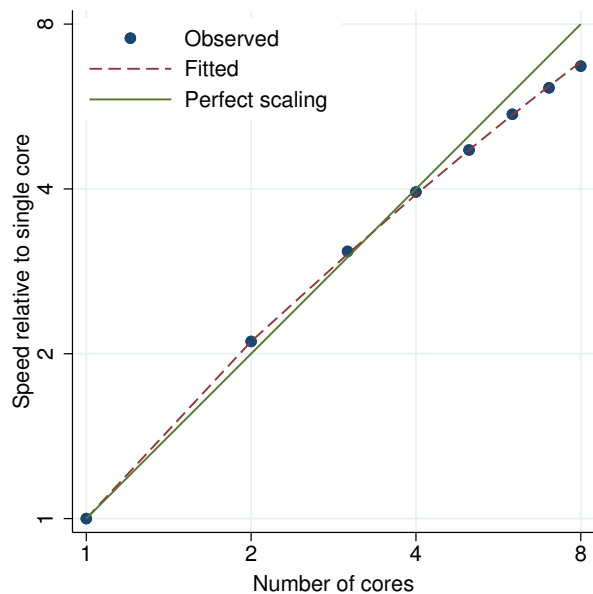


Figure 434. `streg, dist(exp) frailty() shared()` performance plot.

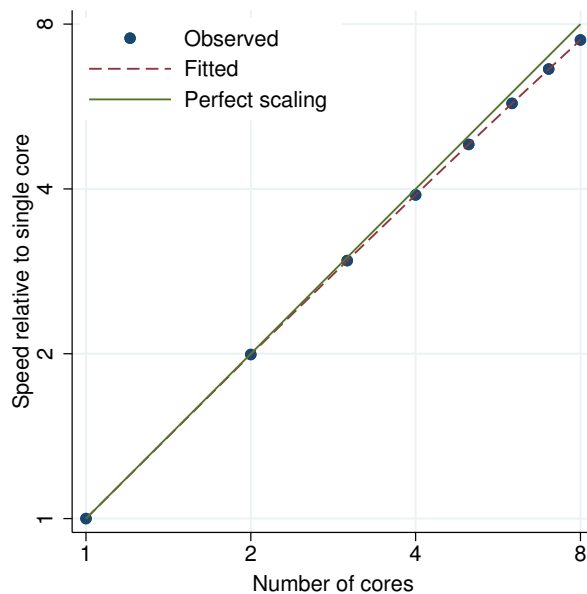


Figure 435. `streg, dist(exp) vce(robust)` performance plot.

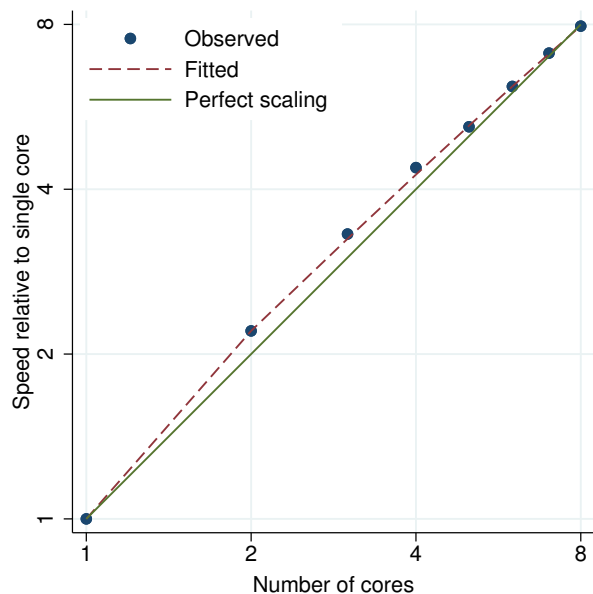


Figure 436. `streg, distribution(ggamma)` performance plot.

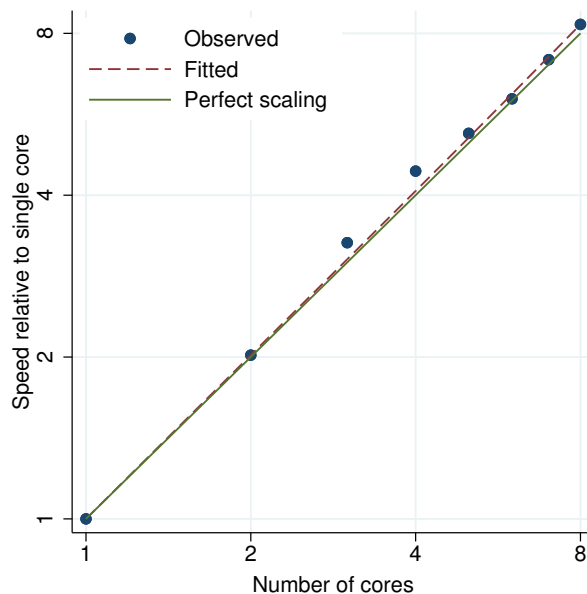


Figure 437. `streg, dist(ggamma) vce(cluster)` performance plot.

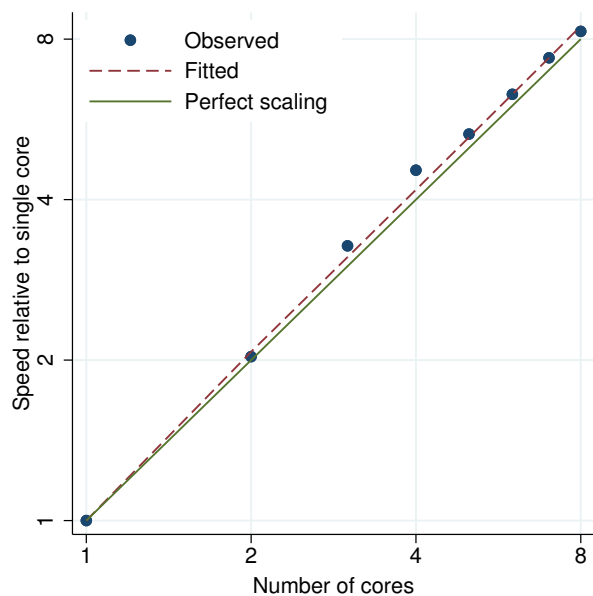


Figure 438. `streg, dist(ggamma) vce(robust)` performance plot.

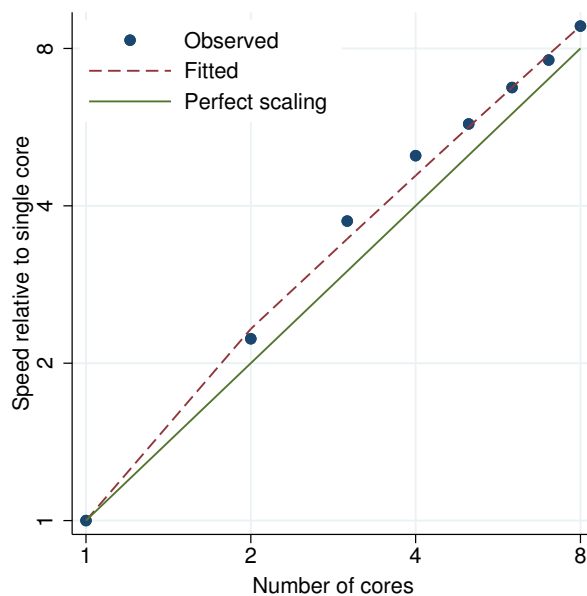


Figure 439. `streg, distribution(gompertz)` performance plot.

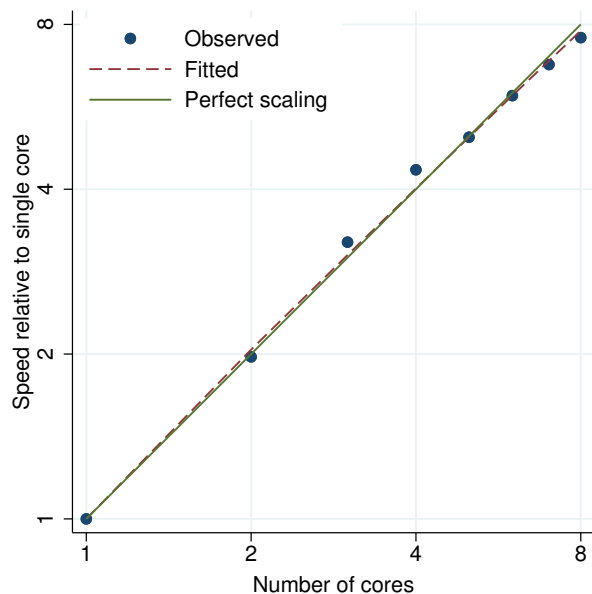


Figure 440. `streg, dist(gompertz)`  
`vce(cluster)` performance plot.

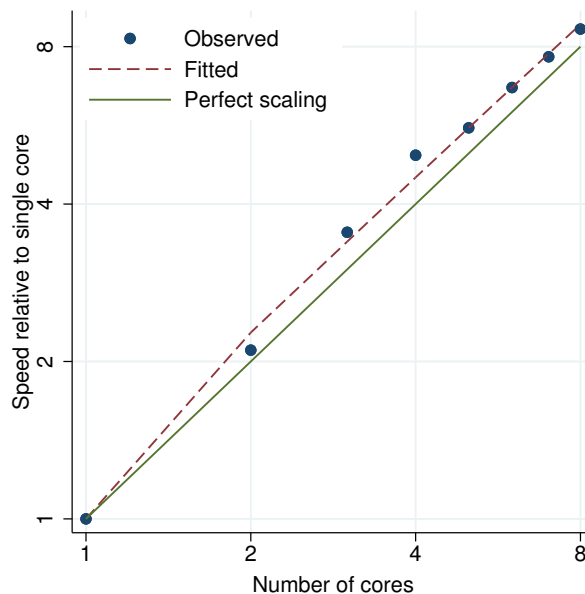


Figure 441. `streg, dist(gompertz)`  
`frailty()` performance plot.

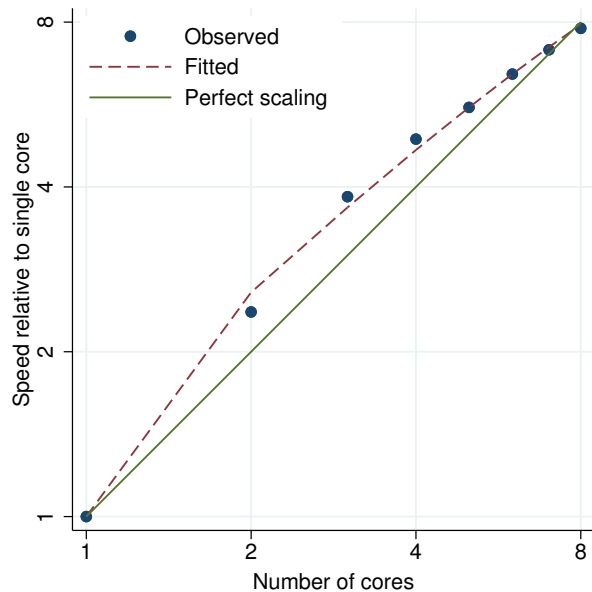


Figure 442. `streg, dist(gomp)` `frailty()`  
`shared()` performance plot.

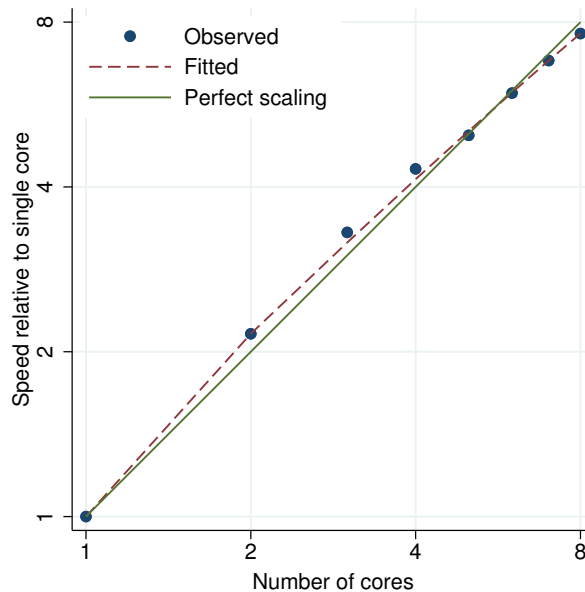


Figure 443. `streg, dist(gompertz)`  
`vce(robust)` performance plot.



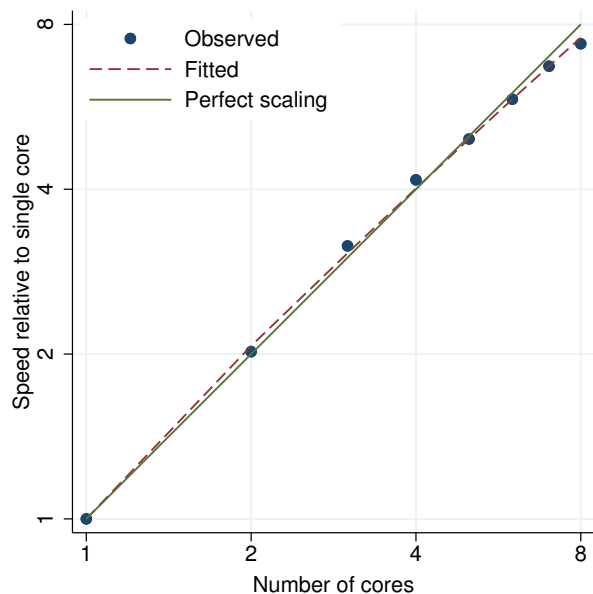


Figure 444. `streg, distribution(llogistic)` performance plot.

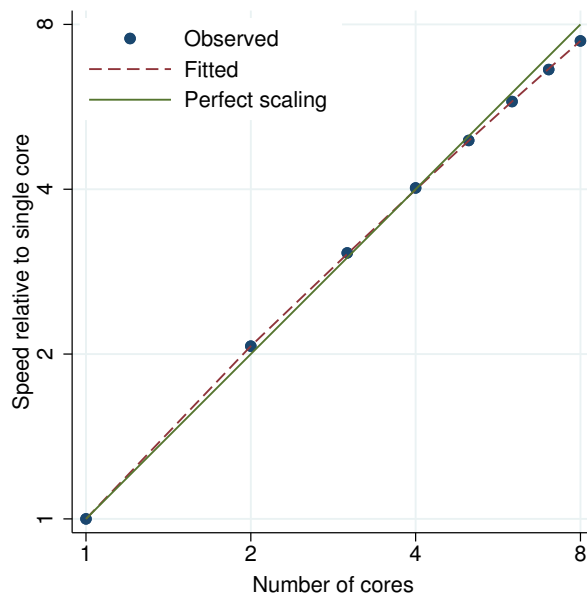


Figure 445. `streg, dist(llogistic) vce(cluster)` performance plot.

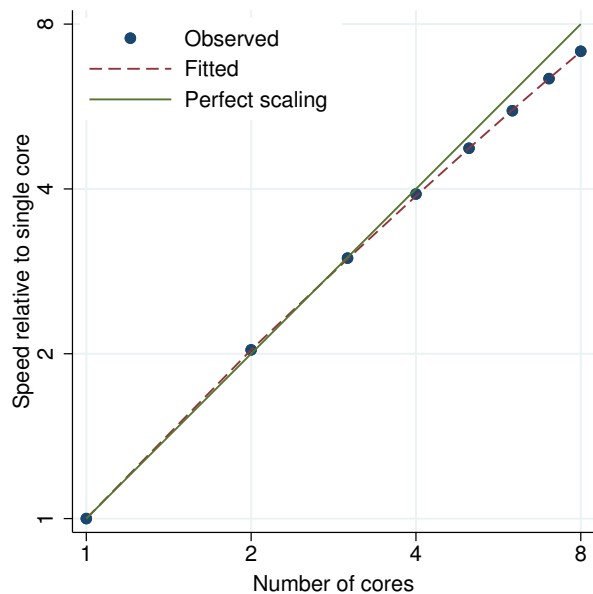


Figure 446. `streg, dist(llogistic) frailty()` performance plot.

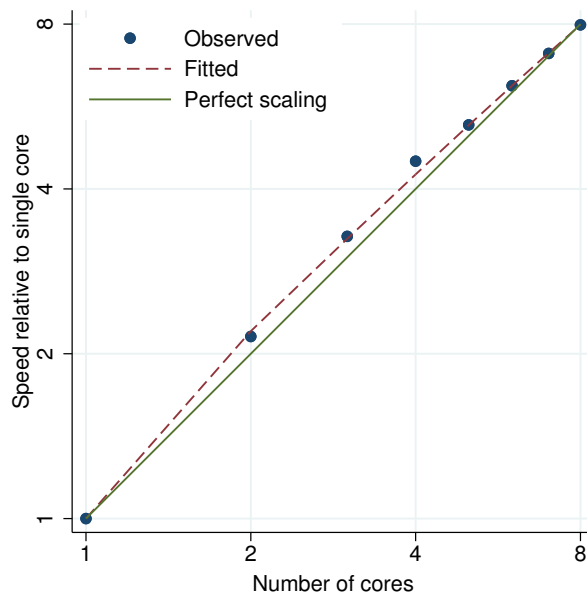


Figure 447. `streg, dist(llog) frailty() shared()` performance plot.

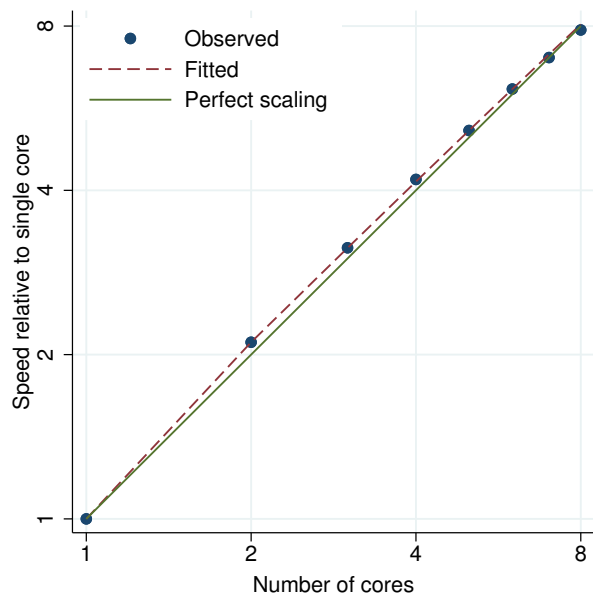


Figure 448. `streg, dist(llogistic)`  
`vce(robust)` performance plot.

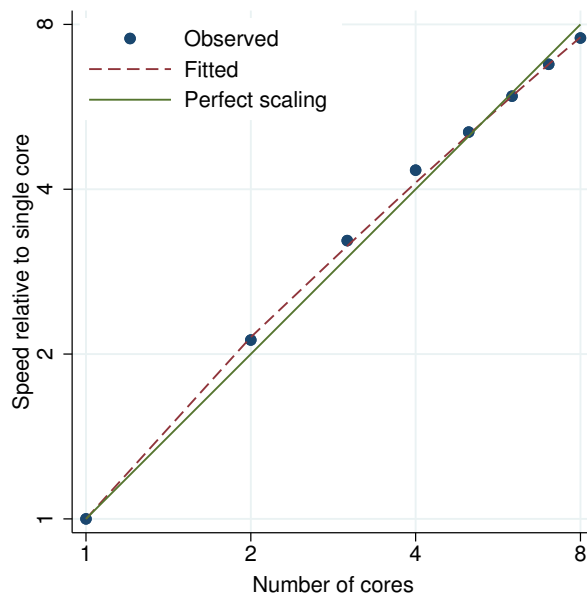


Figure 449. `streg, distribution(lnormal)`  
performance plot.

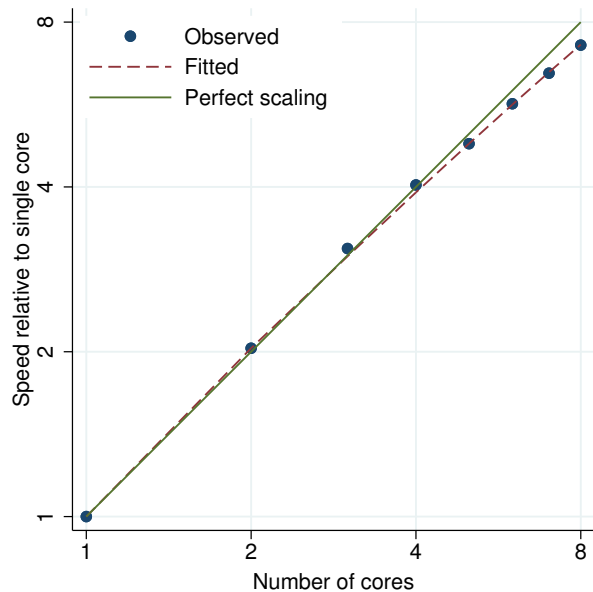


Figure 450. `streg, dist(lnormal)`  
`vce(cluster)` performance plot.

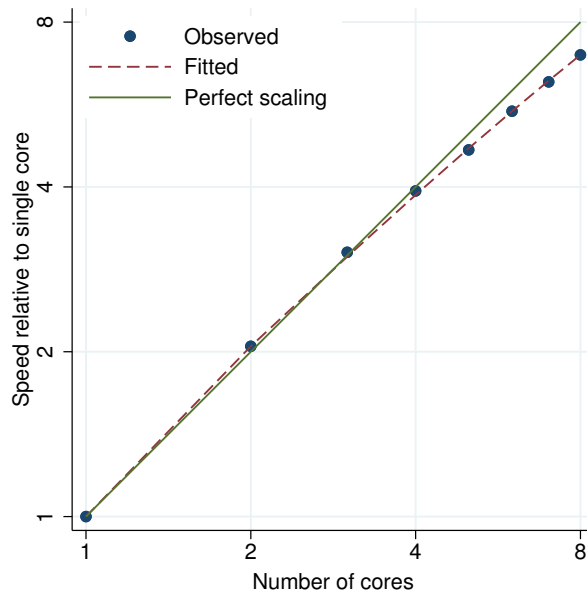


Figure 451. `streg, dist(lnormal)`  
`frailty()` performance plot.

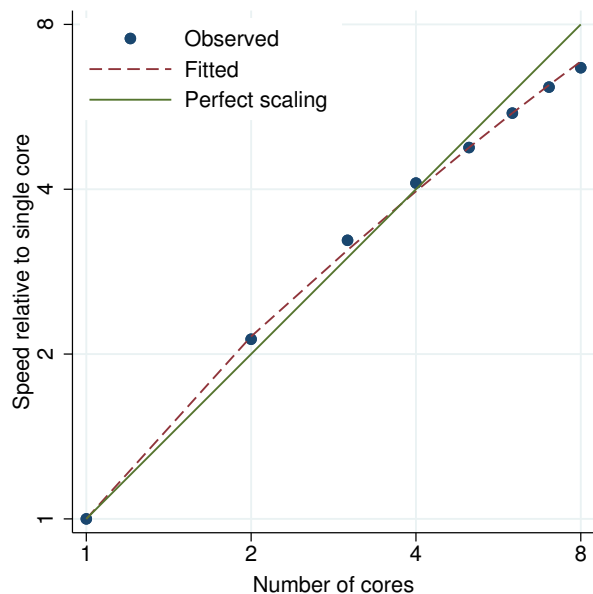


Figure 452. `streg, dist(lnorm) frailty() shared()` performance plot.

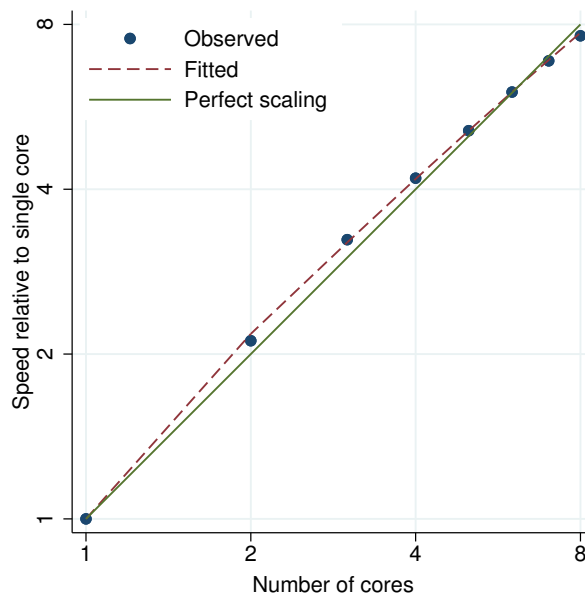


Figure 453. `streg, dist(lnormal) vce(robust)` performance plot.

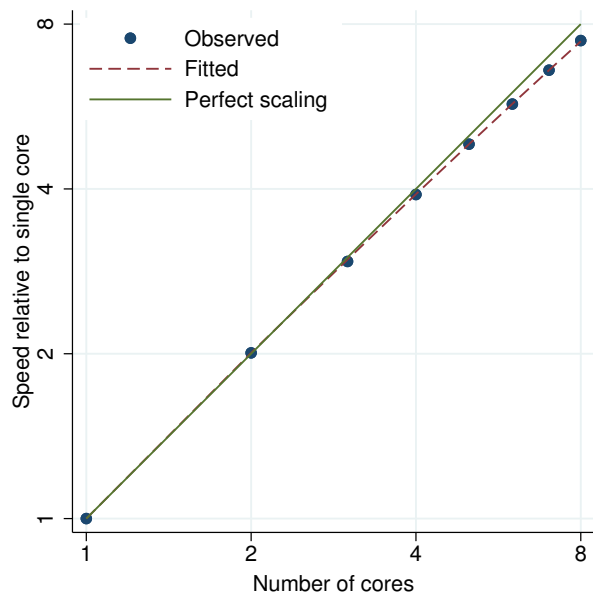


Figure 454. `streg, distribution(weibull)` performance plot.

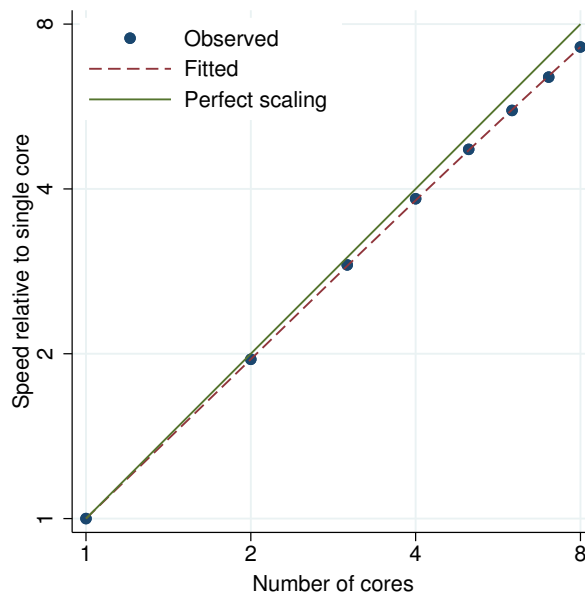


Figure 455. `streg, dist(weibull) vce(cluster)` performance plot.

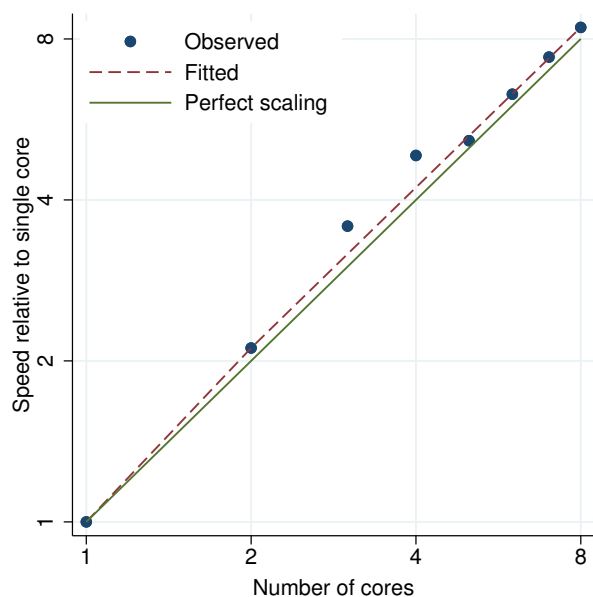


Figure 456. `streg, dist(weibull)`  
`frailty()` performance plot.

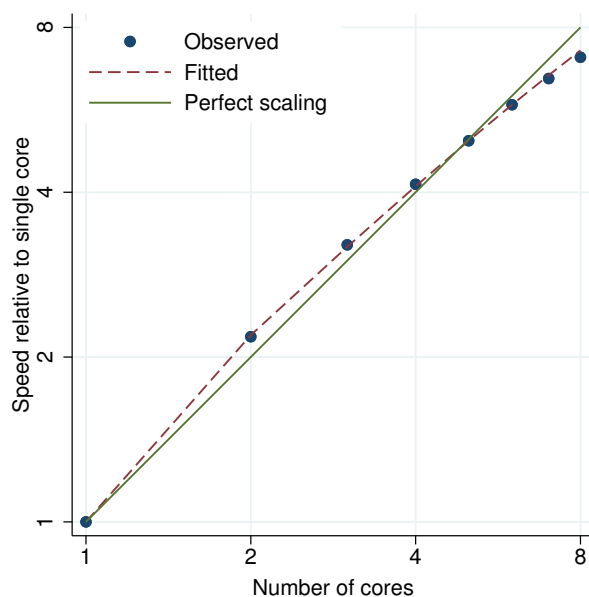


Figure 457. `streg, dist(weib) frailty()`  
`shared()` performance plot.

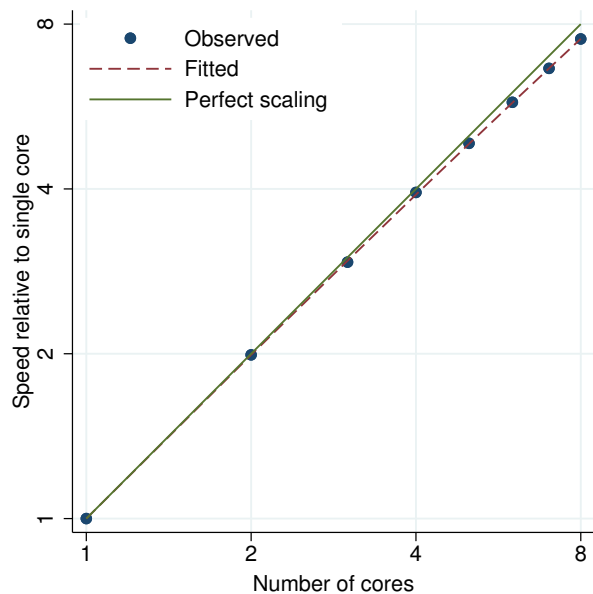


Figure 458. `streg, dist(weibull)`  
`vce(robust)` performance plot.

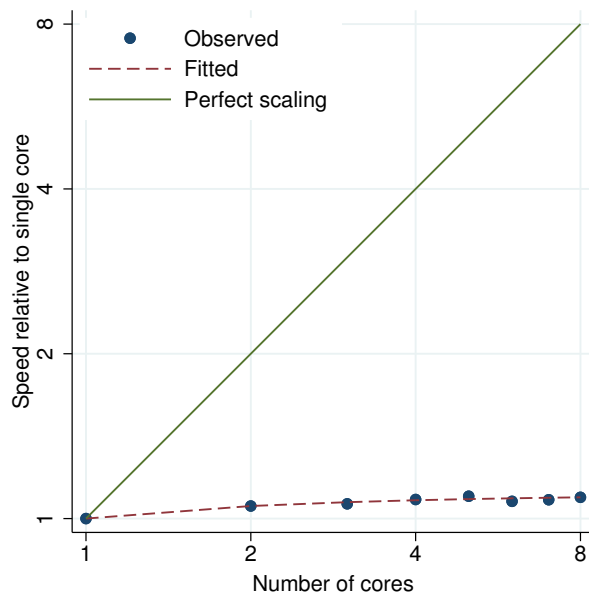
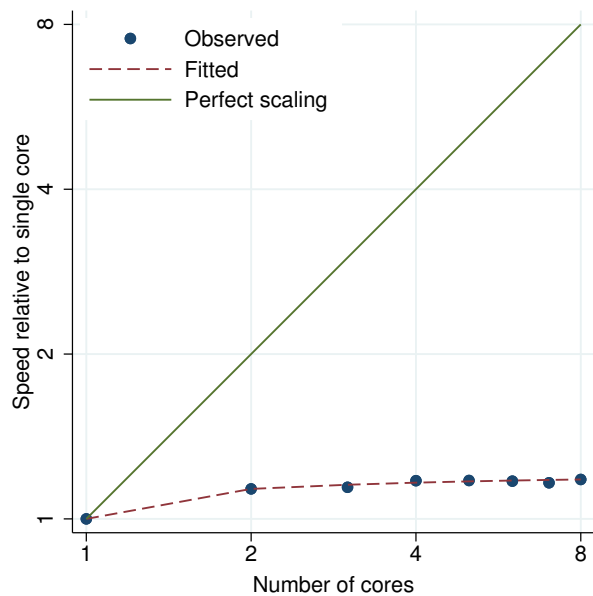
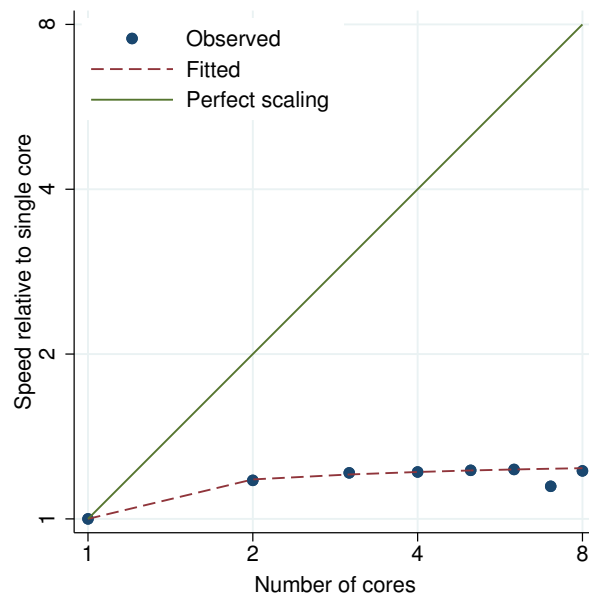
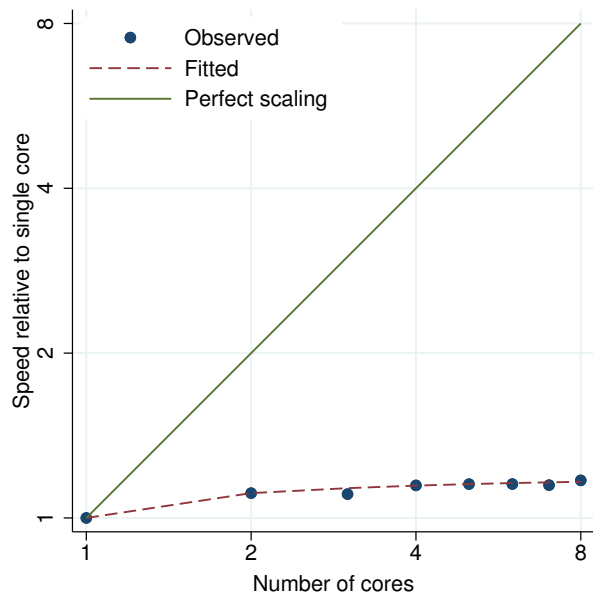
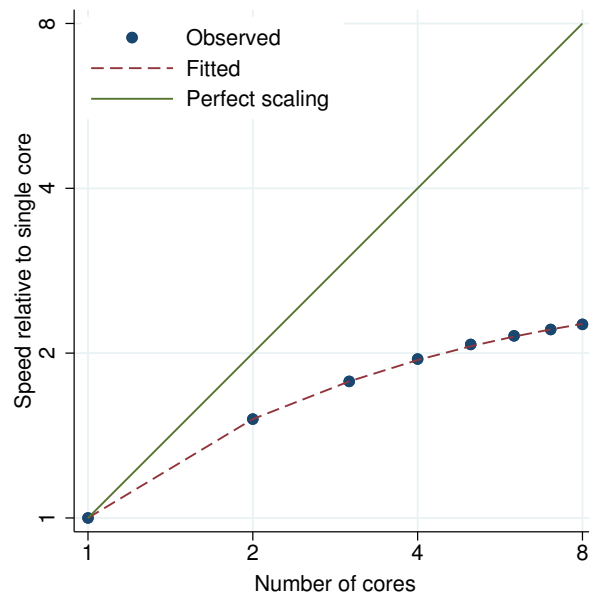
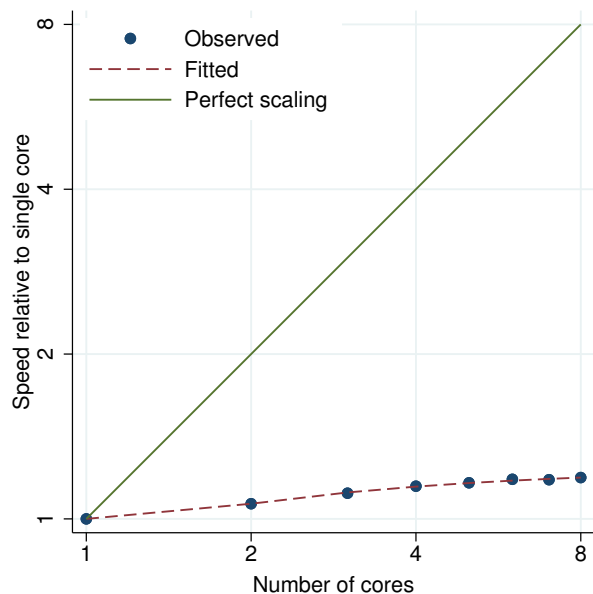
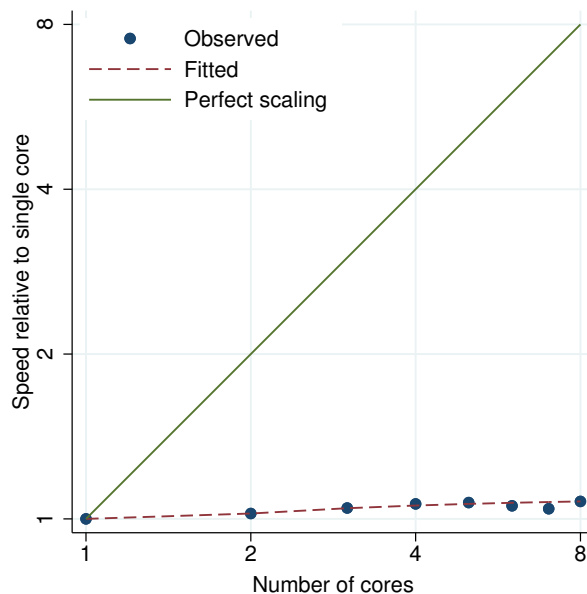
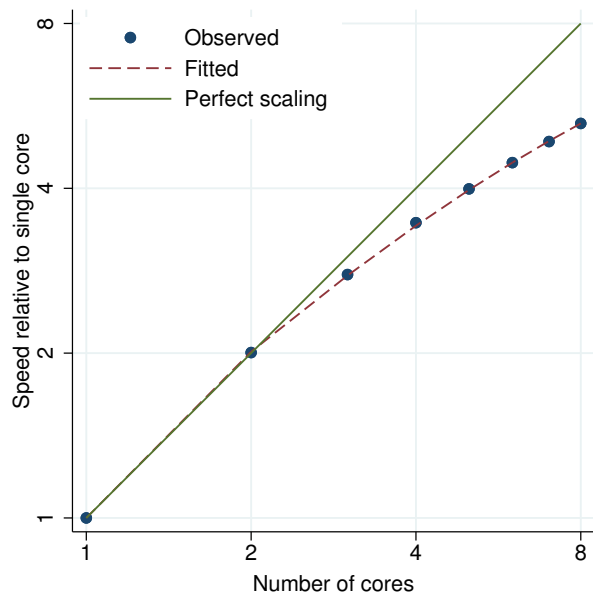
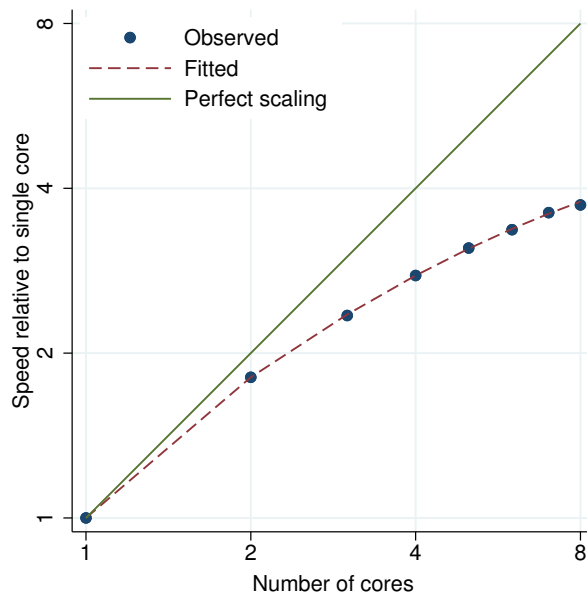


Figure 459. `sts generate` performance plot.

Figure 460. `sts graph` performance plot.Figure 461. `sts list` performance plot.Figure 462. `sts test` performance plot.Figure 463. `stset` performance plot.

Figure 464. `stsplits` performance plot.Figure 465. `stsum` performance plot.Figure 466. `stteffects ipw (weibull)` performance plot.Figure 467. `stteffects ipwra (weibull)` performance plot.

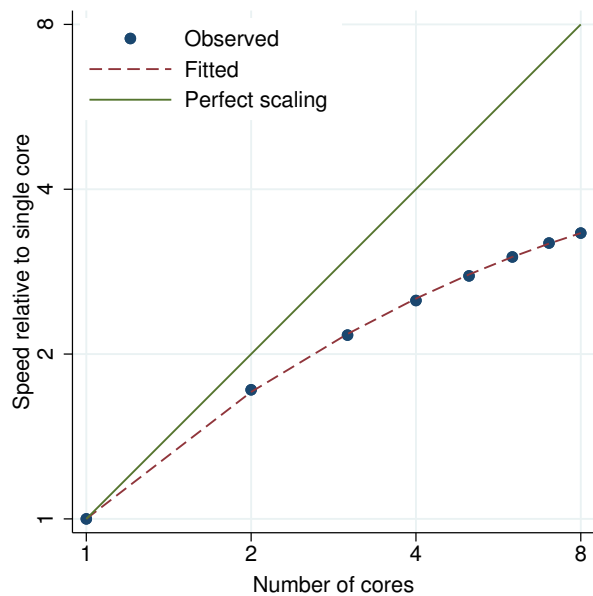


Figure 468. `stteffects ra (weibull)` performance plot.

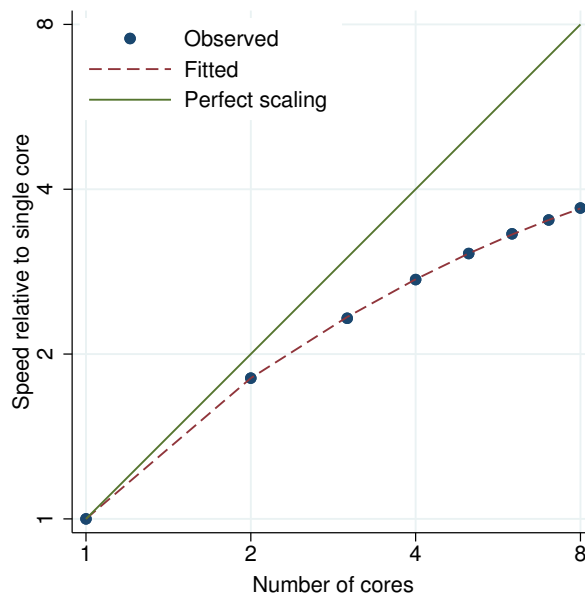


Figure 469. `stteffects wra (weibull)` performance plot.

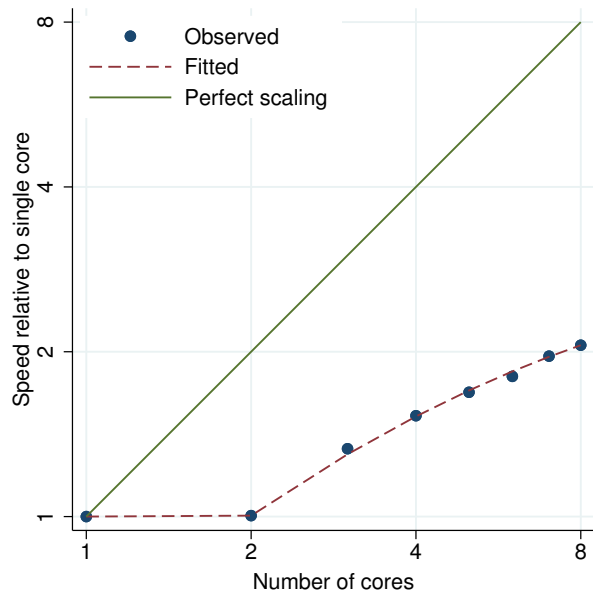


Figure 470. `stvary` performance plot.

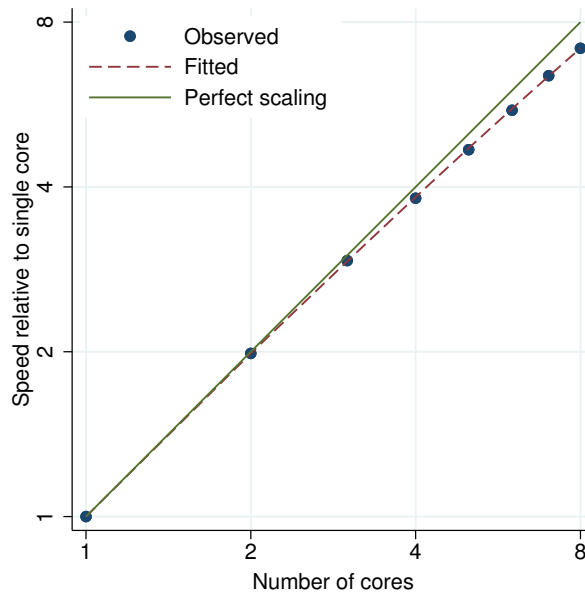
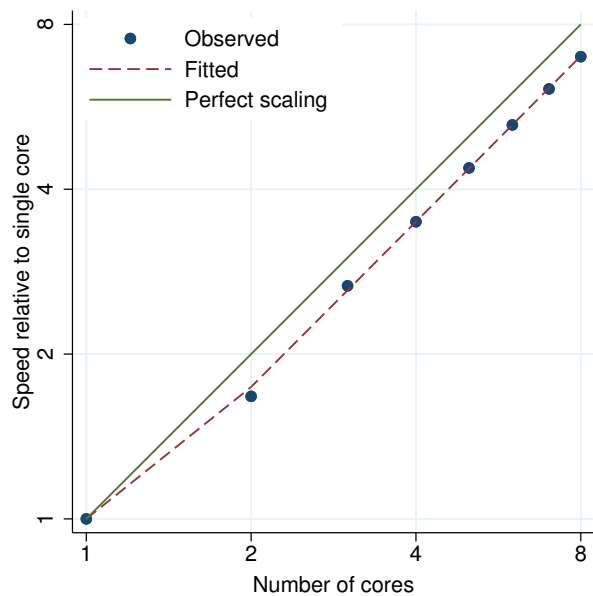
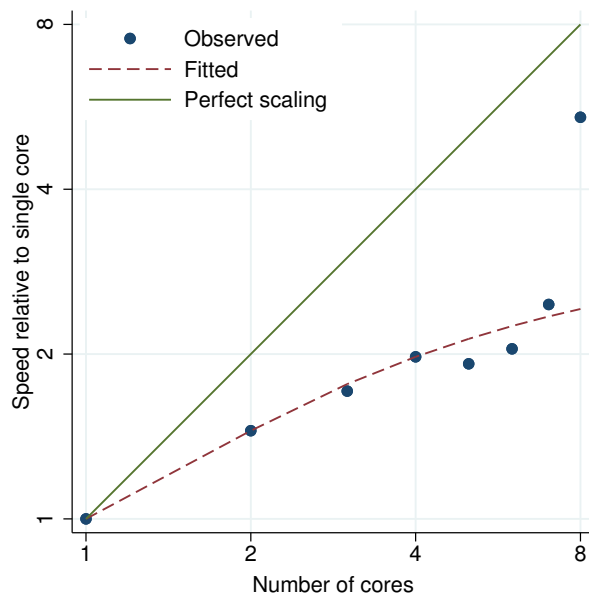
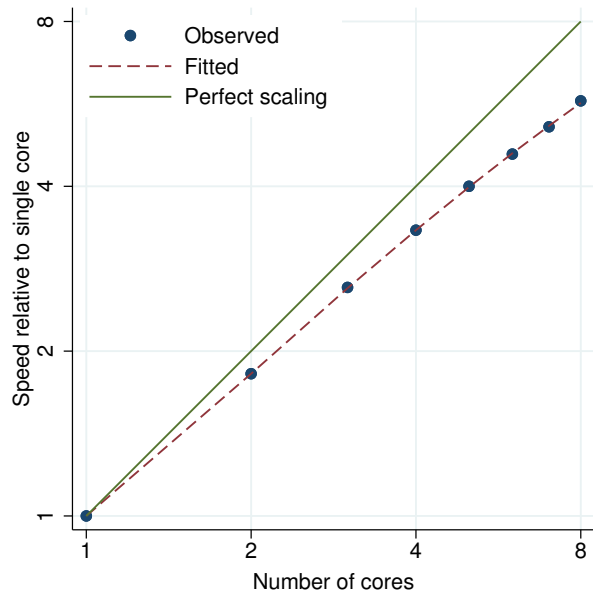
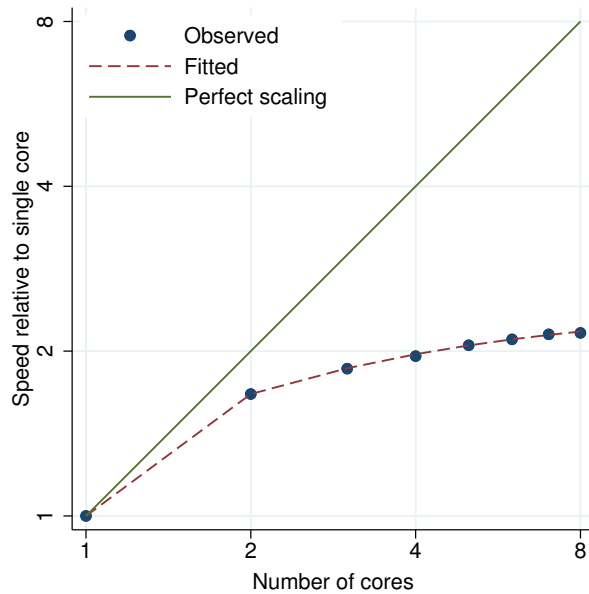


Figure 471. `suest` performance plot.

Figure 472. `summarize` performance plot.Figure 473. `sunflower` performance plot.Figure 474. `sureg` performance plot.Figure 475. `svar` performance plot.



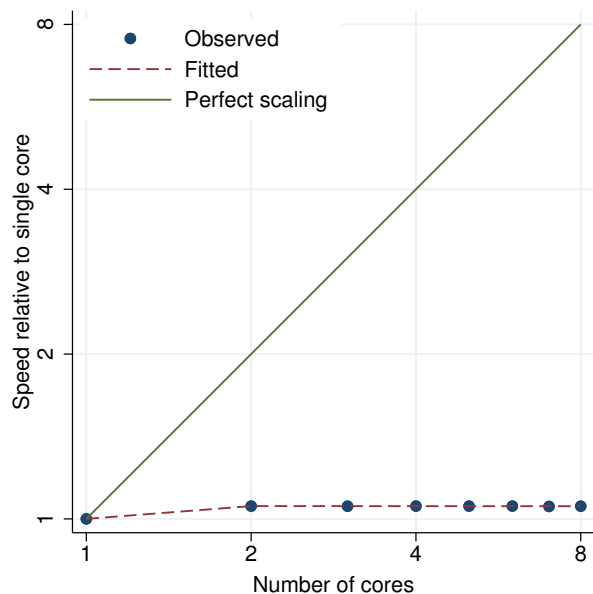


Figure 476. svmat performance plot.

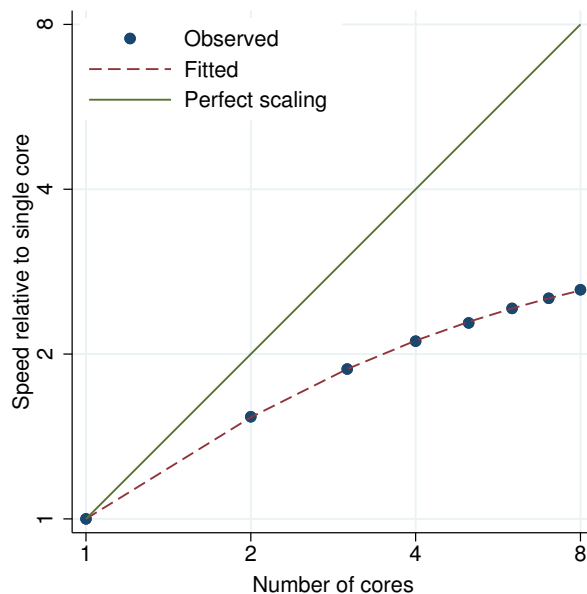


Figure 477. svy brr: logit performance plot.

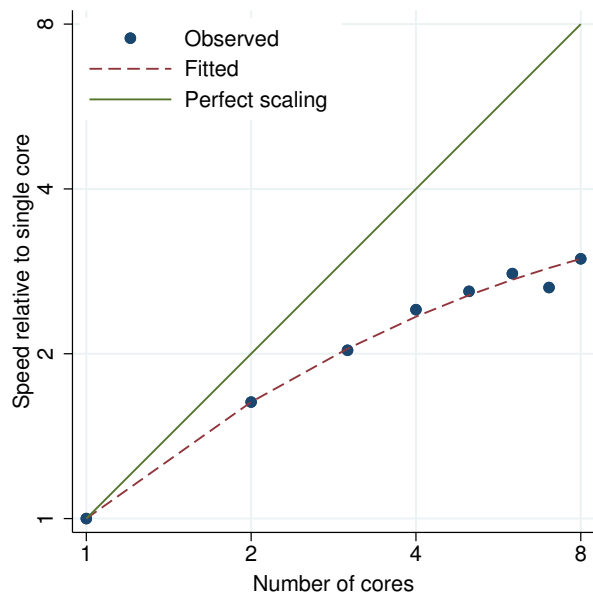


Figure 478. svy brr: poisson performance plot.

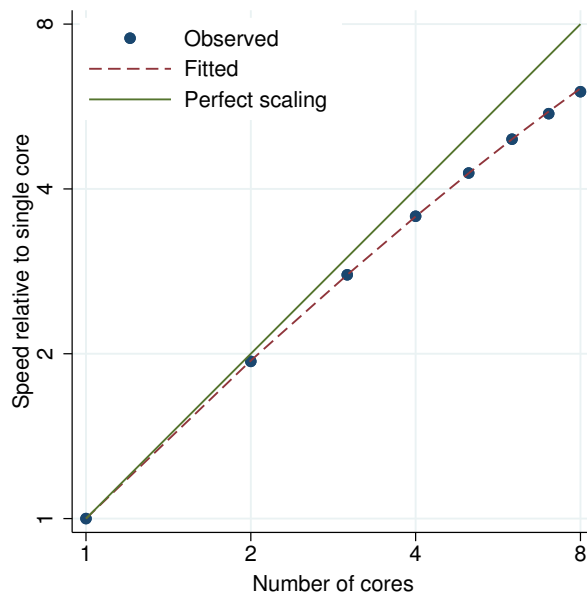


Figure 479. svy brr: regress performance plot.

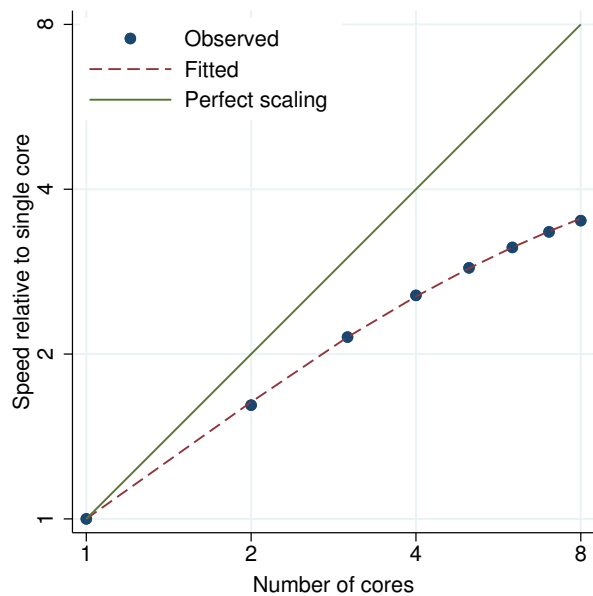


Figure 480. svy jackknife: logit performance plot.

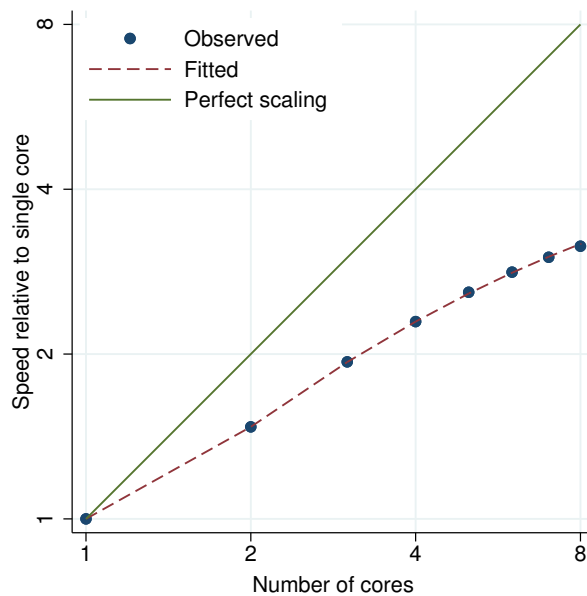


Figure 481. svy jackknife: poisson performance plot.

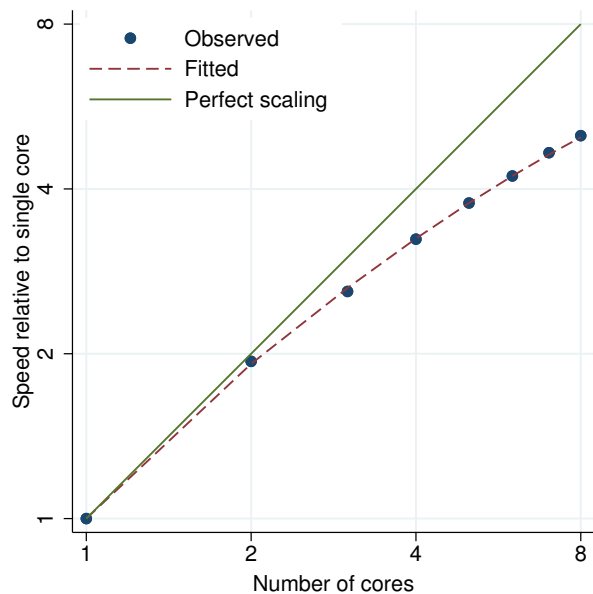


Figure 482. svy jackknife: regress performance plot.

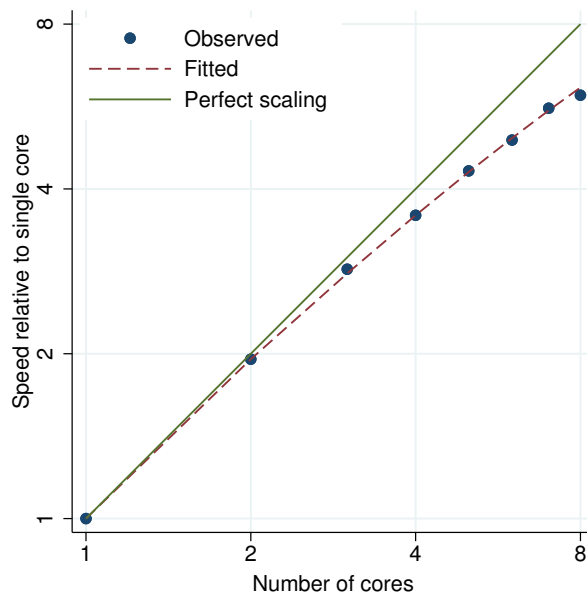
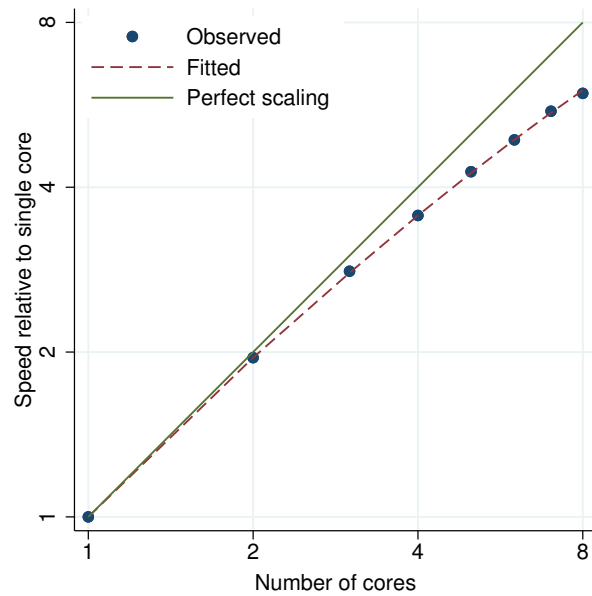
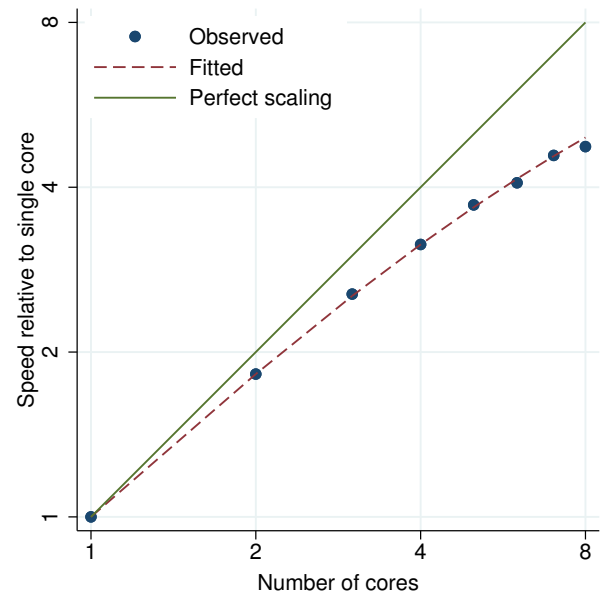
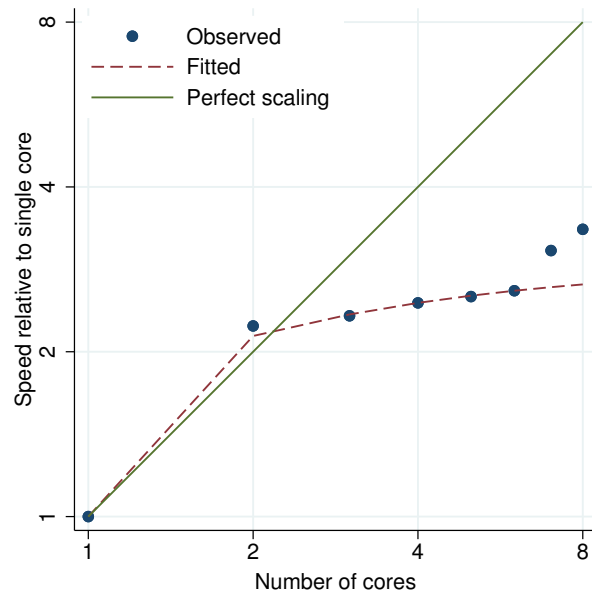
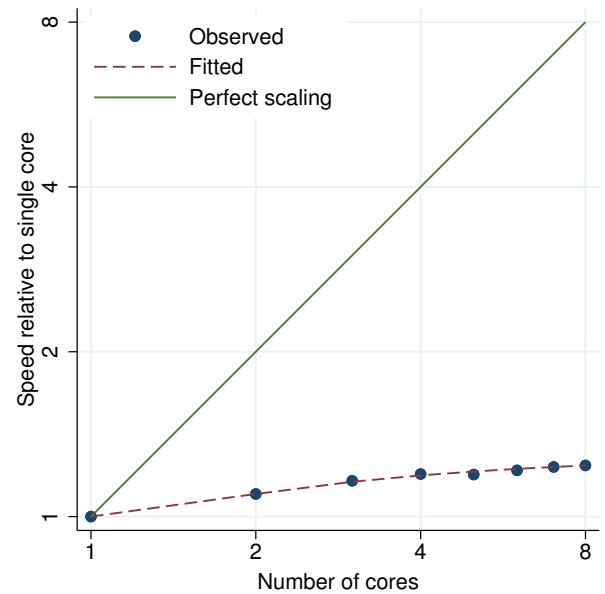
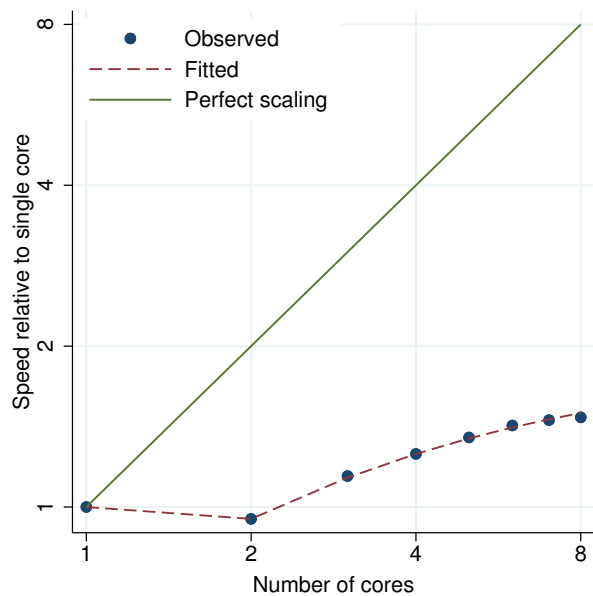
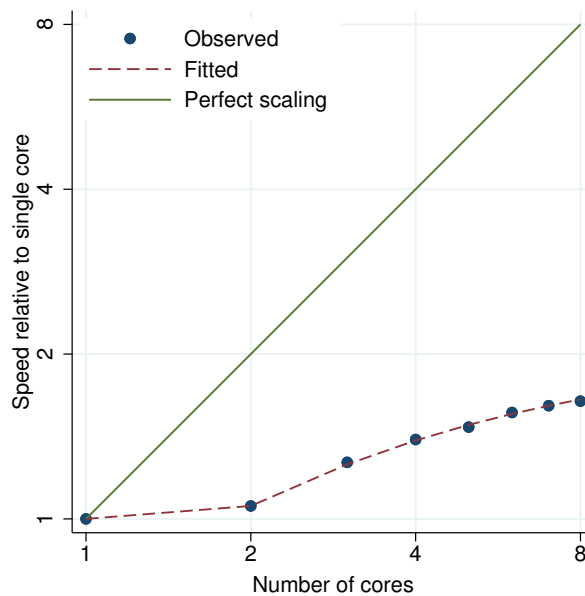
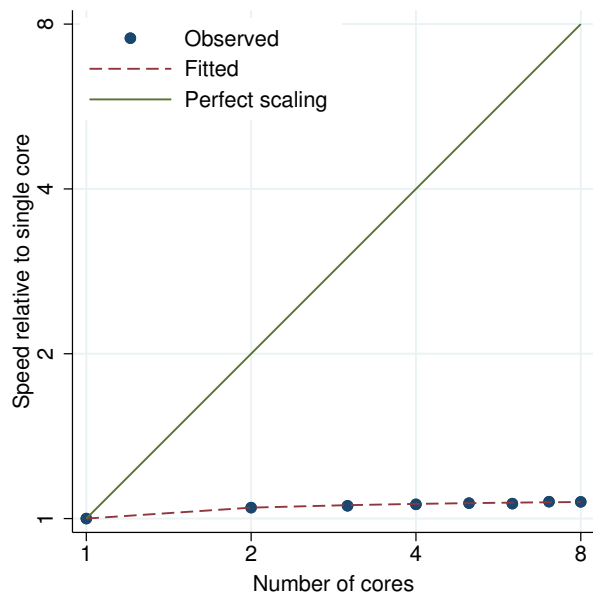
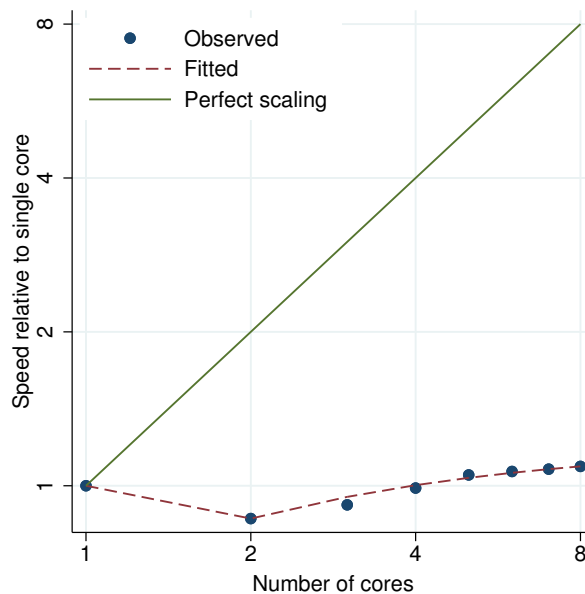
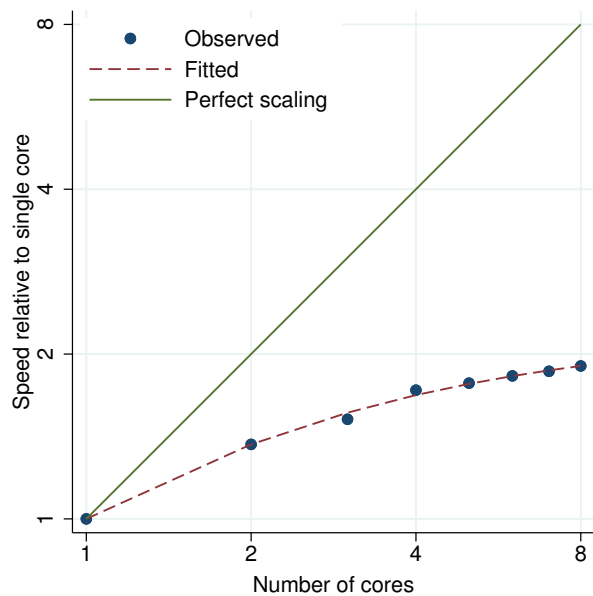
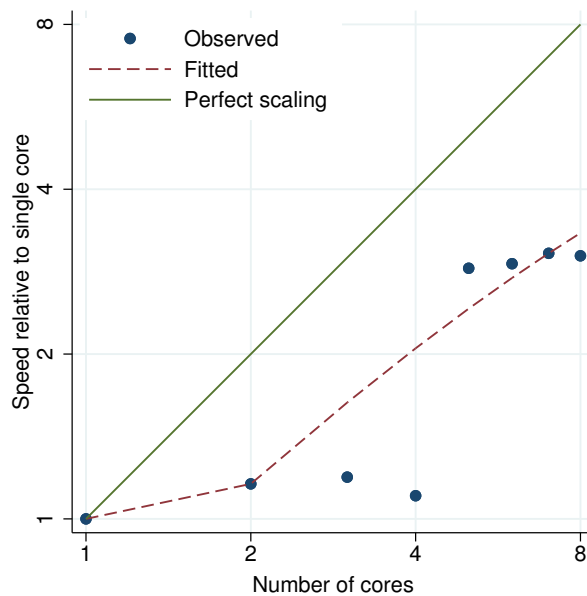
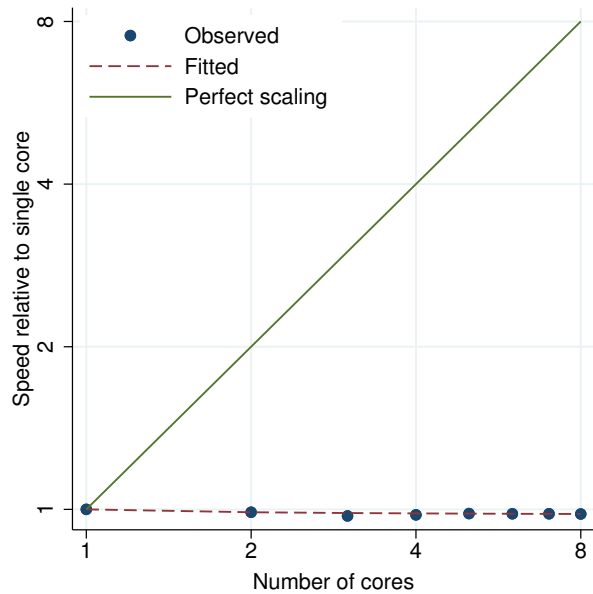
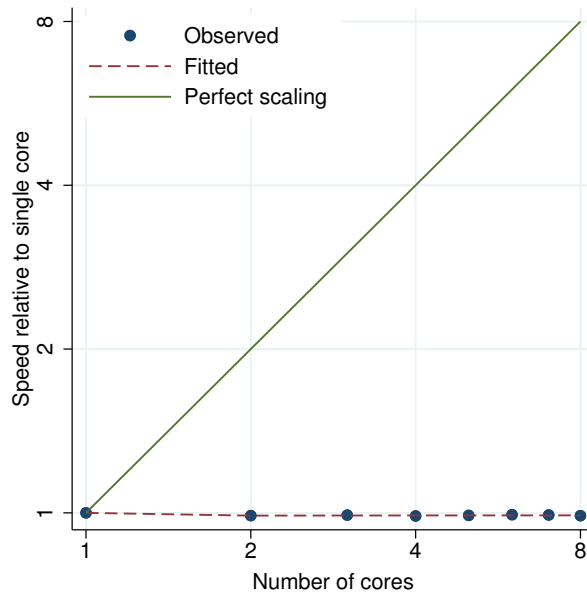


Figure 483. svy linearized: logit performance plot.

Figure 484. `svy linearized: poisson` performance plot.Figure 485. `svy linearized: regress` performance plot.Figure 486. `swilk` performance plot.Figure 487. `symmetry` performance plot.

Figure 488. `table` (one-way) performance plot.Figure 489. `table` (two-way) performance plot.Figure 490. `tabodds` performance plot.Figure 491. `tabodds` (adjusted) performance plot.

Figure 492. `tabstat` performance plot.Figure 493. `tabstat, by()` performance plot.Figure 494. `tabulate` (one-way) performance plot.Figure 495. `tabulate` (two-way) performance plot.

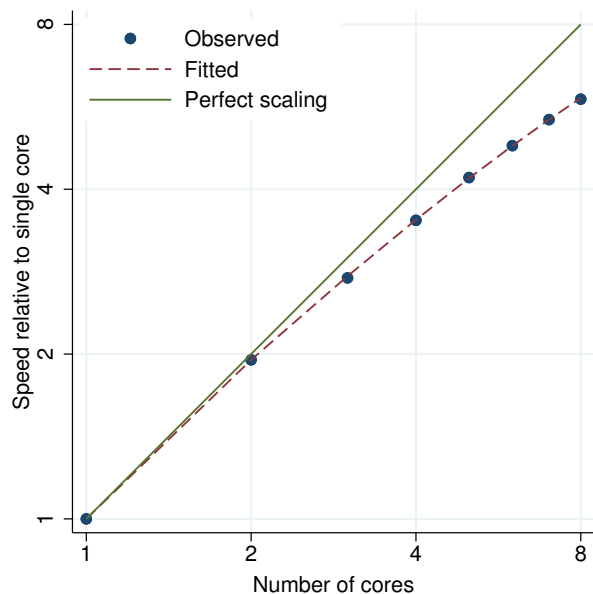


Figure 496. `teffects aipw (linear)` performance plot.

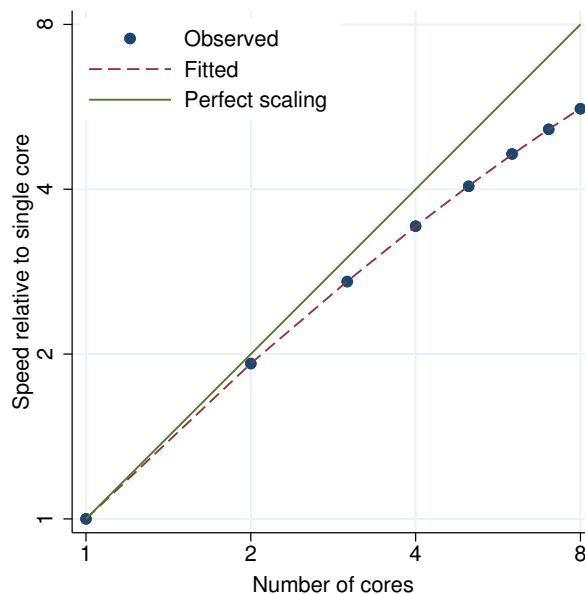


Figure 497. `teffects aipw (probit)` performance plot.

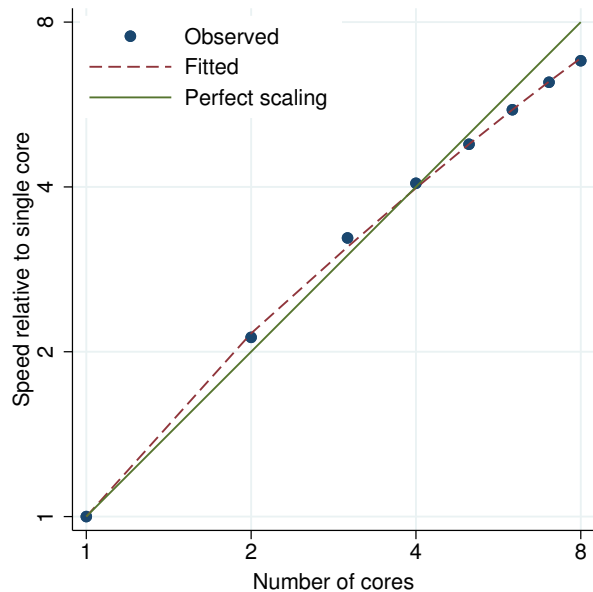


Figure 498. `teffects ipw (logit)` performance plot.

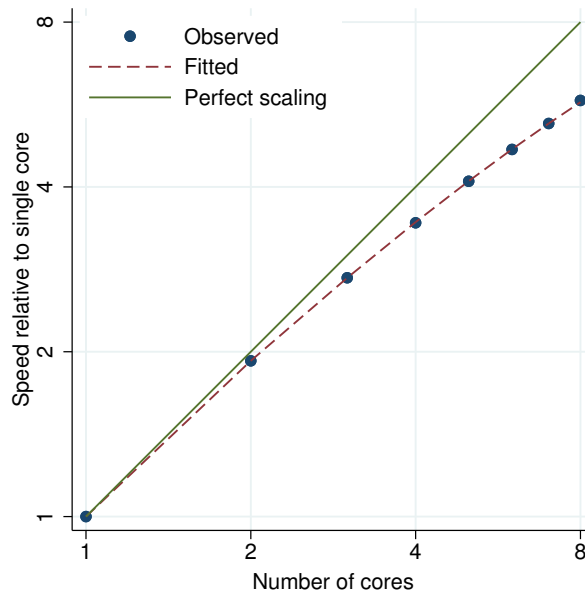


Figure 499. `teffects ipwra (linear)` performance plot.

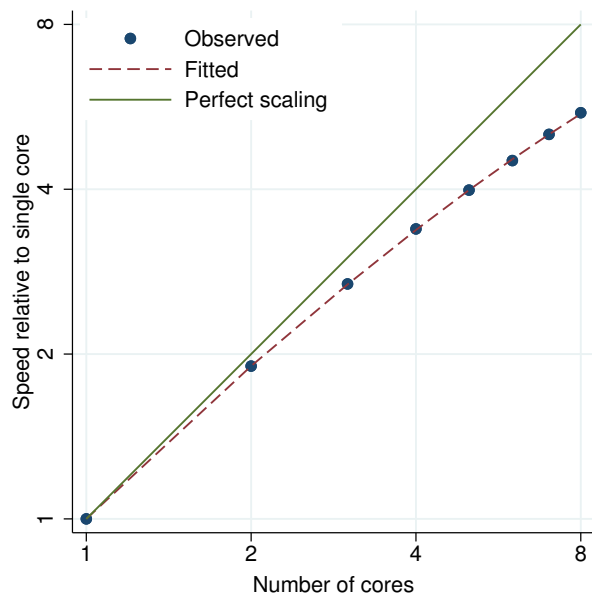


Figure 500. `teffects ipwra (probit)` performance plot.

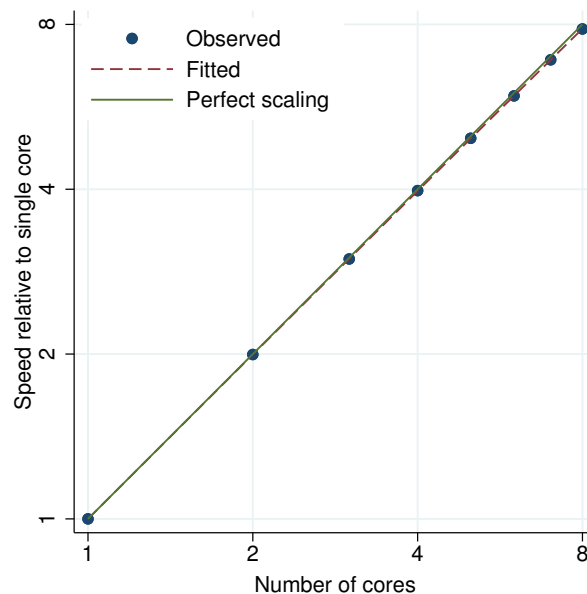


Figure 501. `teffects nnmatch` performance plot.

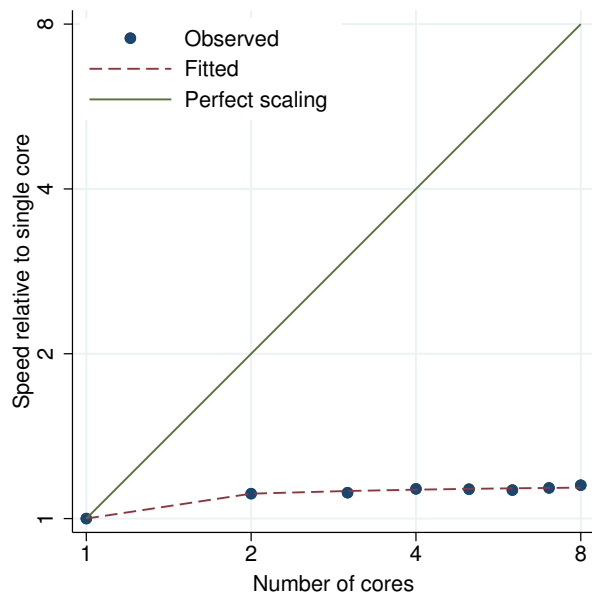


Figure 502. `teffects psmatch, logit` performance plot.

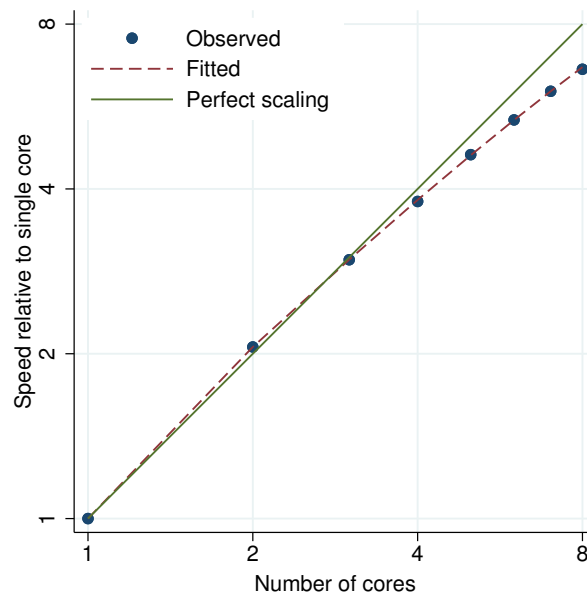


Figure 503. `teffects ra (linear)` performance plot.

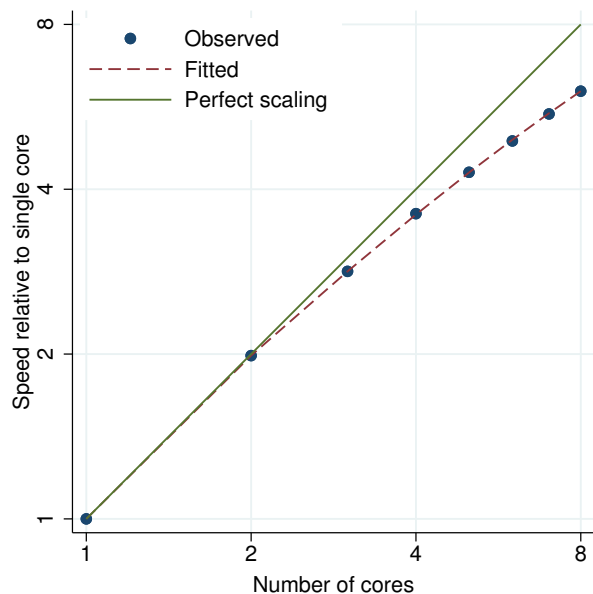


Figure 504. `teffects ra (probit)` performance plot.

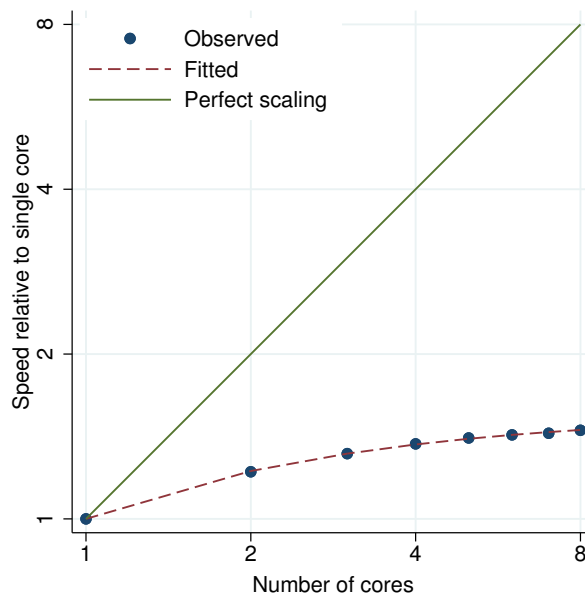


Figure 505. `tetrachoric` performance plot.

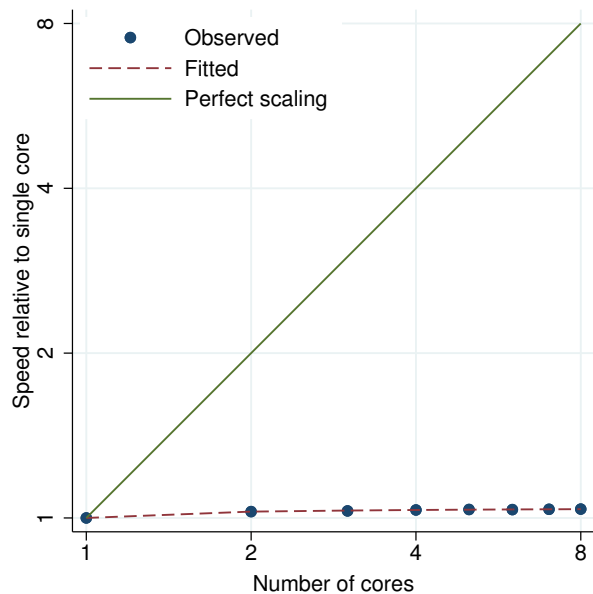


Figure 506. `threshold, threshvar()` performance plot.

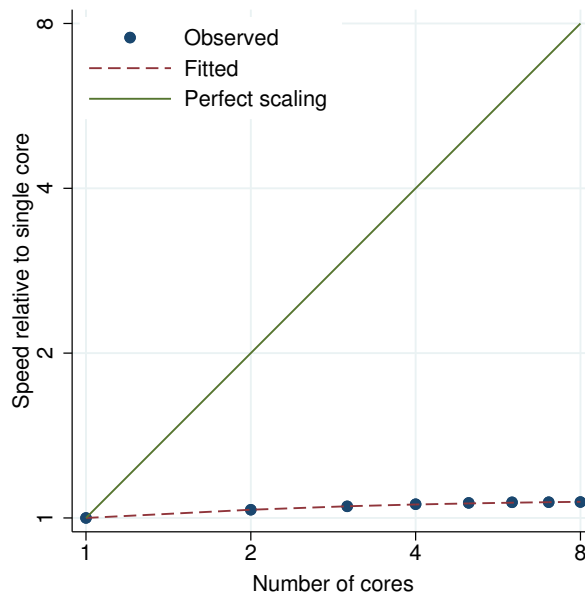
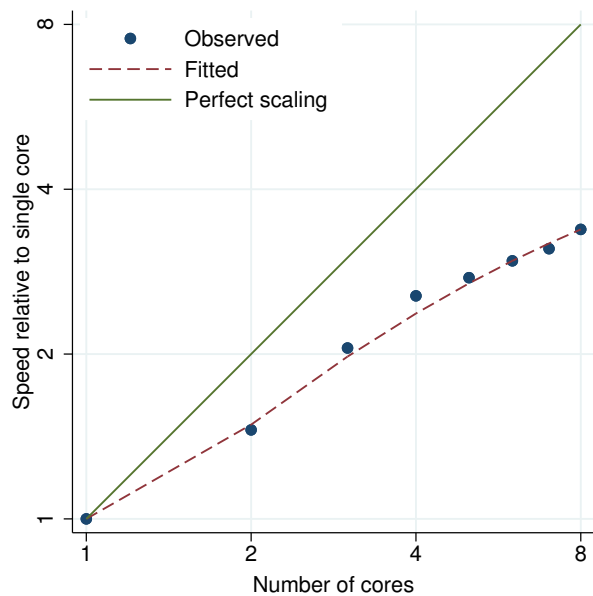
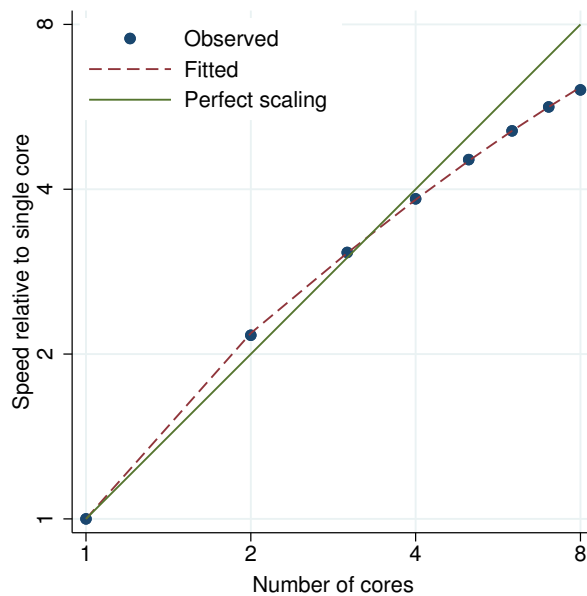
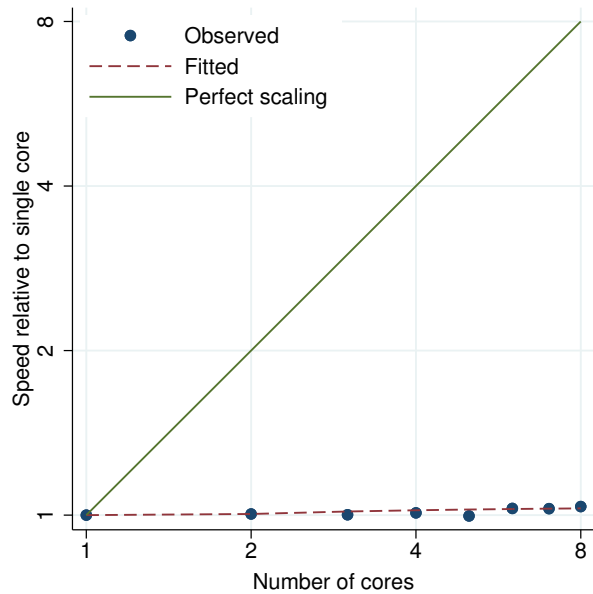
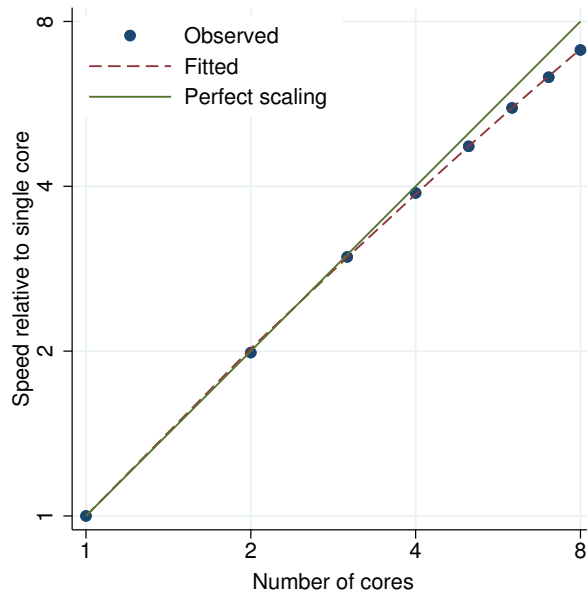
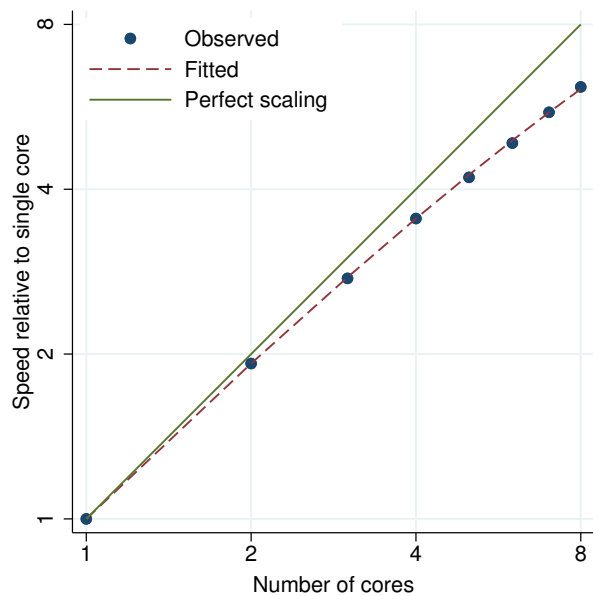
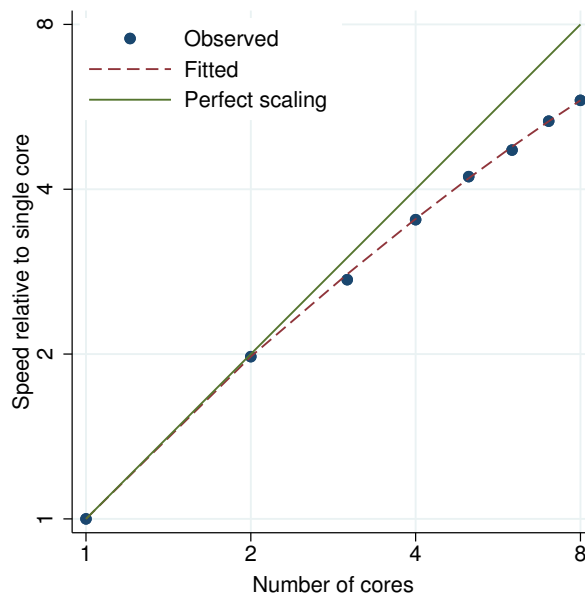
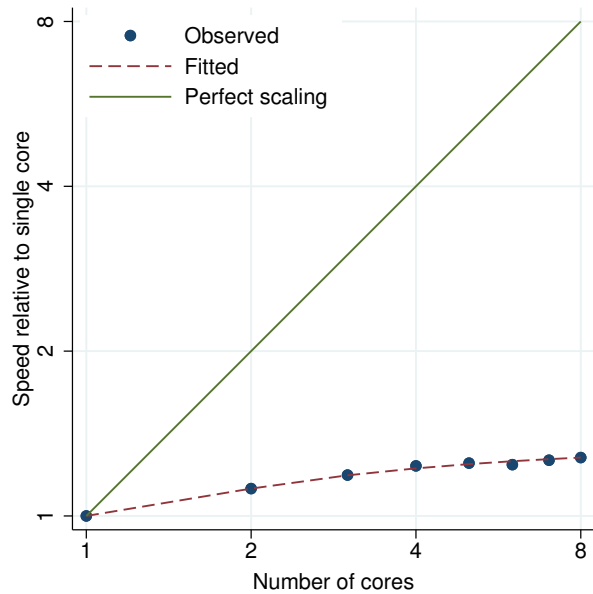
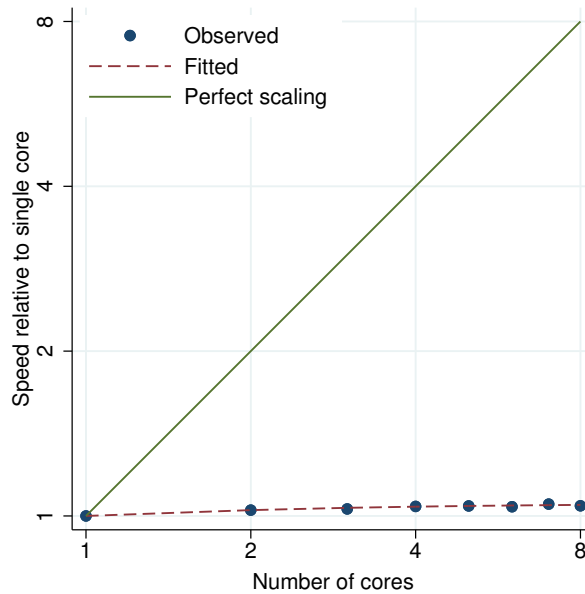
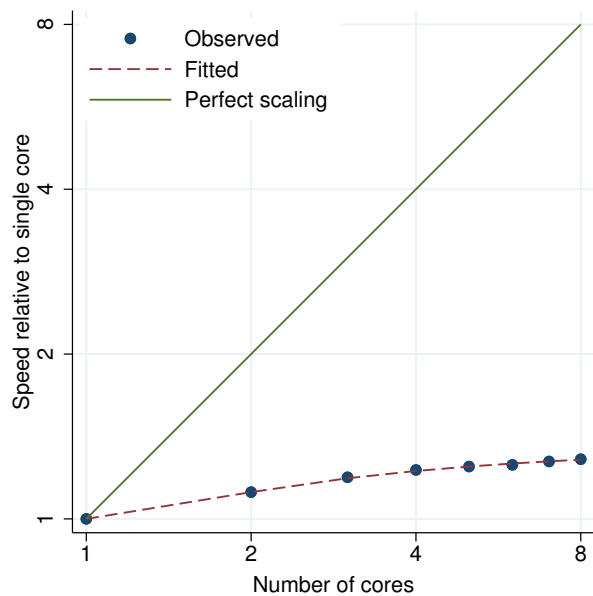
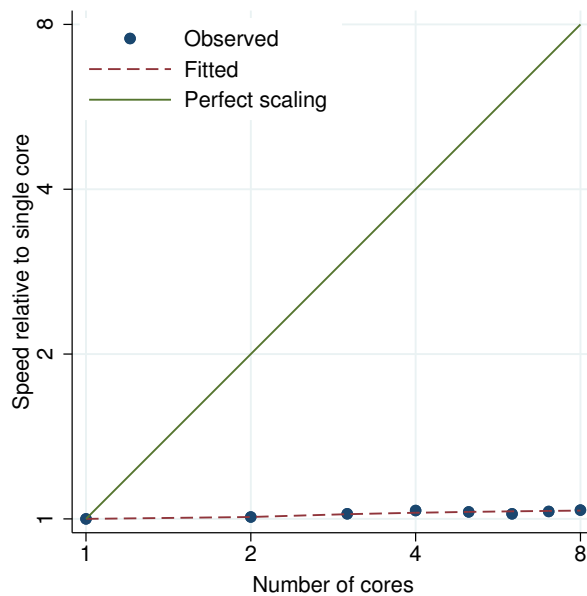
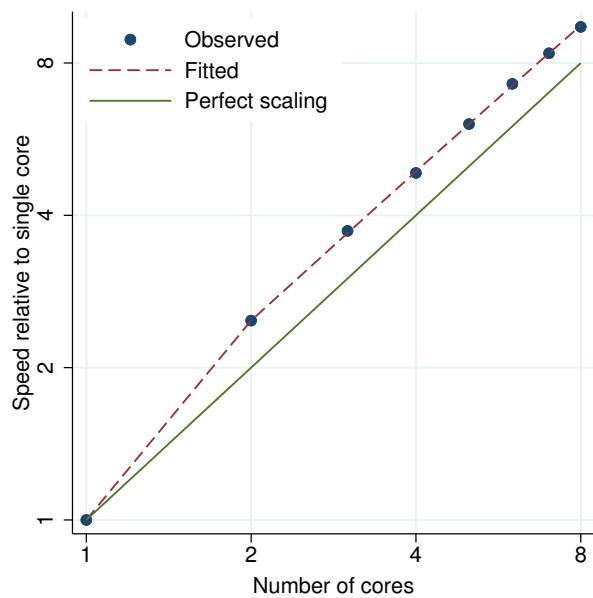
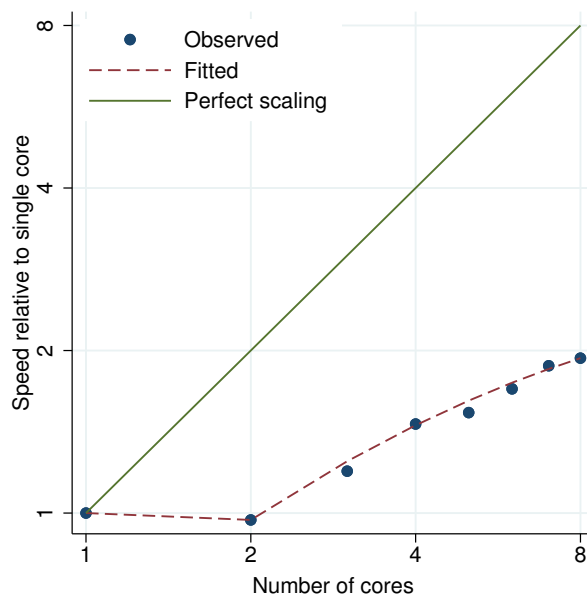


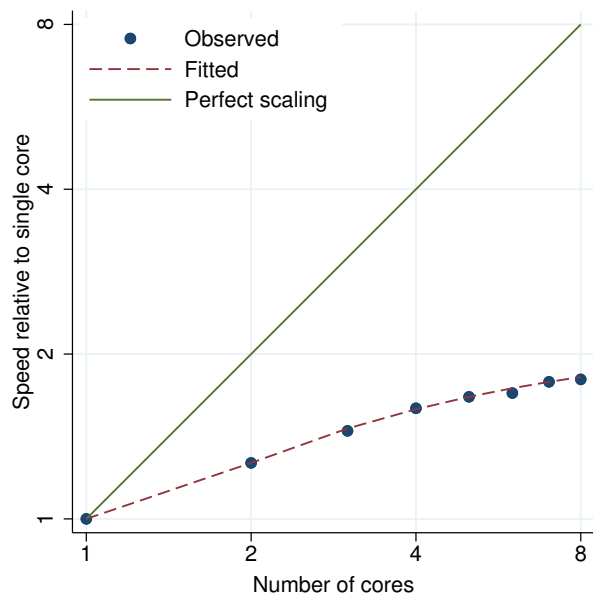
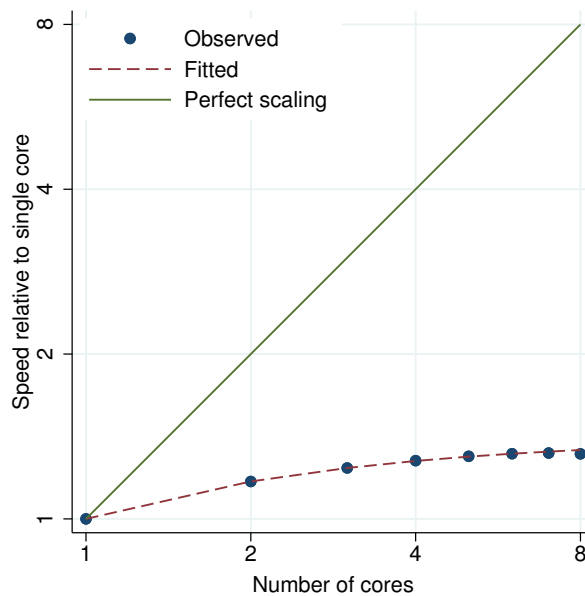
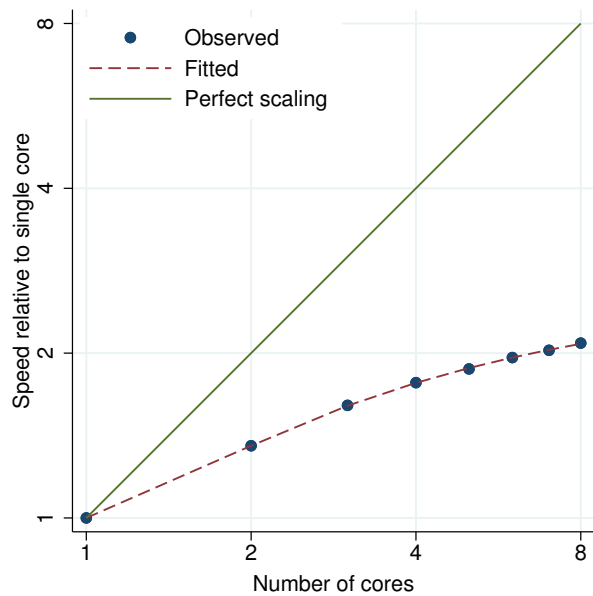
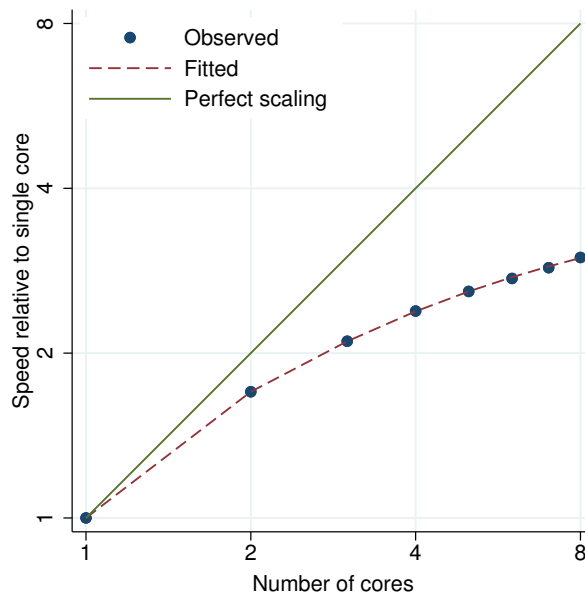
Figure 507. `threshold, threshvar() regionvars()` performance plot.

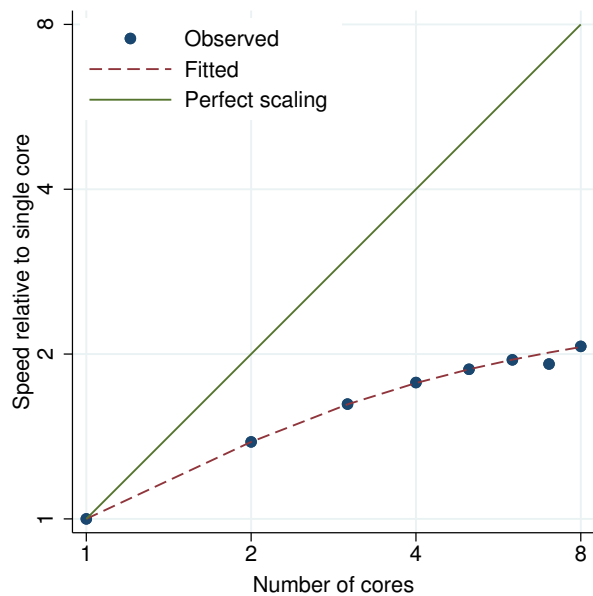
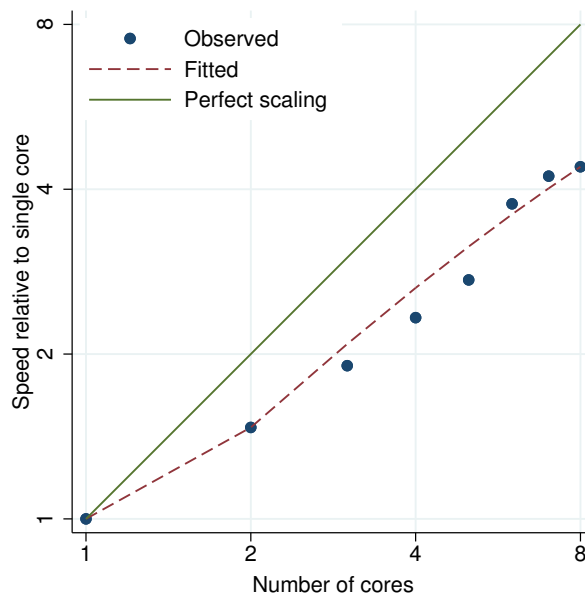
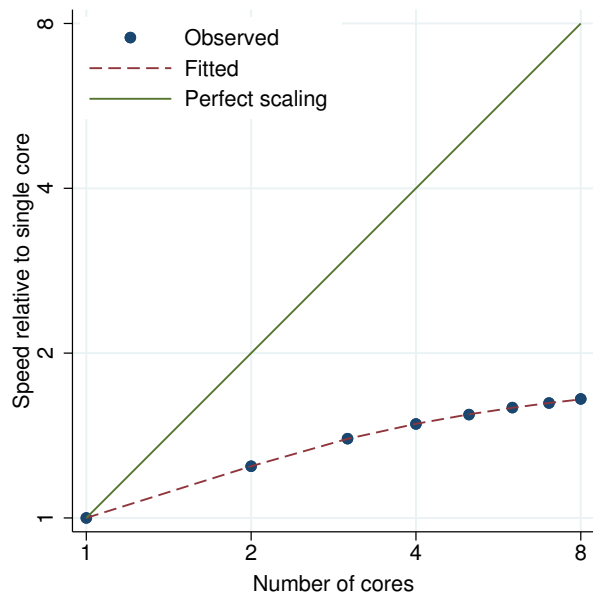
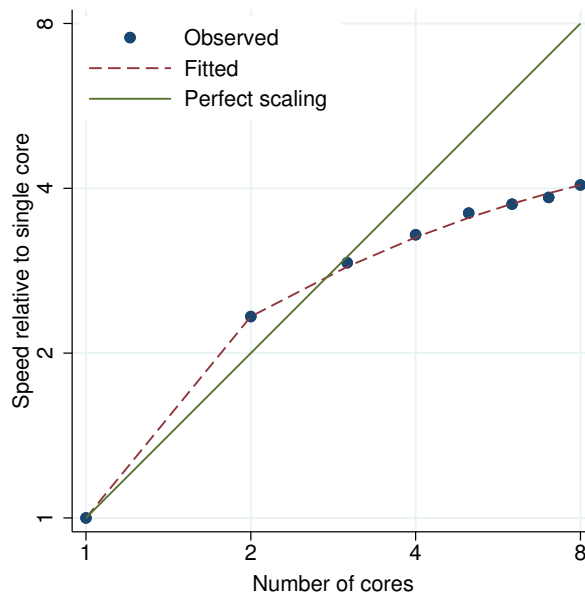


Figure 508. `tnbreg` performance plot.Figure 509. `tobit` performance plot.Figure 510. `tostring` performance plot.Figure 511. `total` performance plot.

Figure 512. `tpoisson` performance plot.Figure 513. `truncreg` performance plot.Figure 514. `tsfilter bk` performance plot.Figure 515. `tsfilter bw` performance plot.

Figure 516. `tsfilter cf` performance plot.Figure 517. `tsfilter hp` performance plot.Figure 518. `tsrevar` performance plot.Figure 519. `tsset` performance plot.

Figure 520. `tssmooth exp` performance plot.Figure 521. `tssmooth ma` performance plot.Figure 522. `tttest1` performance plot.Figure 523. `tttest2` performance plot.

Figure 524. `ttest, by()` performance plot.Figure 525. `twoway fpfit` performance plot.Figure 526. `twoway lfitci` performance plot.Figure 527. `twoway mband` performance plot.

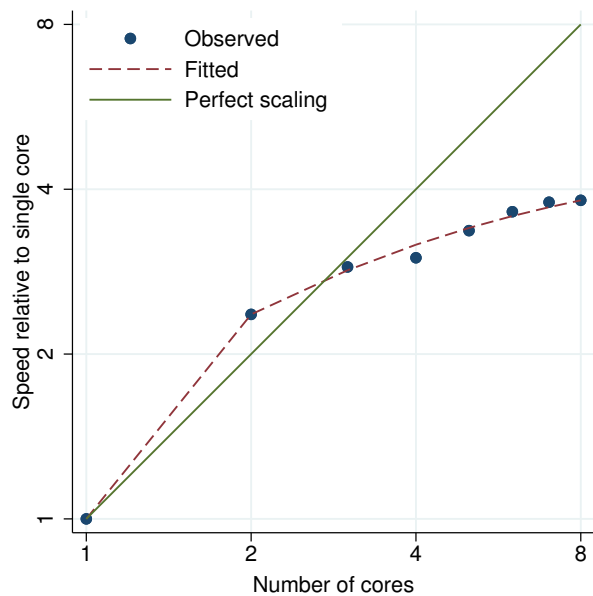


Figure 528. twoway mspline performance plot.

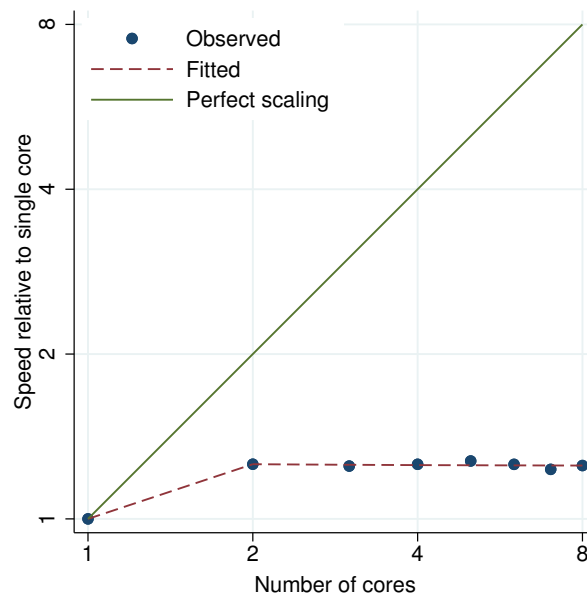


Figure 529. ucm, model(rwdrift) performance plot.

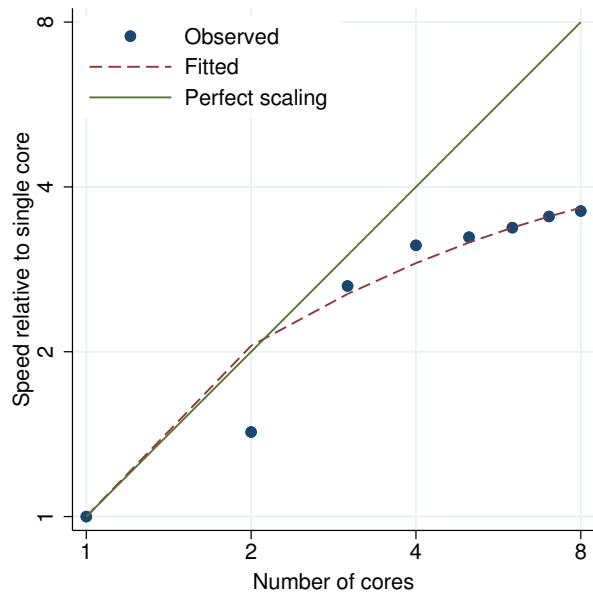


Figure 530. var performance plot.

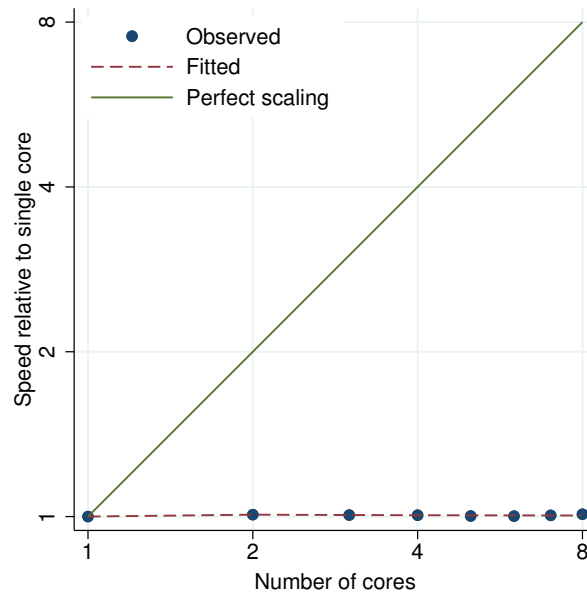
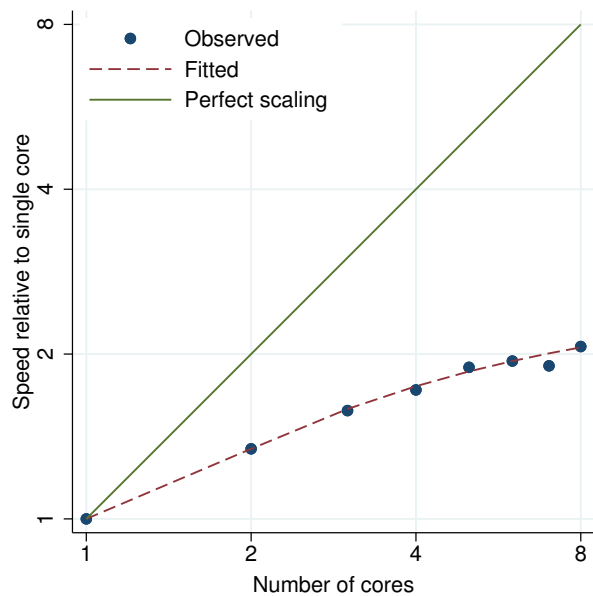
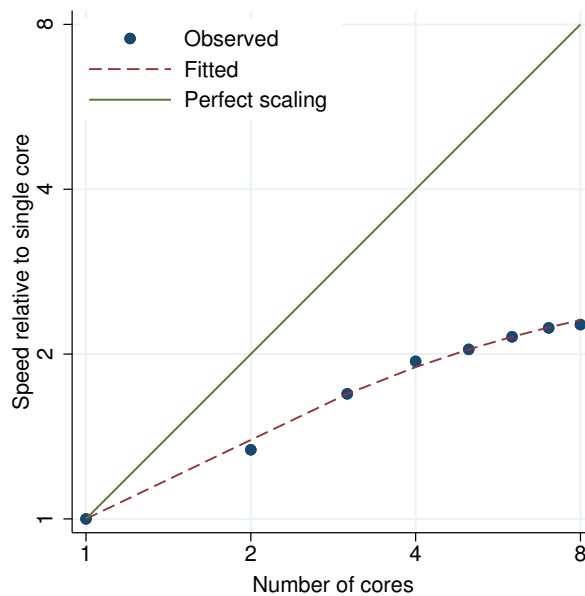
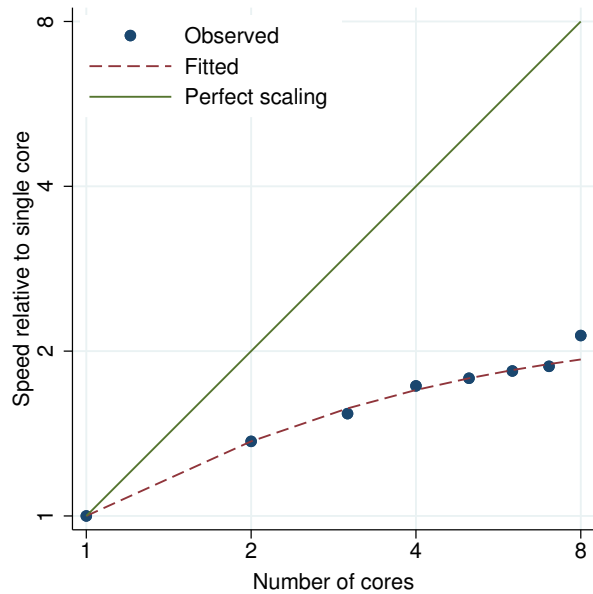
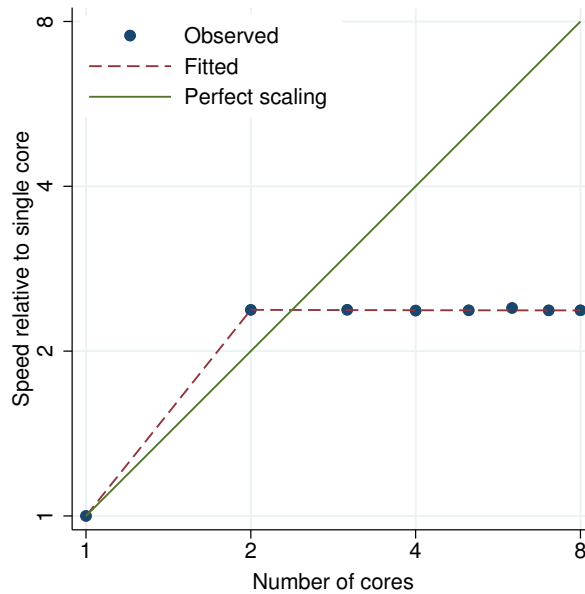
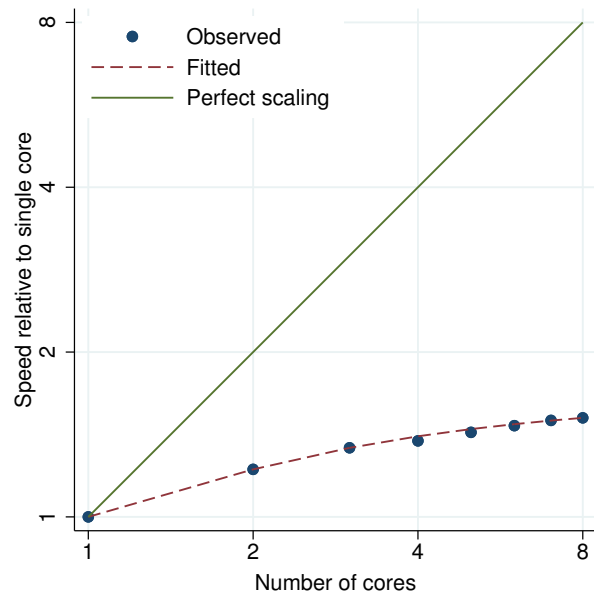
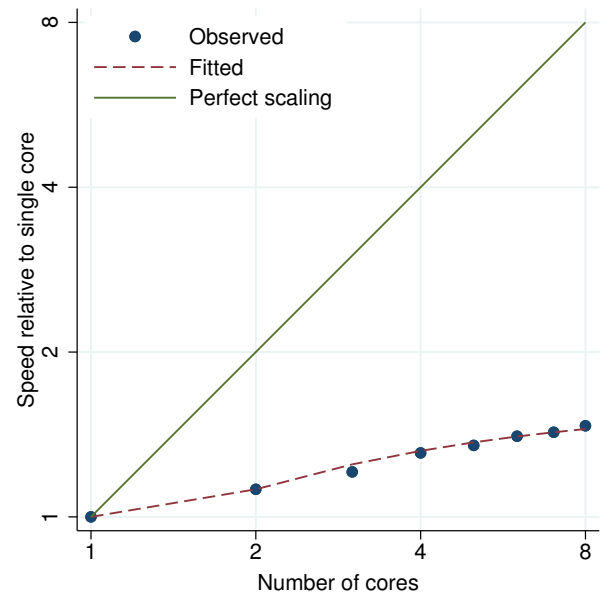
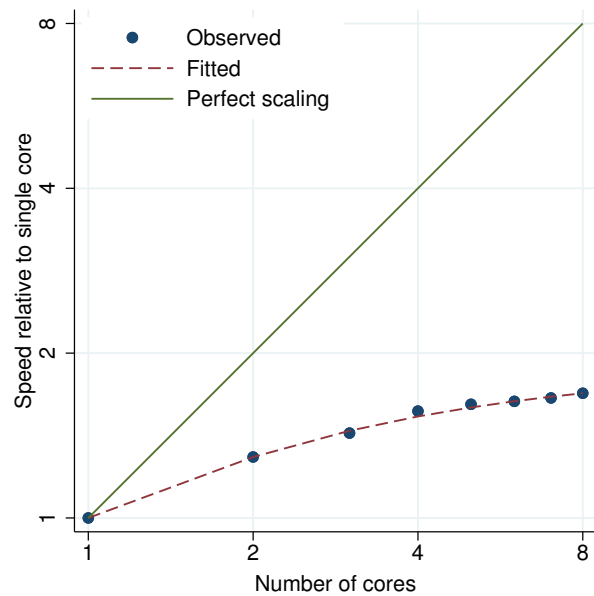
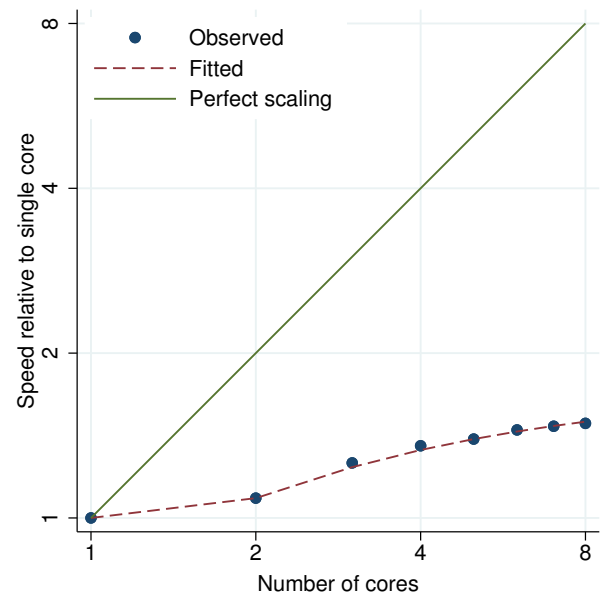
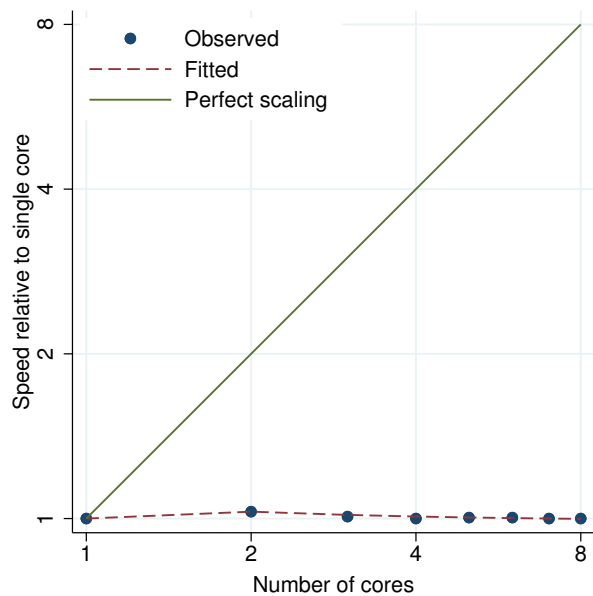
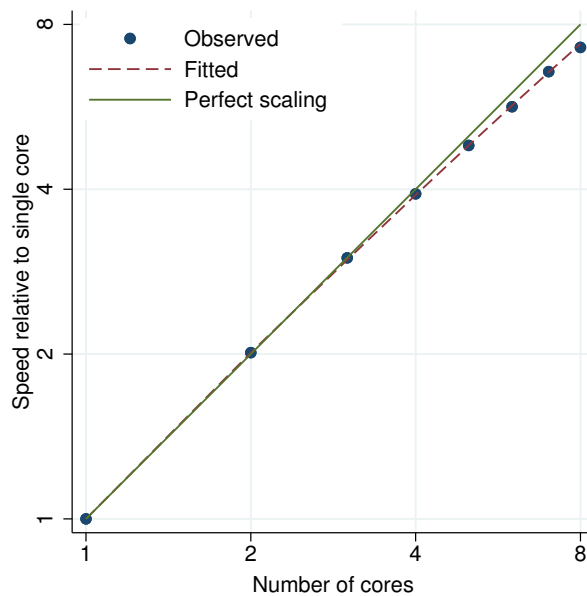
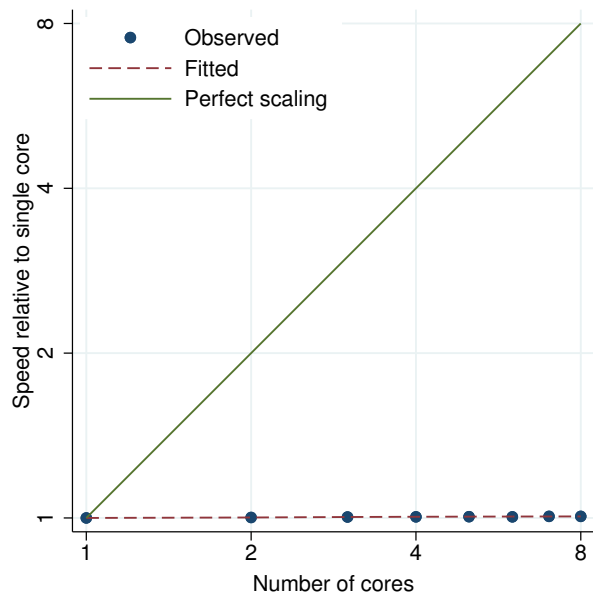
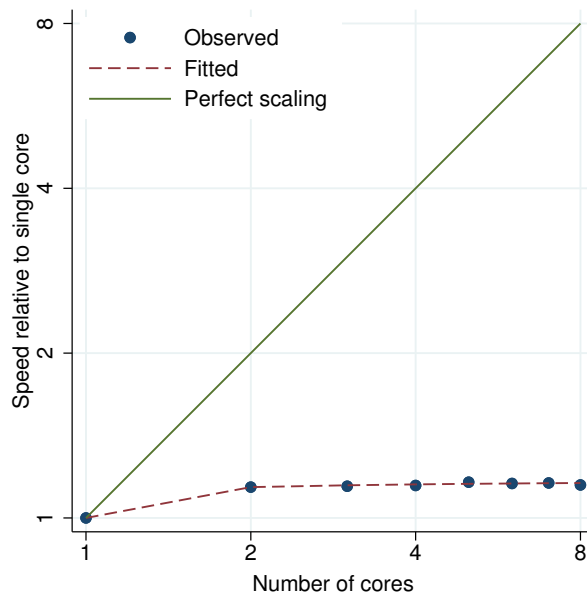


Figure 531. vargranger performance plot.

Figure 532. `varlmar` performance plot.Figure 533. `varnorm` performance plot.Figure 534. `varsoc` performance plot.Figure 535. `varstable` performance plot.

Figure 536. `vec` performance plot.Figure 537. `vec1mar` performance plot.Figure 538. `vecnorm` performance plot.Figure 539. `vecrank` performance plot.



Figure 540. `vecstable` performance plot.Figure 541. `vwls` performance plot.Figure 542. `wntestb` performance plot.Figure 543. `wntestq` performance plot.

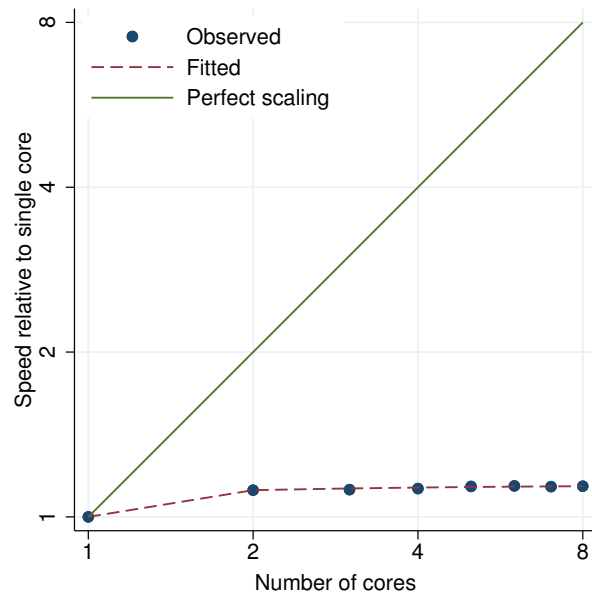


Figure 544. xcorr performance plot.

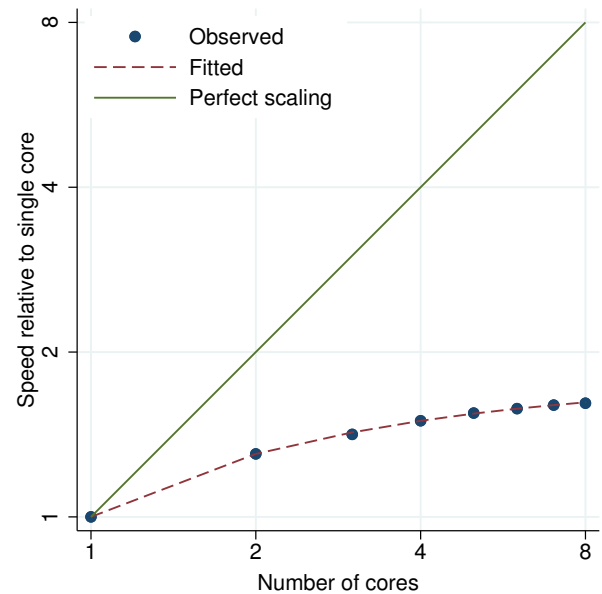


Figure 545. xpologit performance plot.

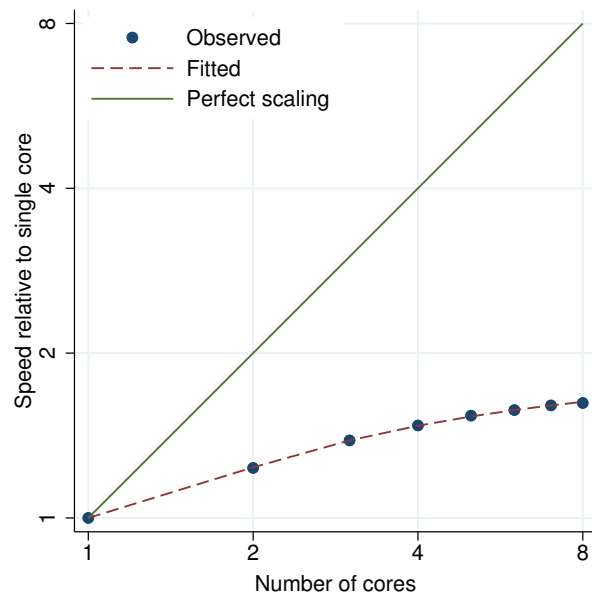


Figure 546. xpoppoisson performance plot.

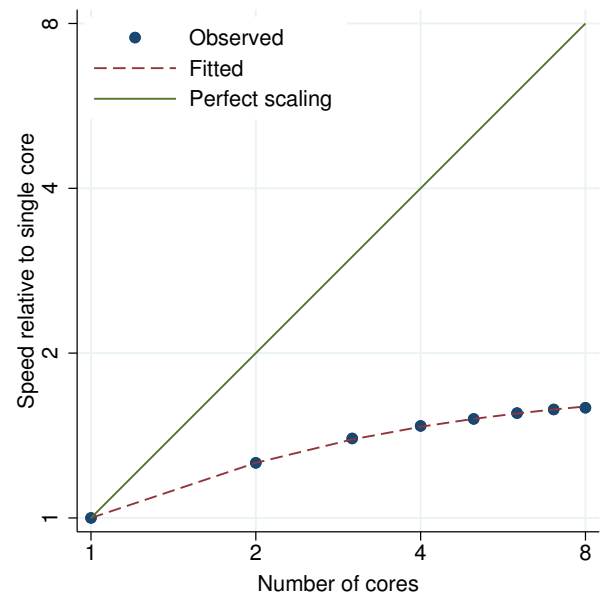


Figure 547. xporegress performance plot.

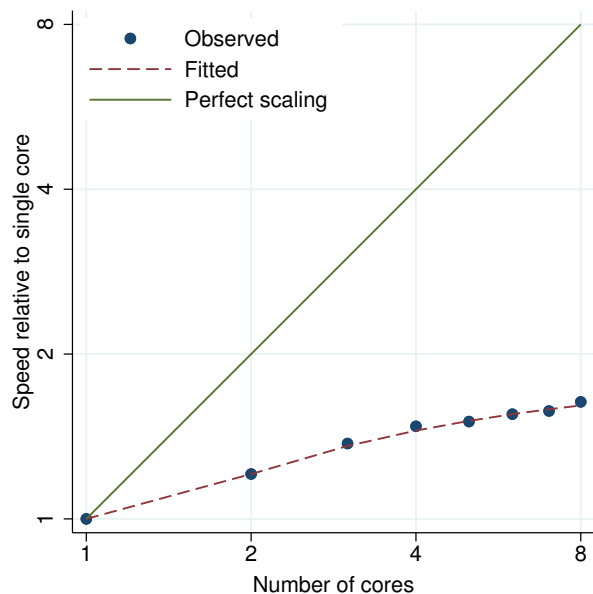


Figure 548. xtabond performance plot.

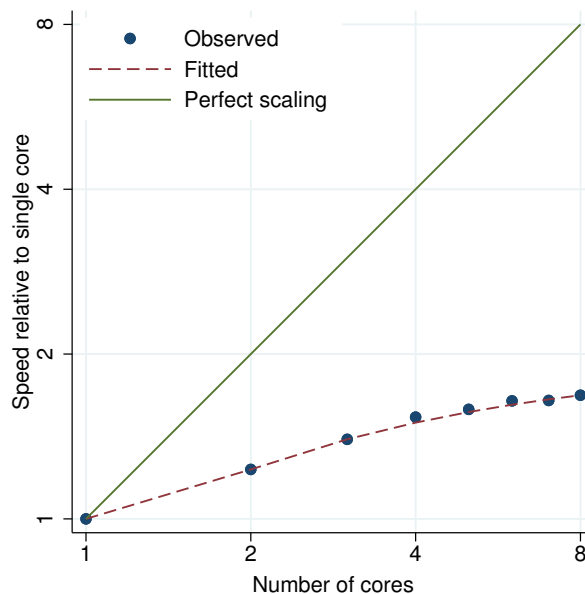


Figure 549. xtabond, twostep performance plot.

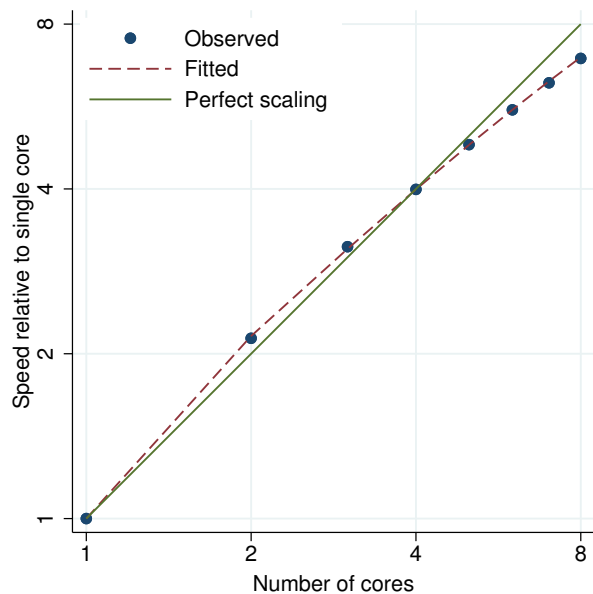


Figure 550. xtcloglog, re performance plot.

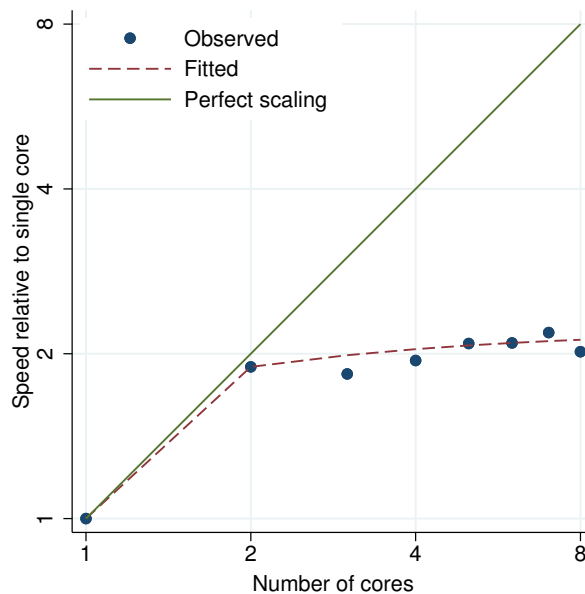


Figure 551. xtdata, be performance plot.

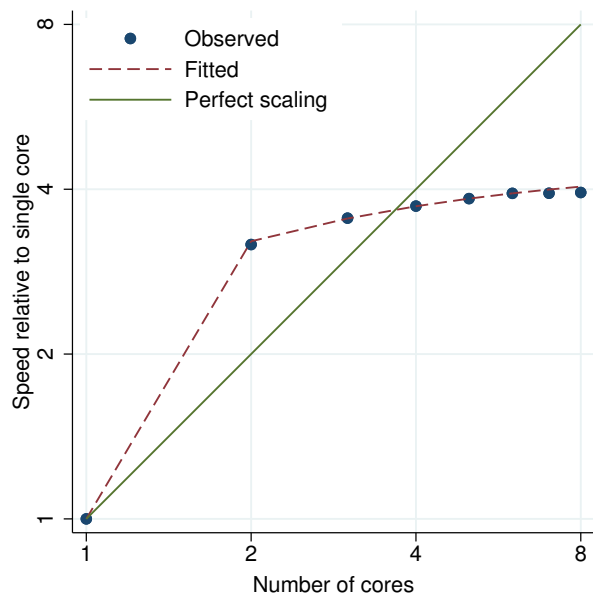


Figure 552. xtdata, fe performance plot.

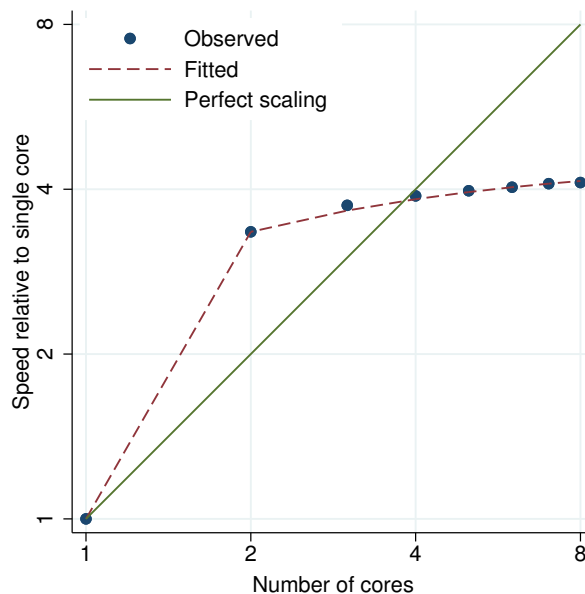


Figure 553. xtdata, re performance plot.

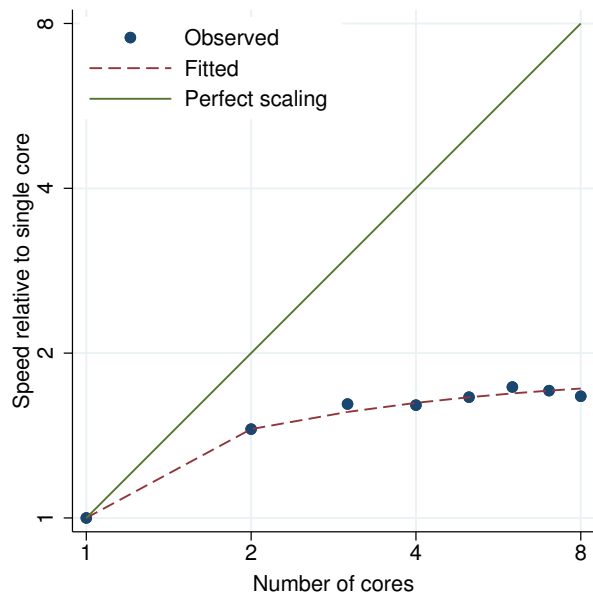


Figure 554. xtdpd performance plot.

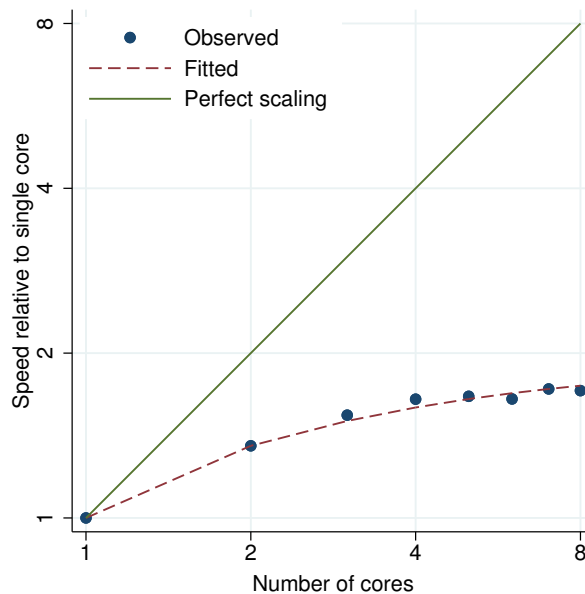
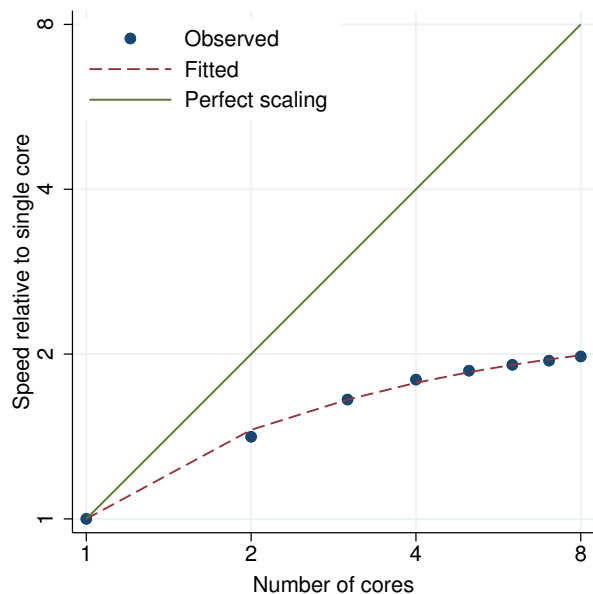
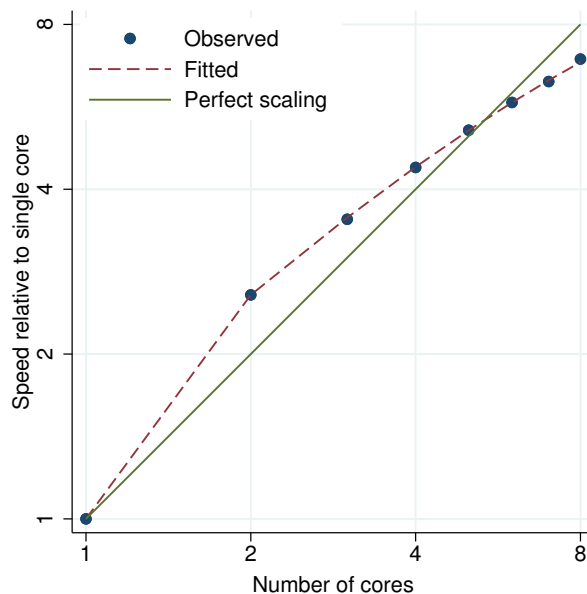
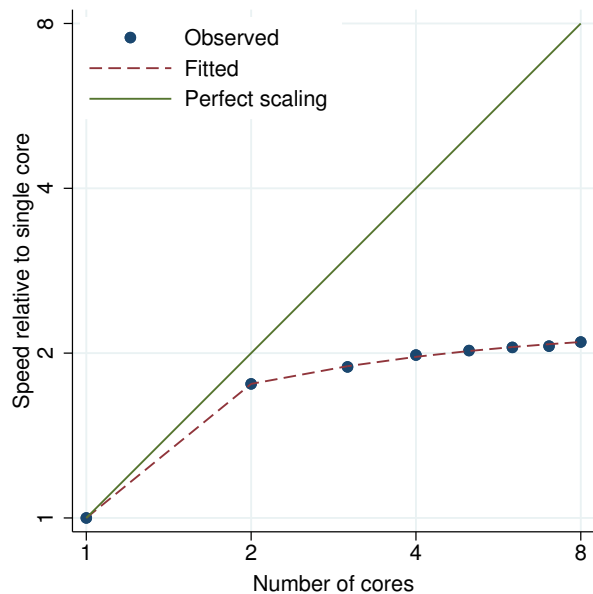
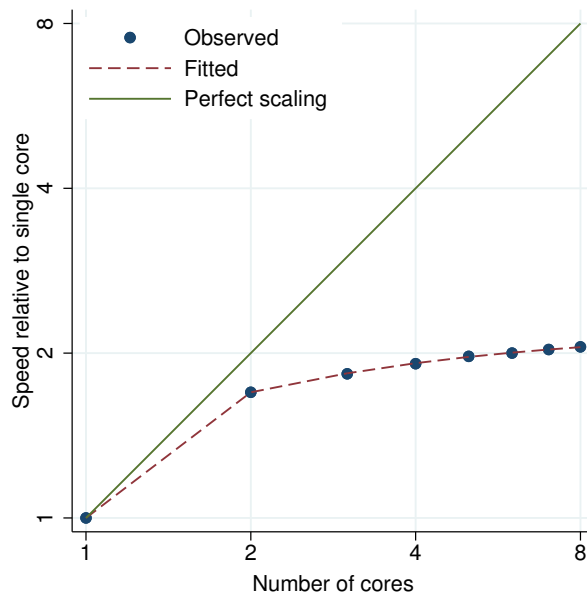


Figure 555. xtdpdsys performance plot.

Figure 556. `xtregress` performance plot.Figure 557. `xtfrontier` performance plot.Figure 558. `xtgee`, `family(gaussian)`  
`corr(ar2)` performance plot.Figure 559. `xtgee`, `fam(gauss)`  
`corr(unstruct)` performance plot.

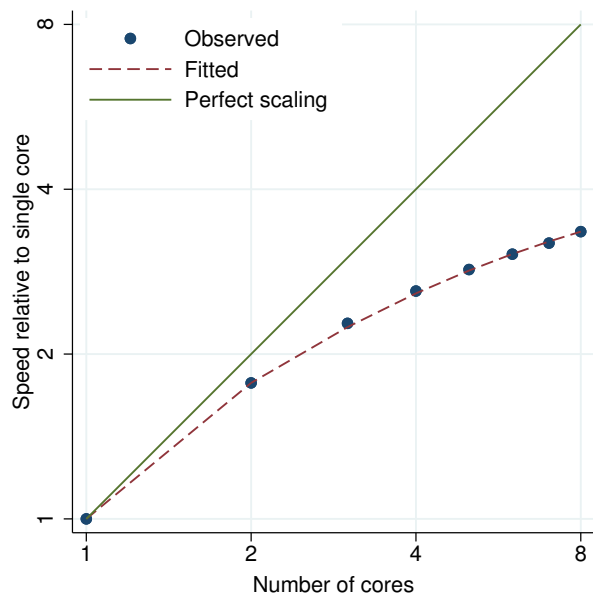


Figure 560. xtcloglog, pa performance plot.

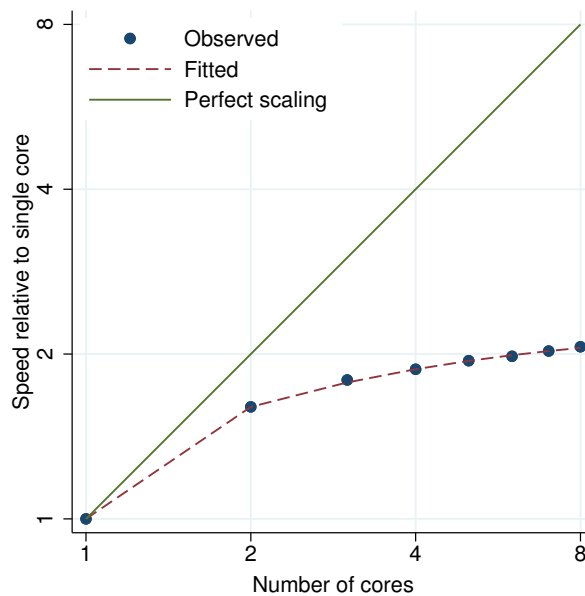


Figure 561. xtlogit, pa performance plot.

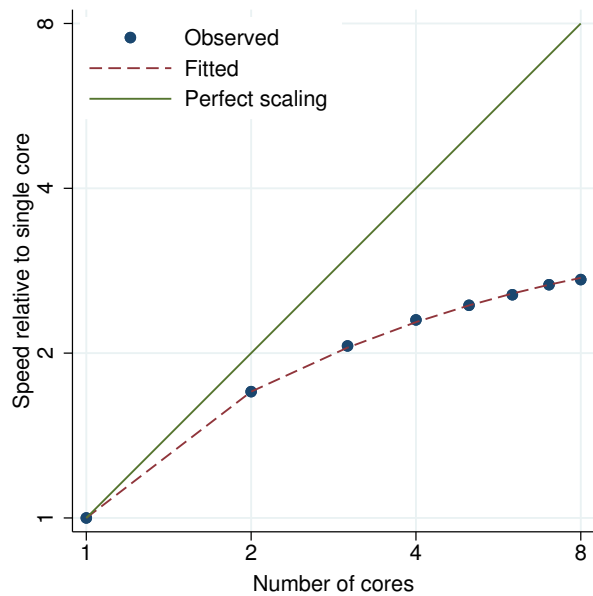


Figure 562. xtnbreg, pa performance plot.

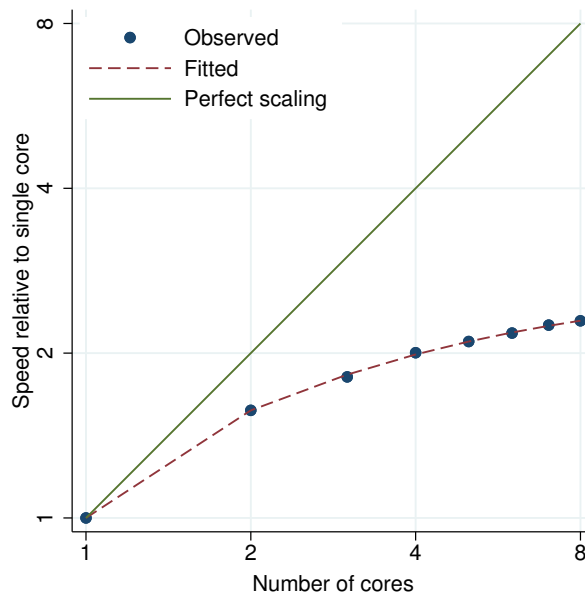


Figure 563. xtpoisson, pa performance plot.

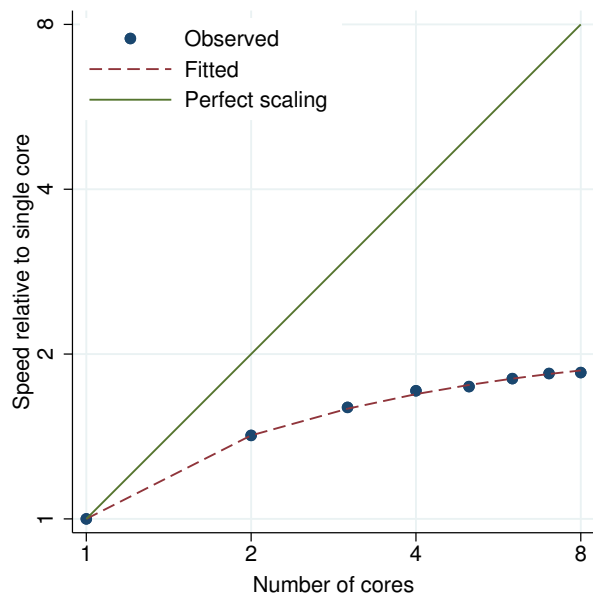


Figure 564. xtprobit, pa performance plot.

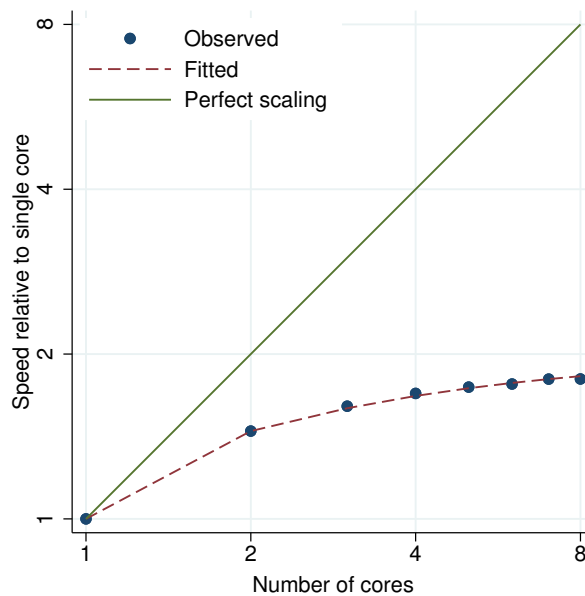


Figure 565. xtreg, pa performance plot.

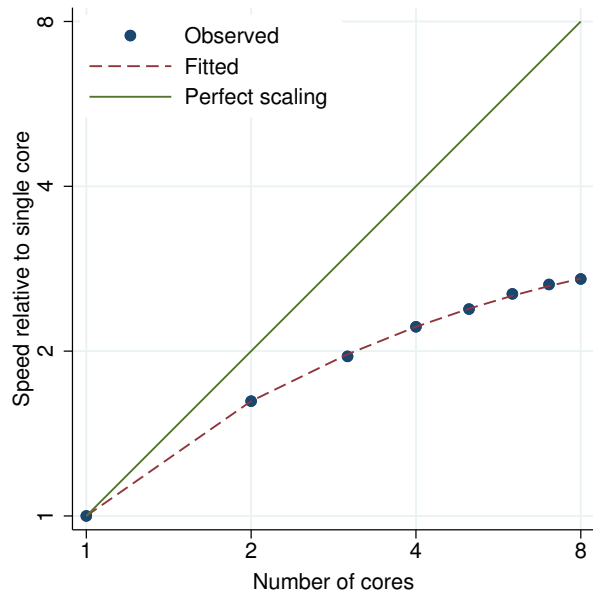


Figure 566. xtglm performance plot.

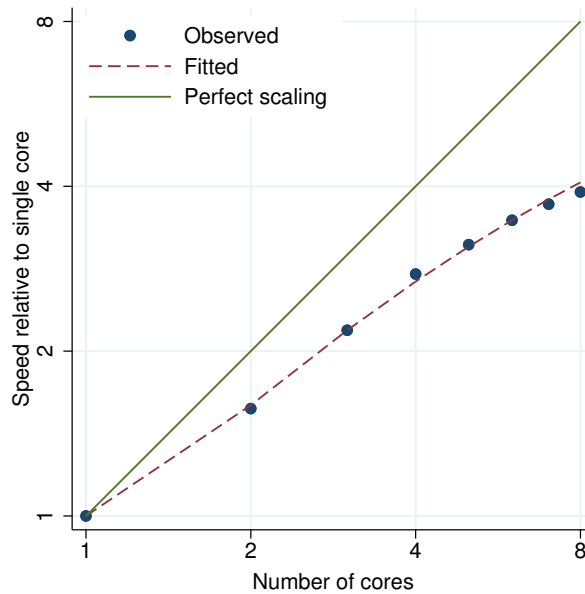
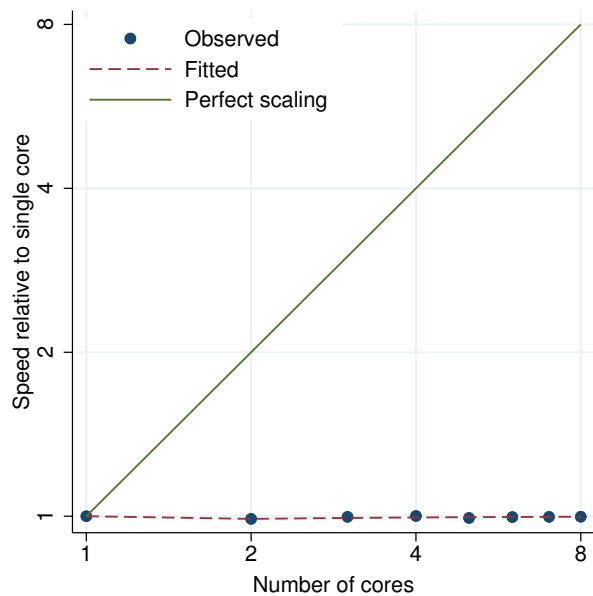
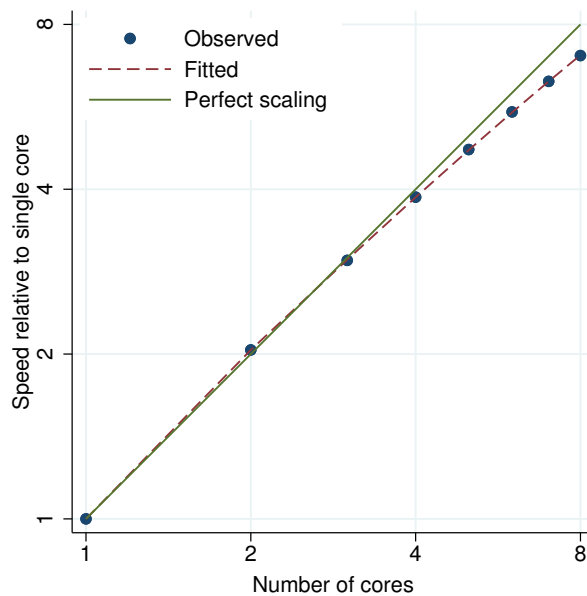
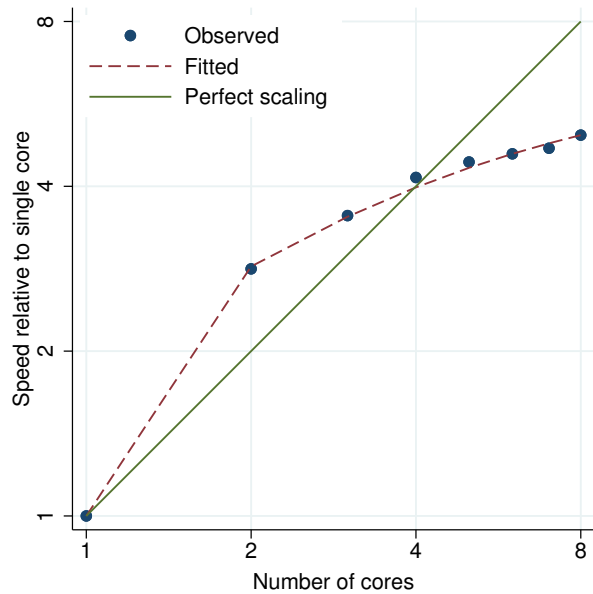
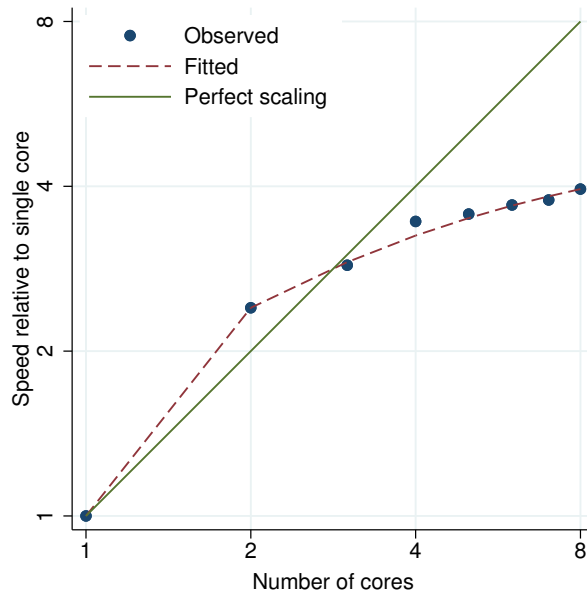


Figure 567. xthtaylor performance plot.

Figure 568. `xtile` performance plot.Figure 569. `xtintreg` performance plot.Figure 570. `xtivreg, be` performance plot.Figure 571. `xtivreg, fd` performance plot.



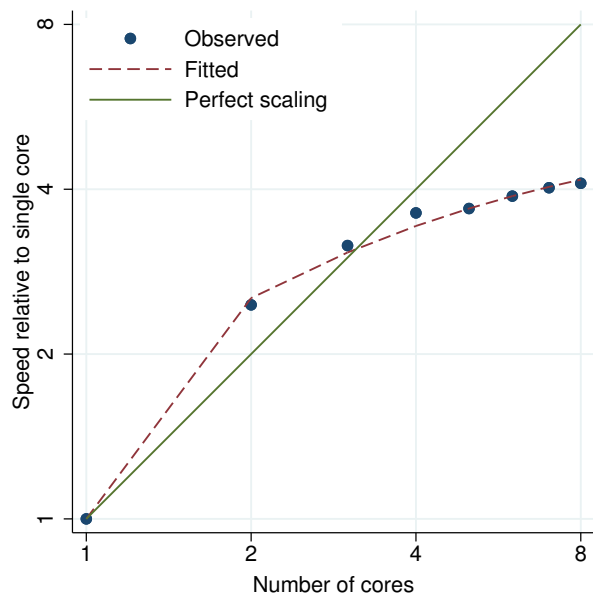


Figure 572. xtivreg, fe performance plot.

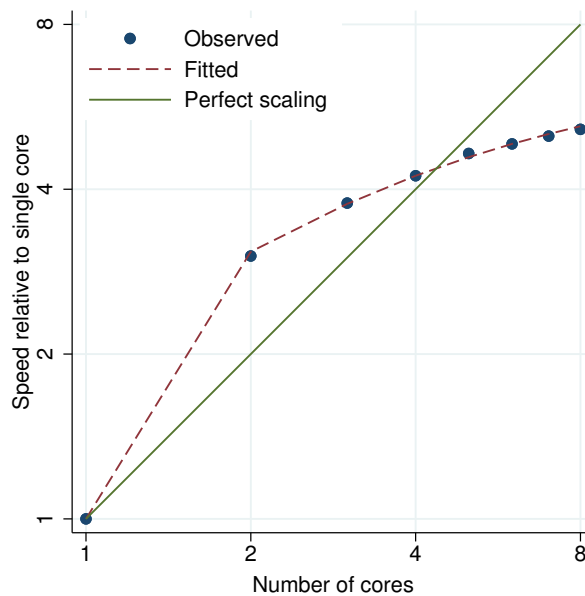


Figure 573. xtivreg, re performance plot.

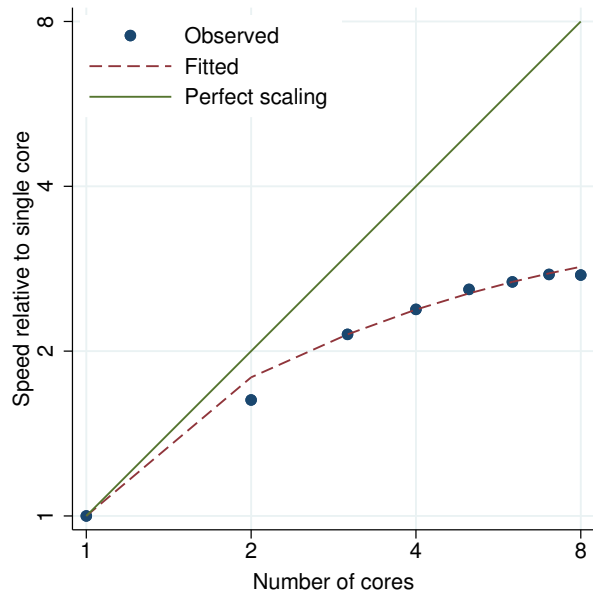


Figure 574. xtlogit, fe performance plot.

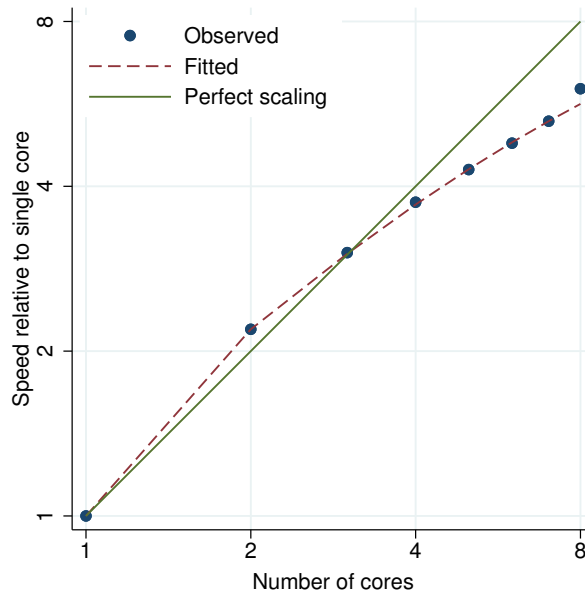


Figure 575. xtlogit, re performance plot.

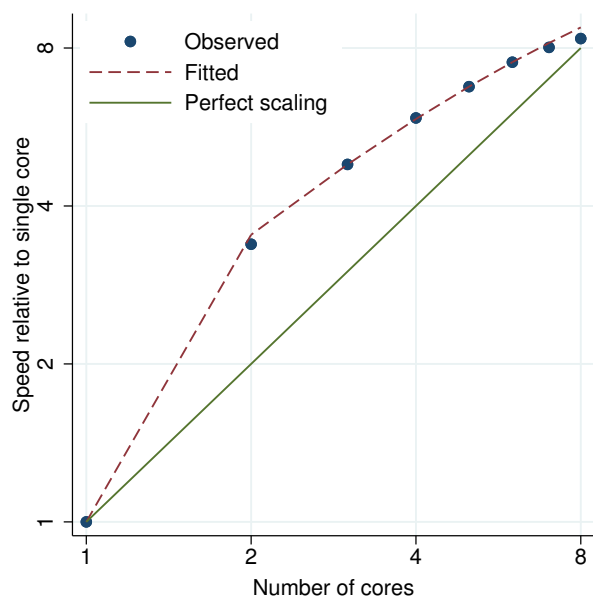


Figure 576. xtnbreg, fe performance plot.

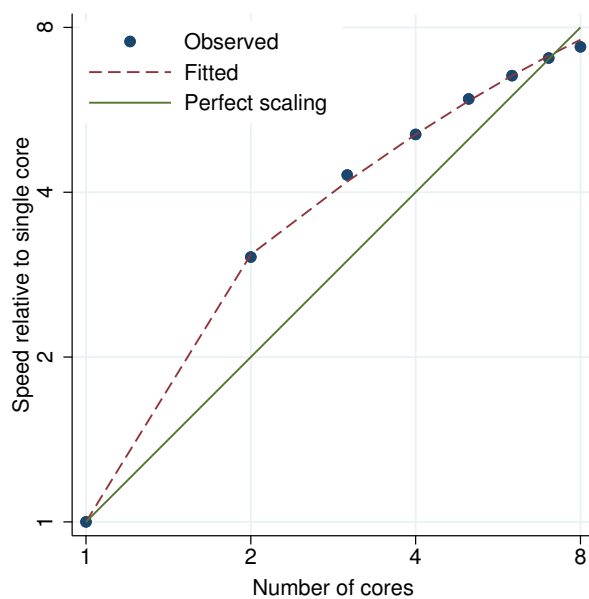


Figure 577. xtnbreg, re performance plot.

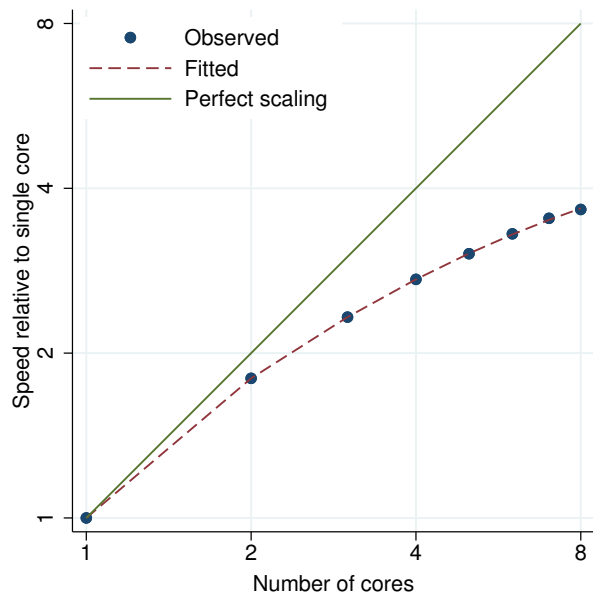


Figure 578. xtologit performance plot.

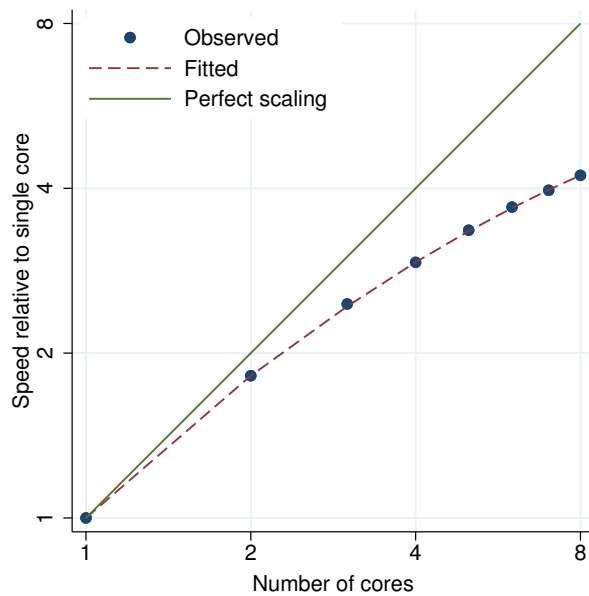


Figure 579. xtoprobit performance plot.

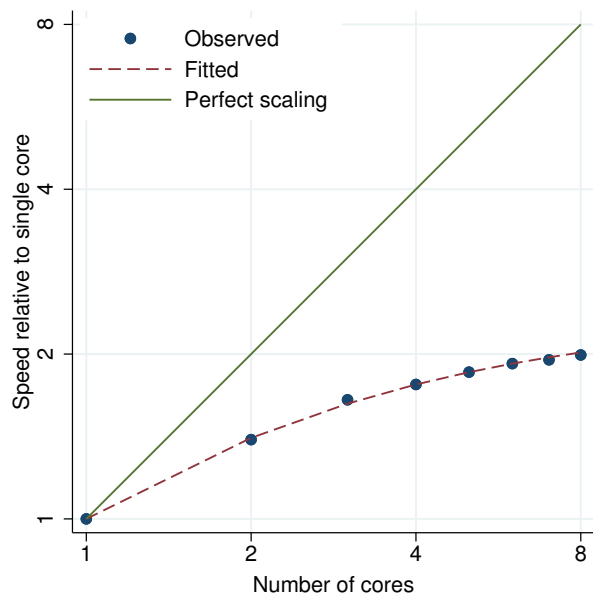


Figure 580. xtpcse performance plot.

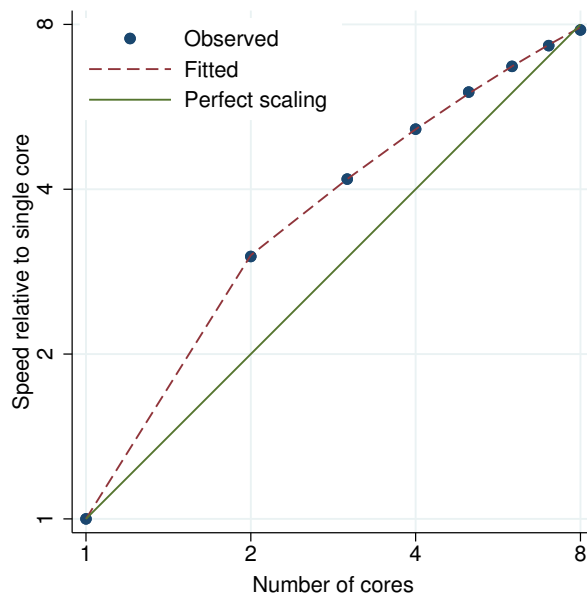


Figure 581. xtpoisson, fe performance plot.

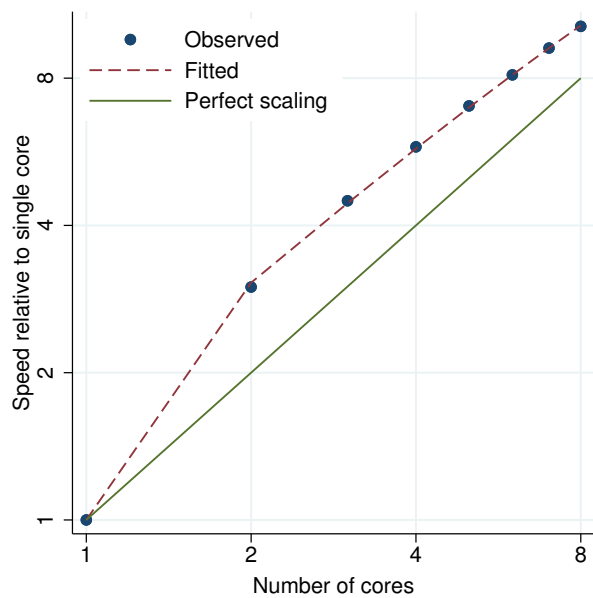


Figure 582. xtpoisson, re performance plot.

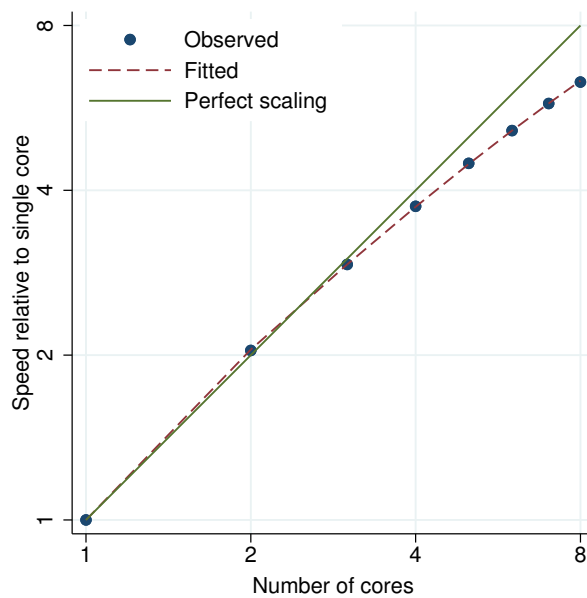


Figure 583. xtprobit, re performance plot.

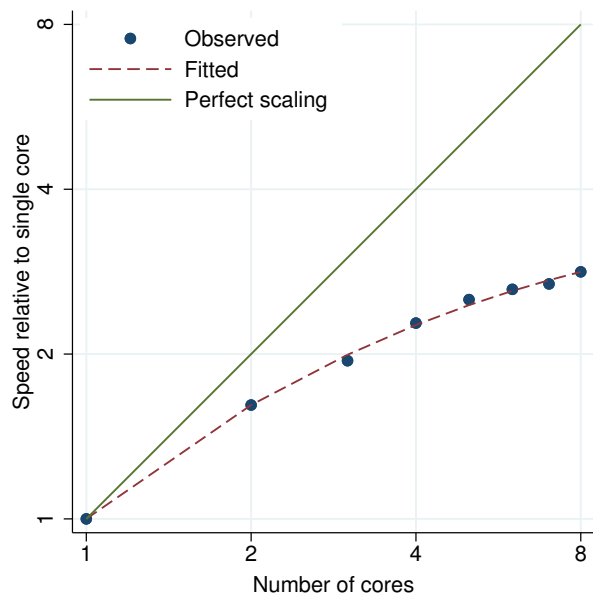


Figure 584. xtrc performance plot.

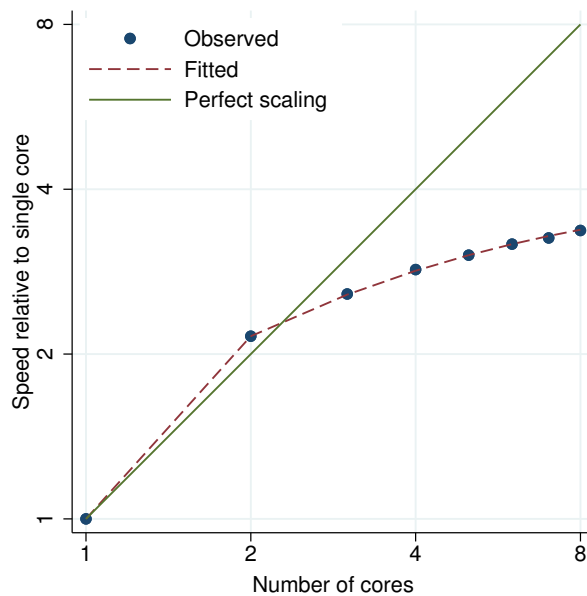


Figure 585. xtreg, be performance plot.

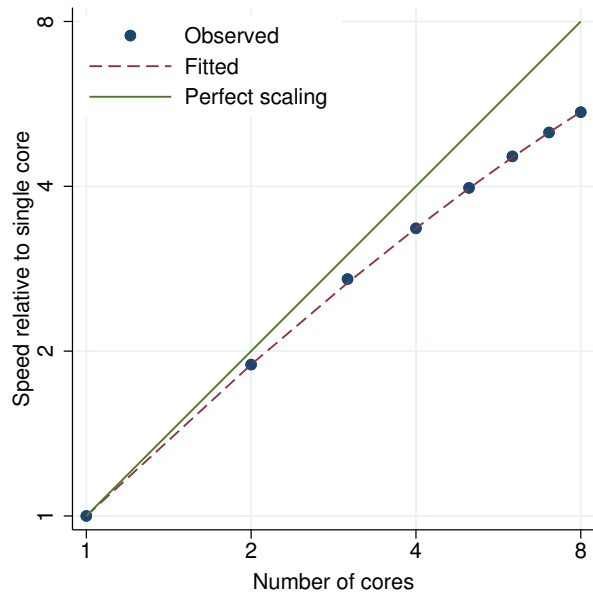


Figure 586. xtreg, fe performance plot.

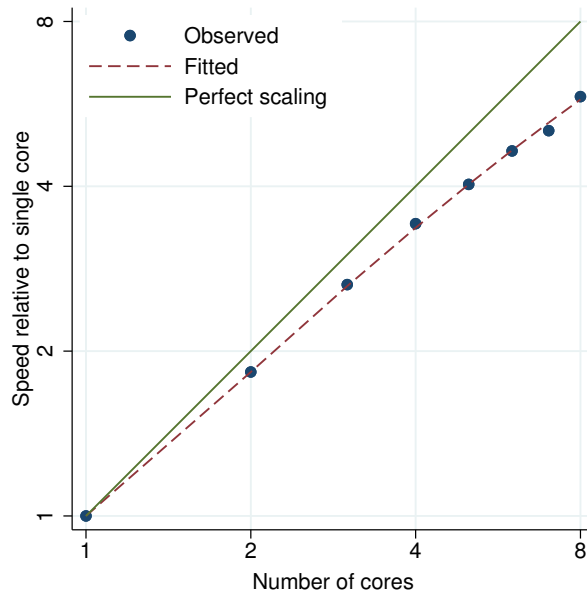
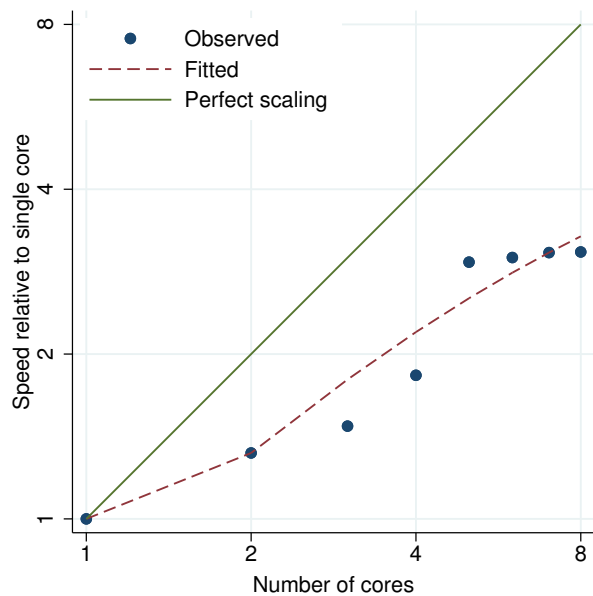
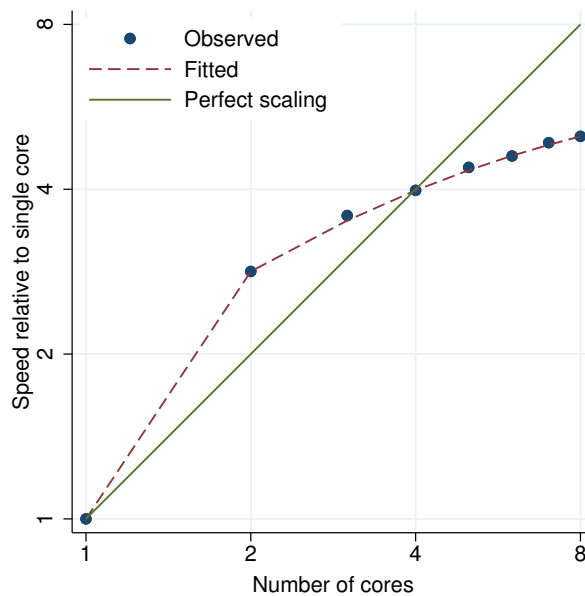
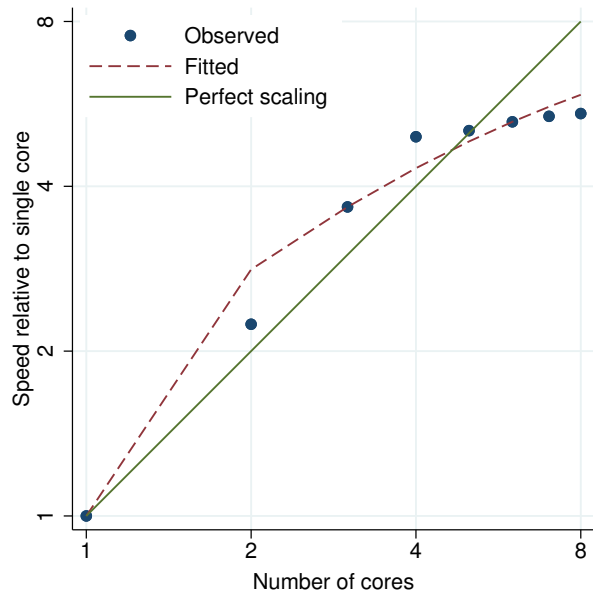
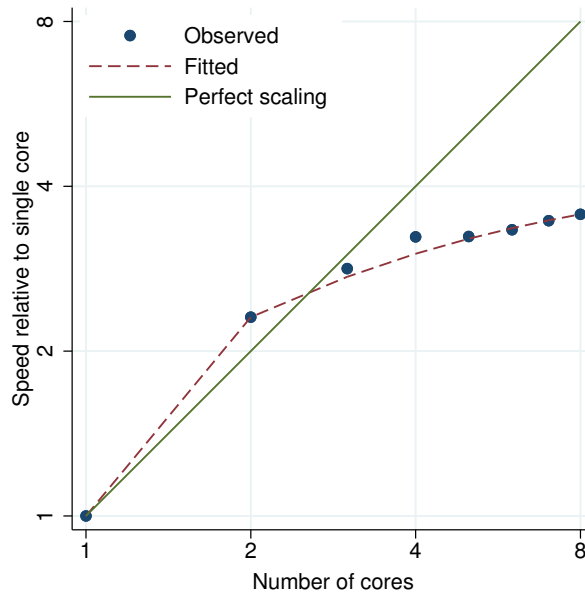
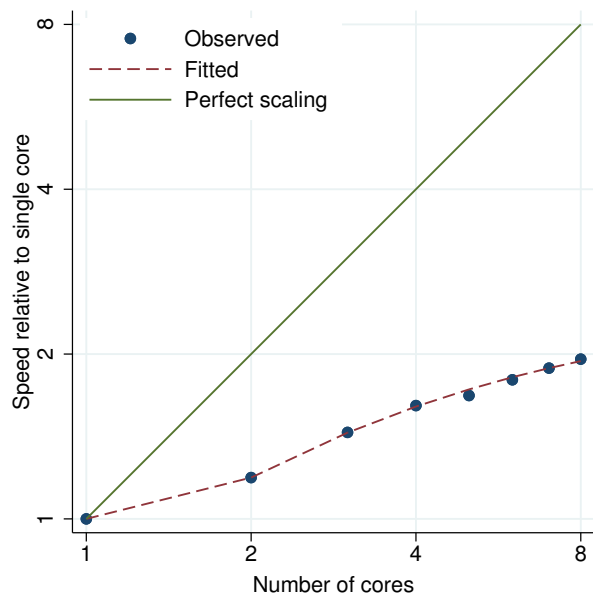
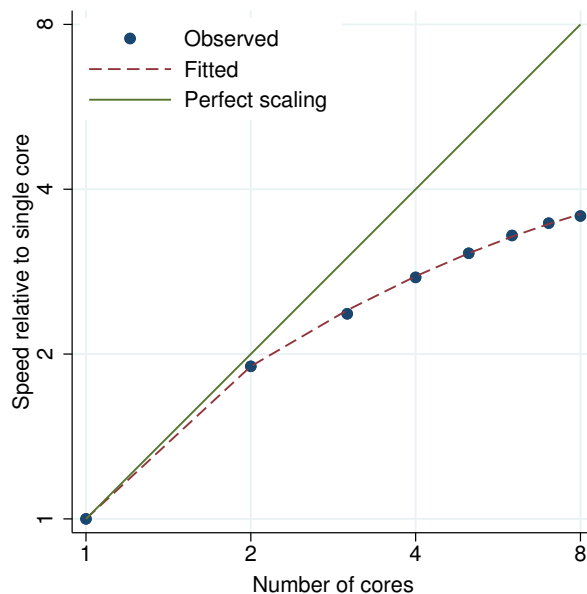
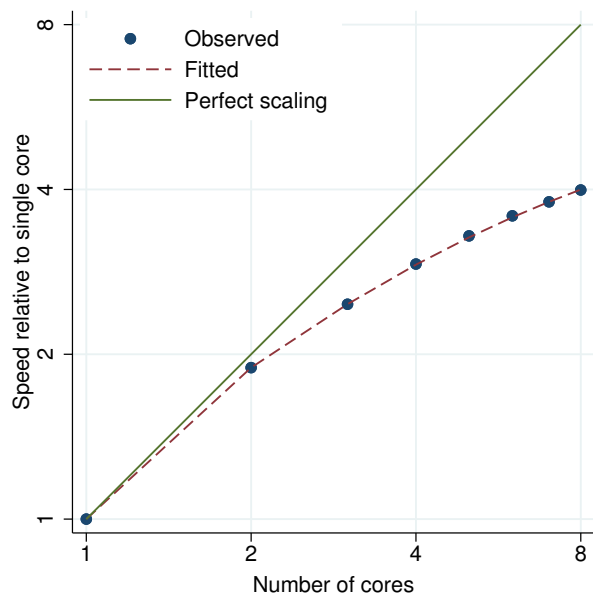
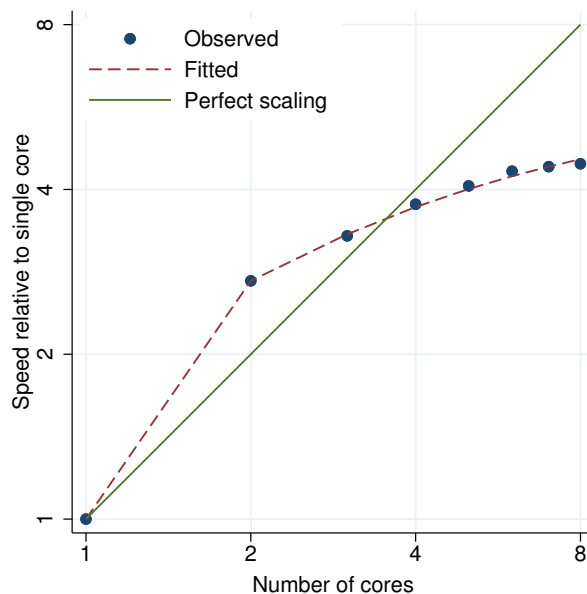


Figure 587. xtreg, fe vce(robust) performance plot.

Figure 588. `xtreg, mle` performance plot.Figure 589. `xtreg, re` performance plot.Figure 590. `xtregar, fe` performance plot.Figure 591. `xtregar, re` performance plot.

Figure 592. `xtset` performance plot.Figure 593. `xtstreg`,  
`distribution(exponential)` performance plot.Figure 594. `xtstreg`,  
`distribution(weibull)` performance plot.Figure 595. `xtsum` performance plot.

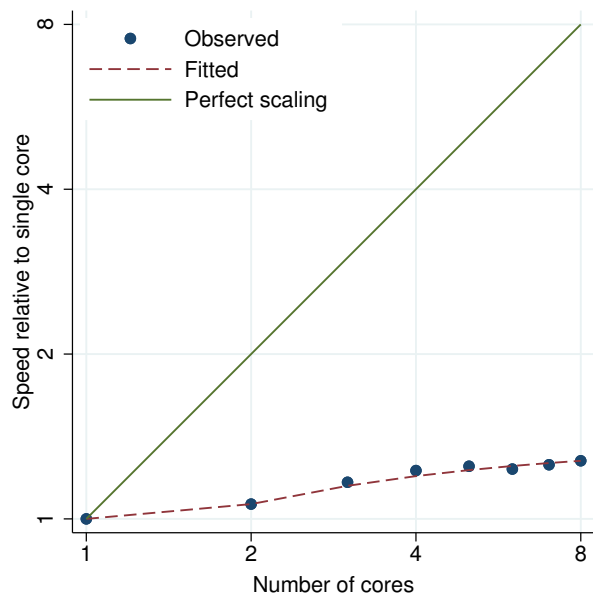


Figure 596. xttab performance plot.

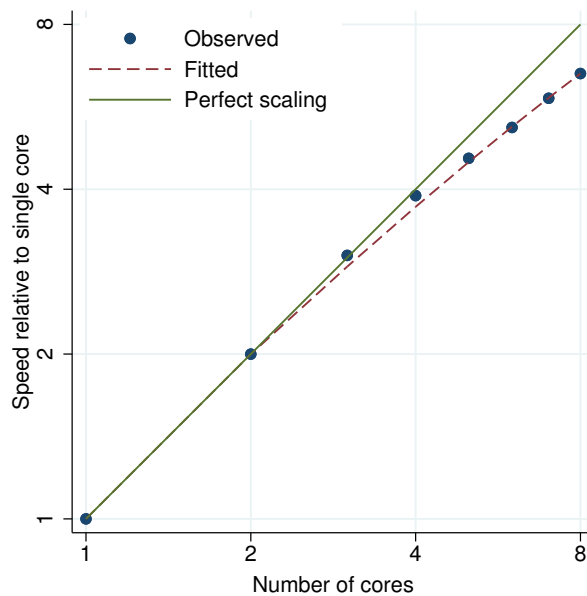


Figure 597. xttobit performance plot.

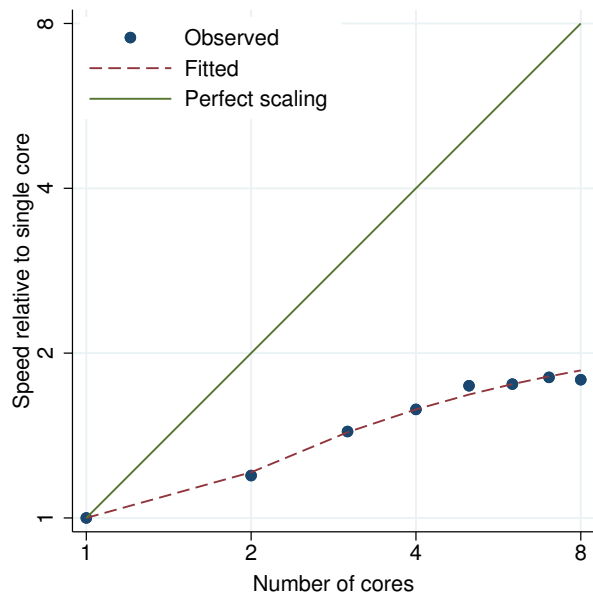


Figure 598. xtunitroot breitung performance plot.

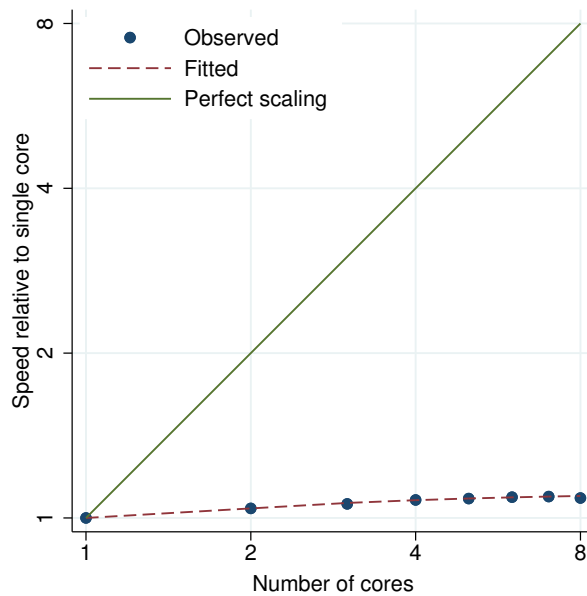
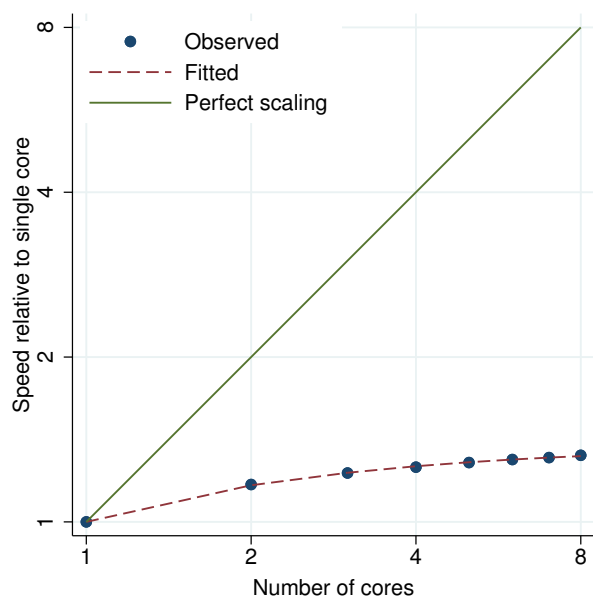
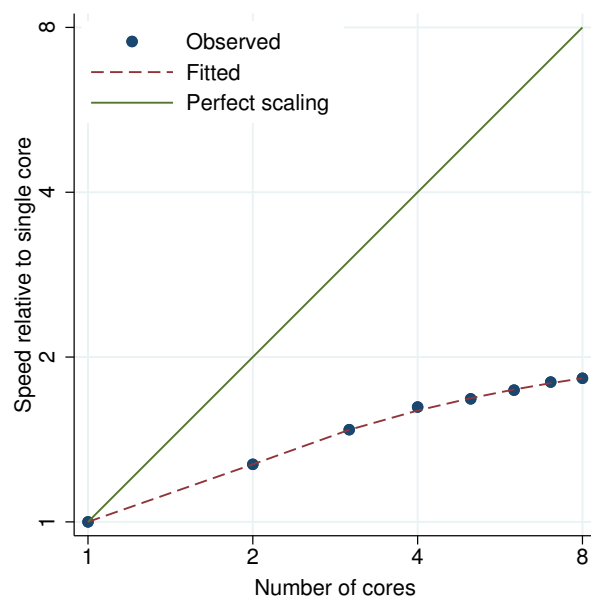
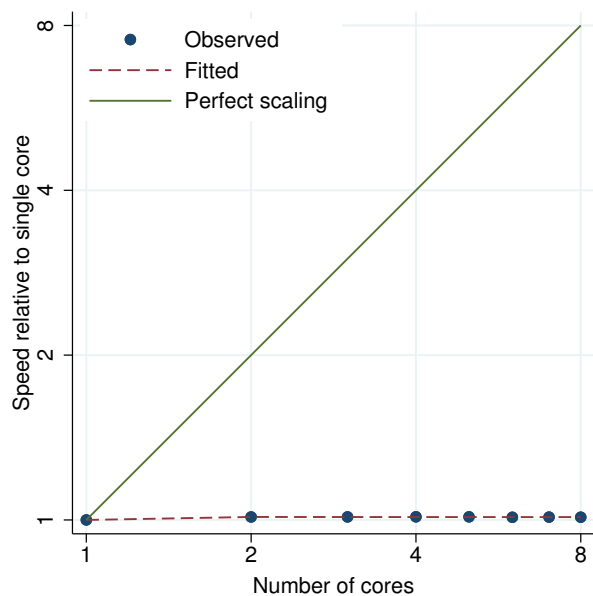
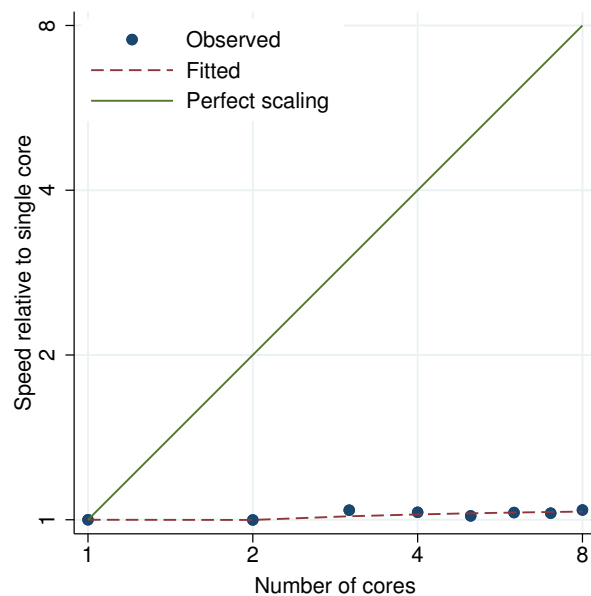


Figure 599. xtunitroot fisher performance plot.

Figure 600. `xtunitroot hadri` performance plot.Figure 601. `xtunitroot ht` performance plot.Figure 602. `xtunitroot ips` performance plot.Figure 603. `xtunitroot llc` performance plot.



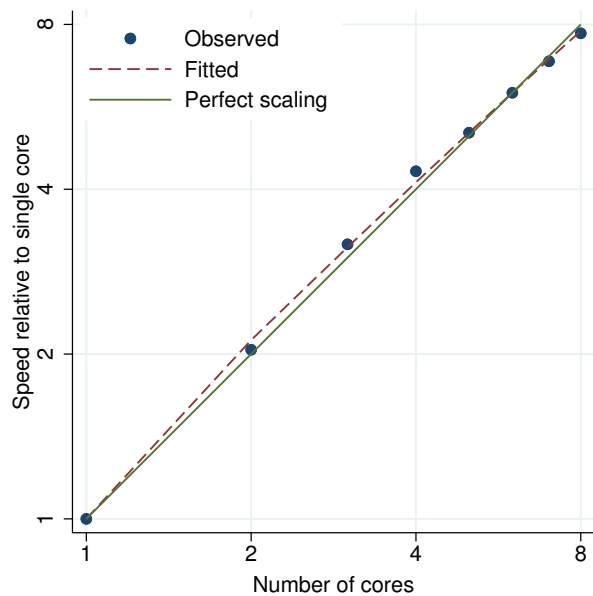


Figure 604. zlnb performance plot.

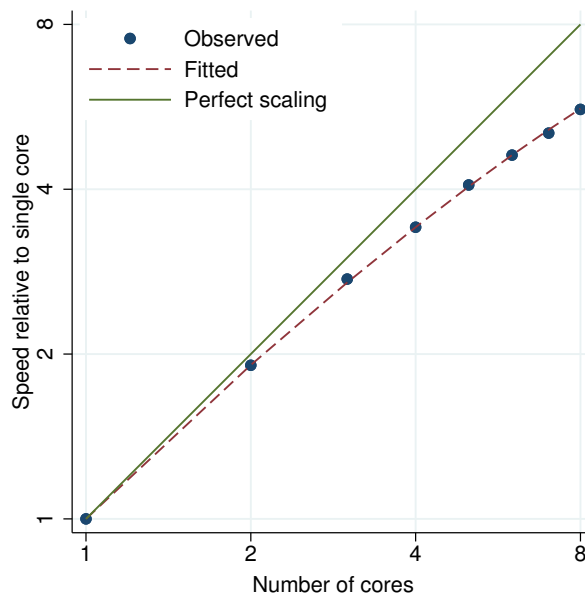


Figure 605. zioprobit performance plot.

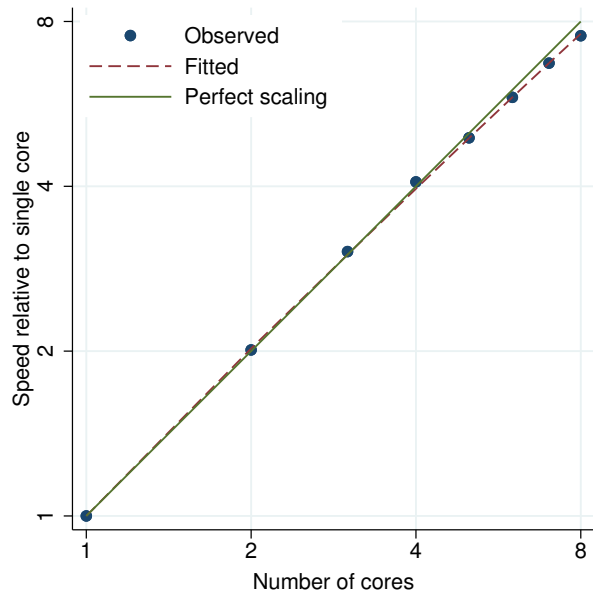


Figure 606. zip performance plot.

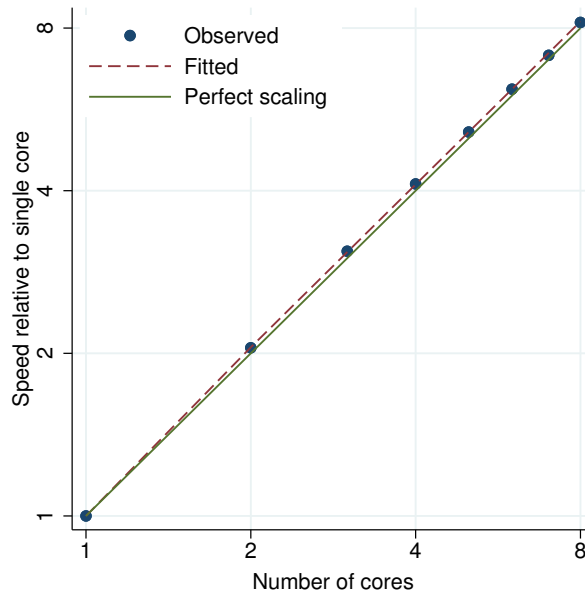
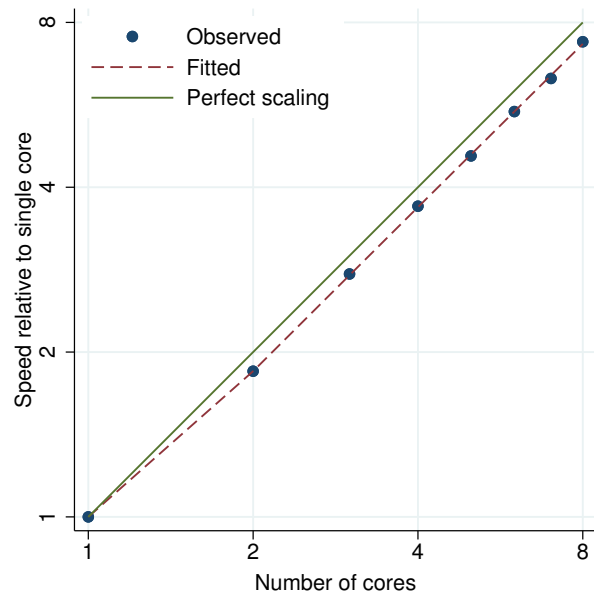
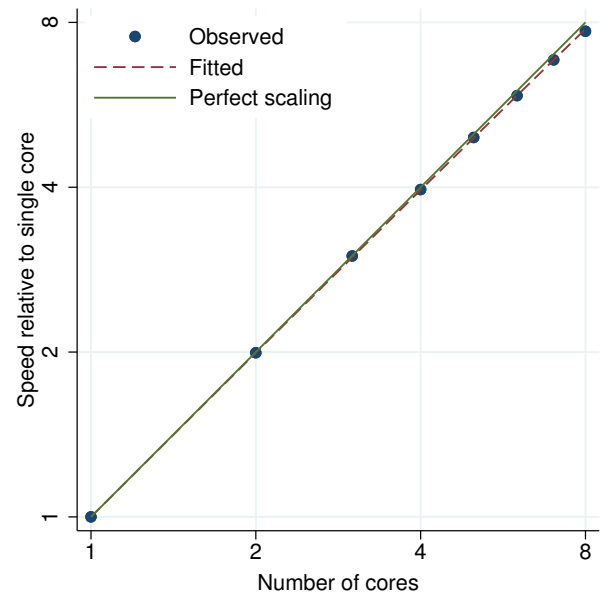


Figure 607. \_predict, xb performance plot.

Figure 608. `_rmcoll` performance plot.Figure 609. `_robust` performance plot.

## B Performance assessment graphs for high-end servers

Performance graphs of all 590 commands running on high-end servers are presented below.

These graphs are similar to the graphs from appendix [A](#) except that here the speeds are evaluated up to 40 cores.

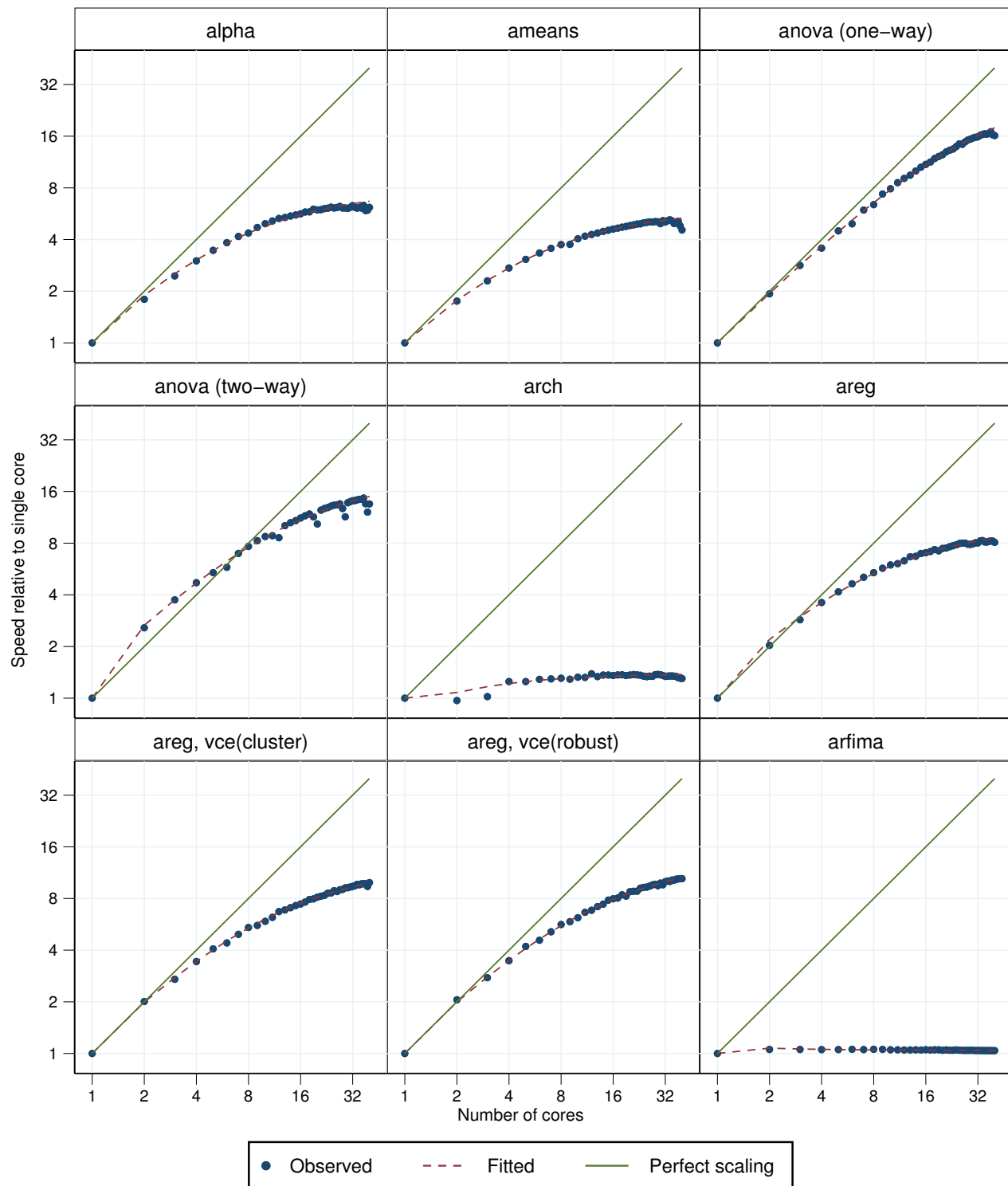


Figure 610. Parallelization performance plots.

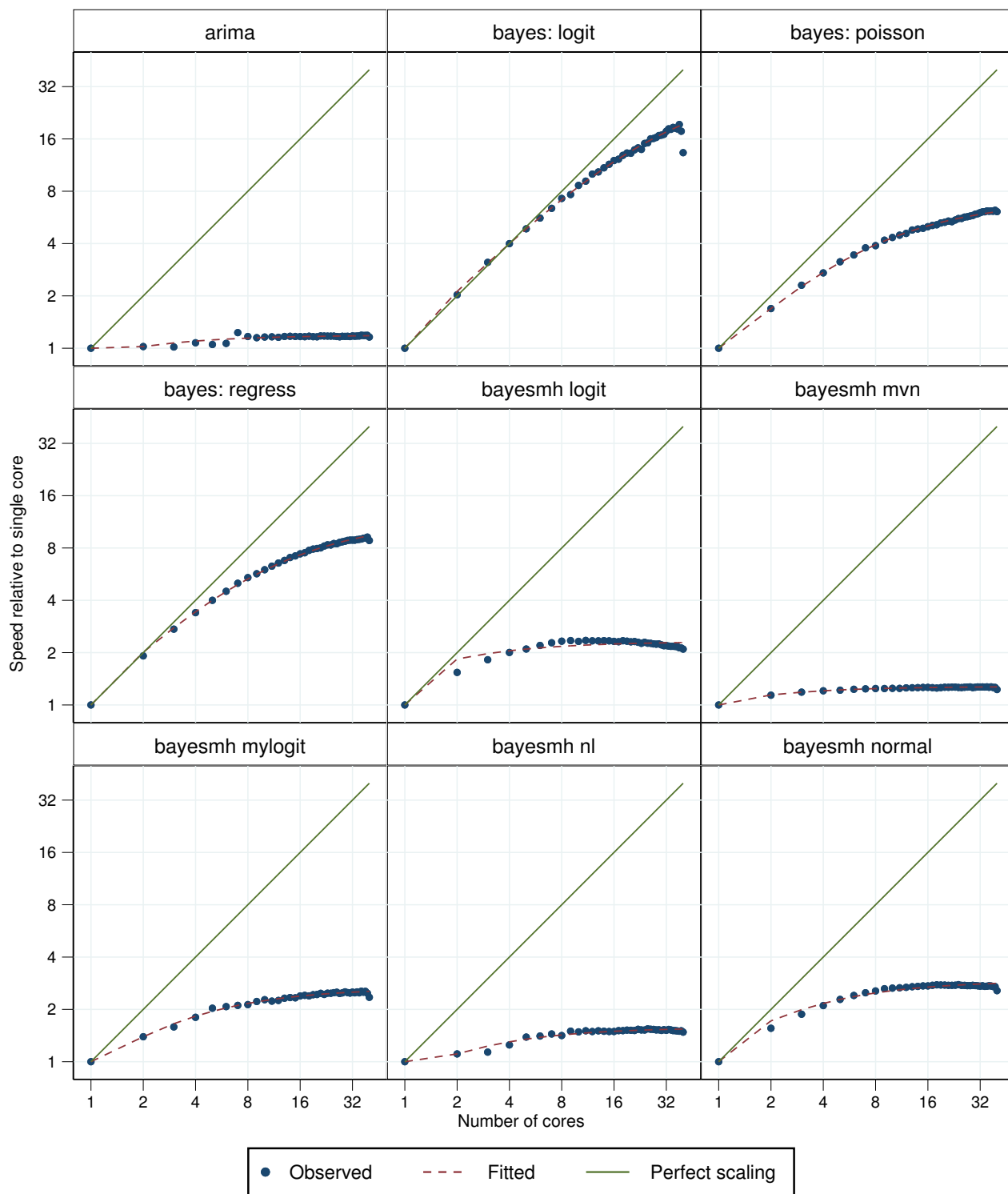


Figure 611. Parallelization performance plots.

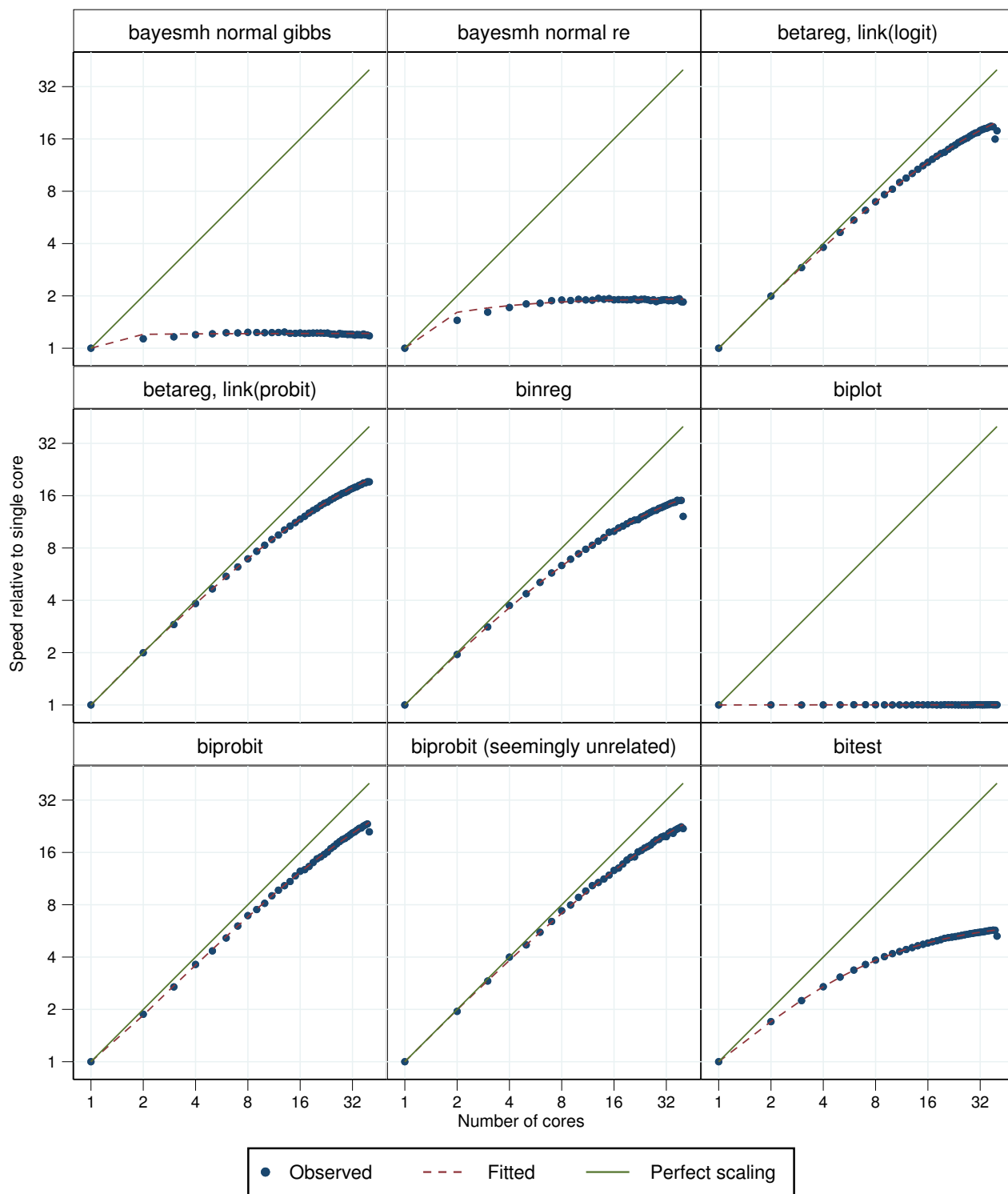


Figure 612. Parallelization performance plots.

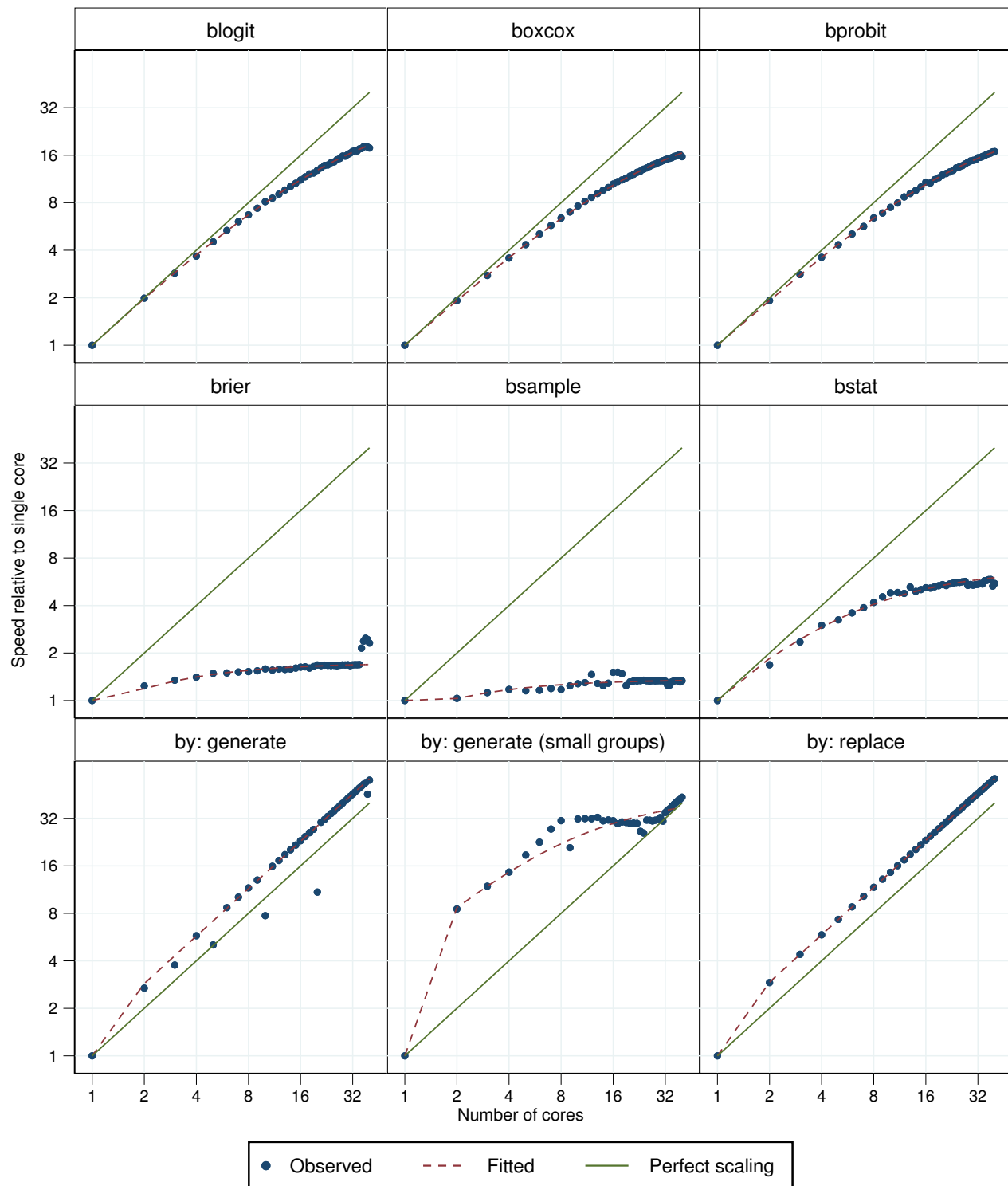


Figure 613. Parallelization performance plots.

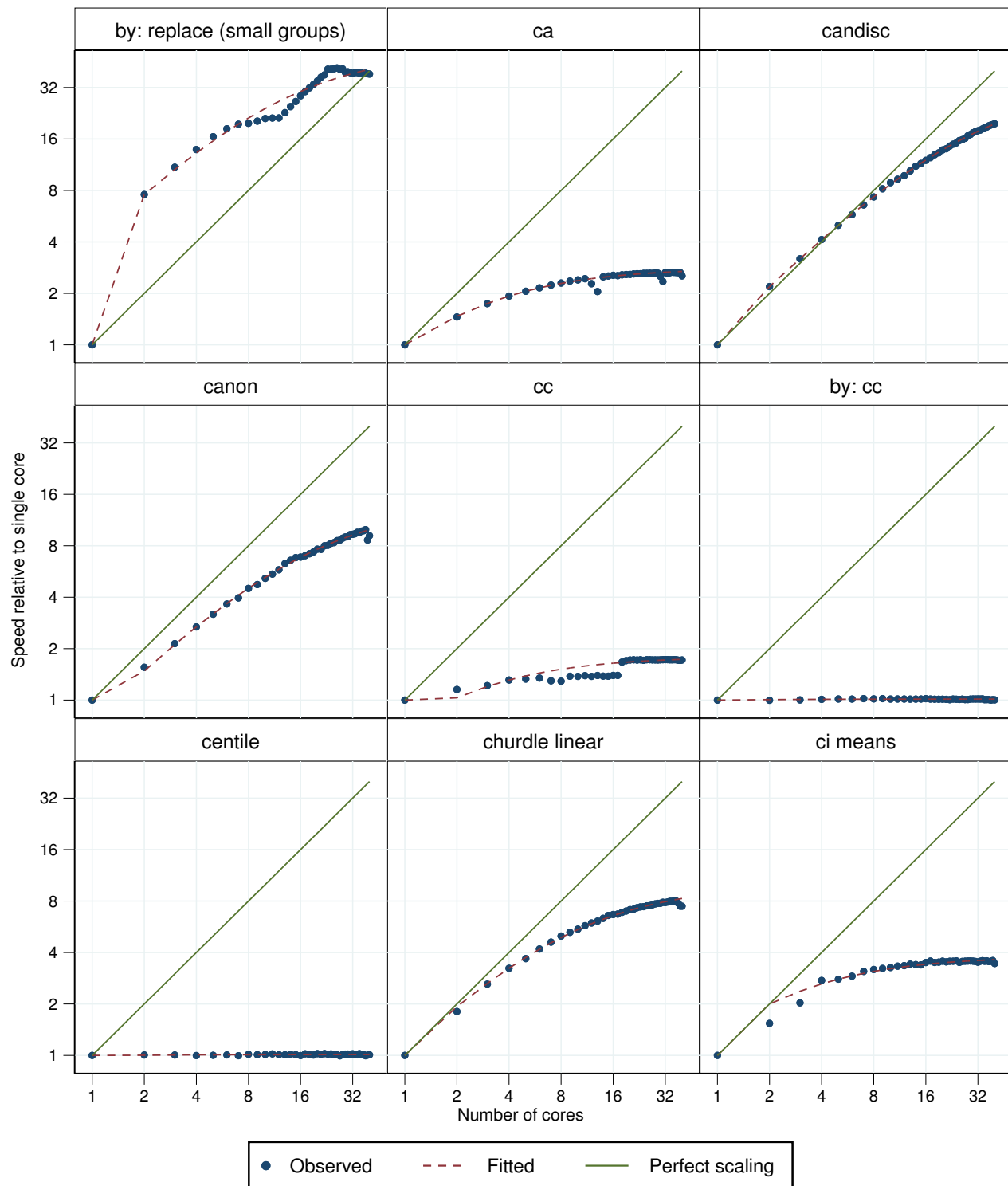


Figure 614. Parallelization performance plots.



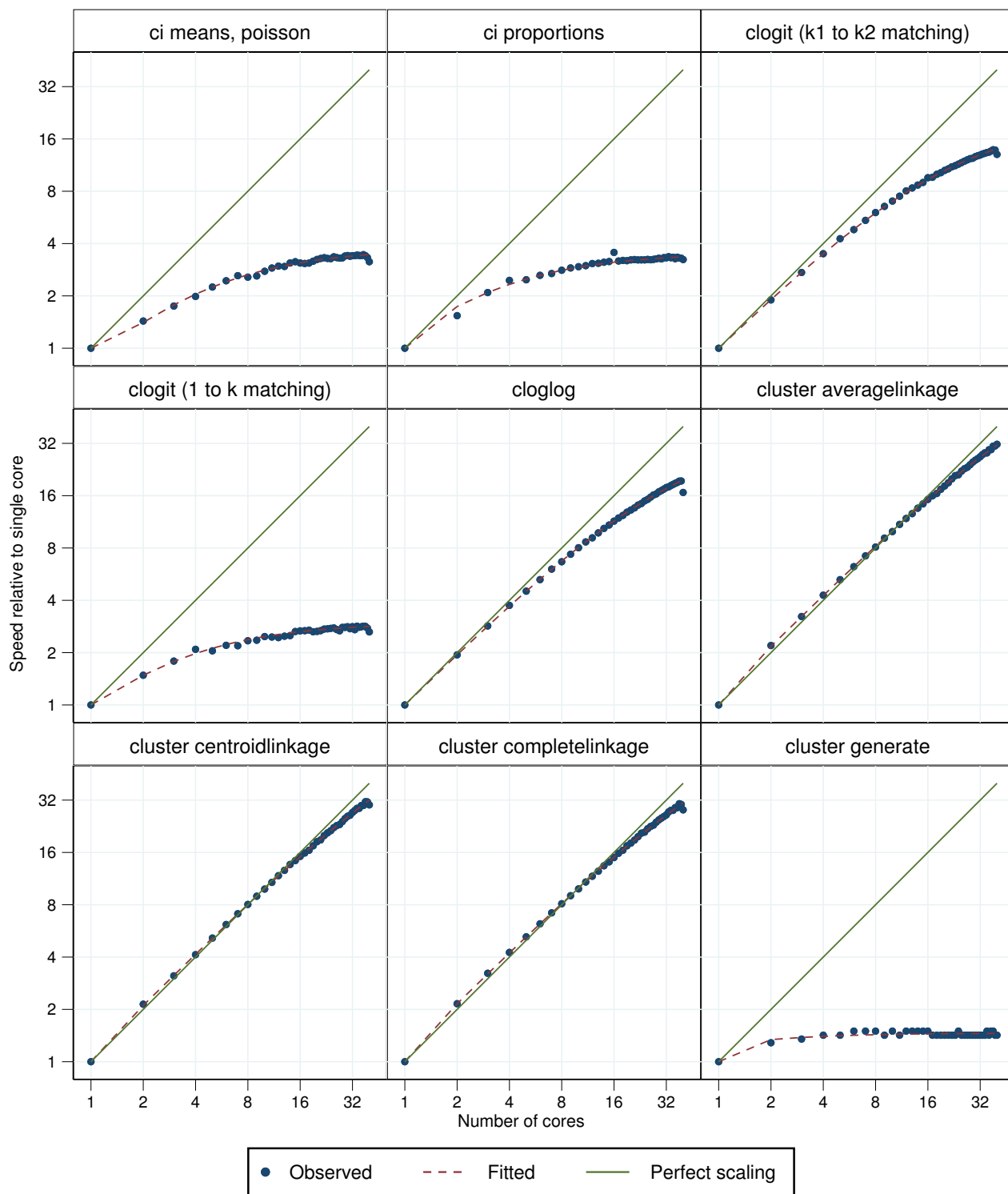


Figure 615. Parallelization performance plots.

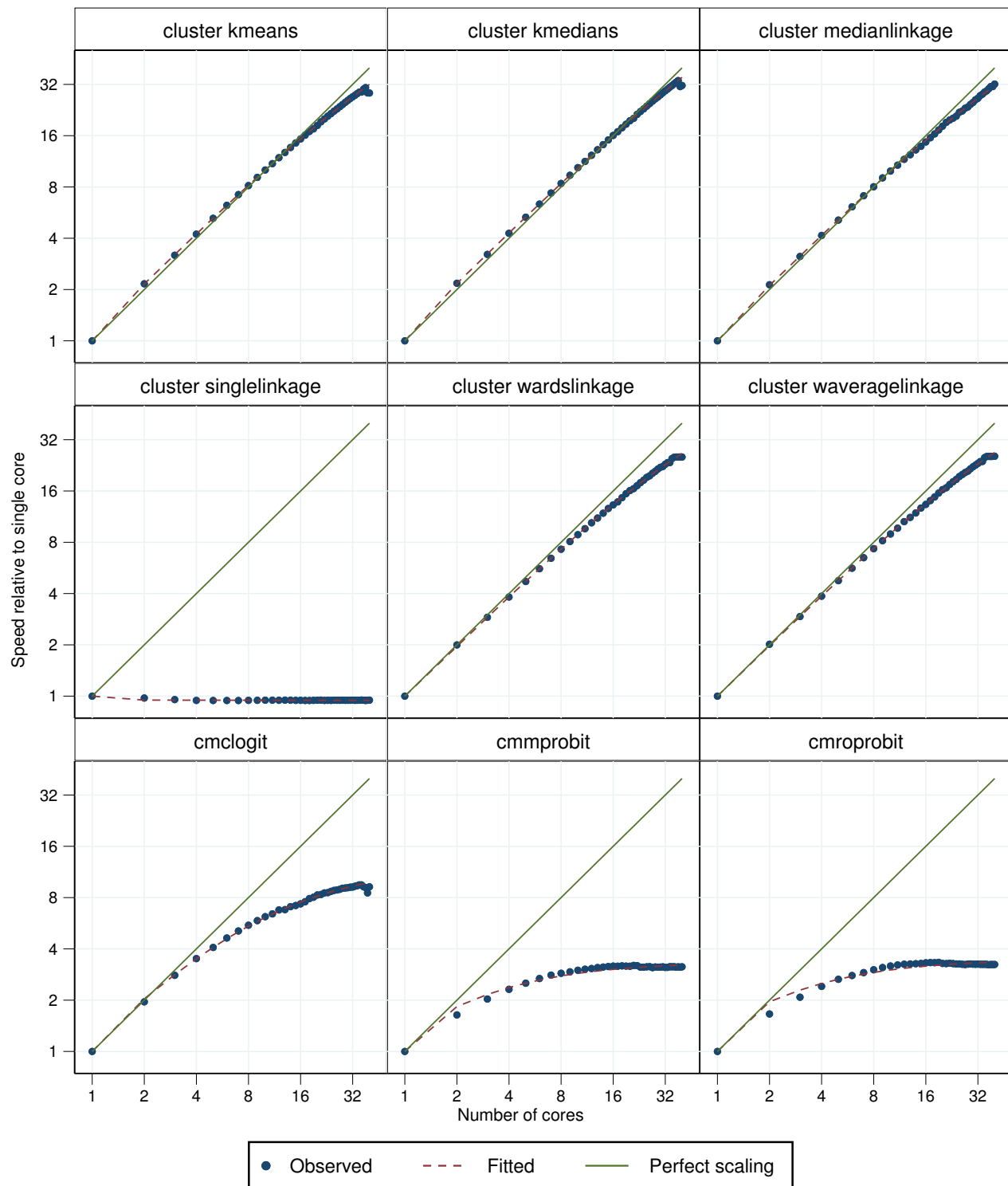


Figure 616. Parallelization performance plots.

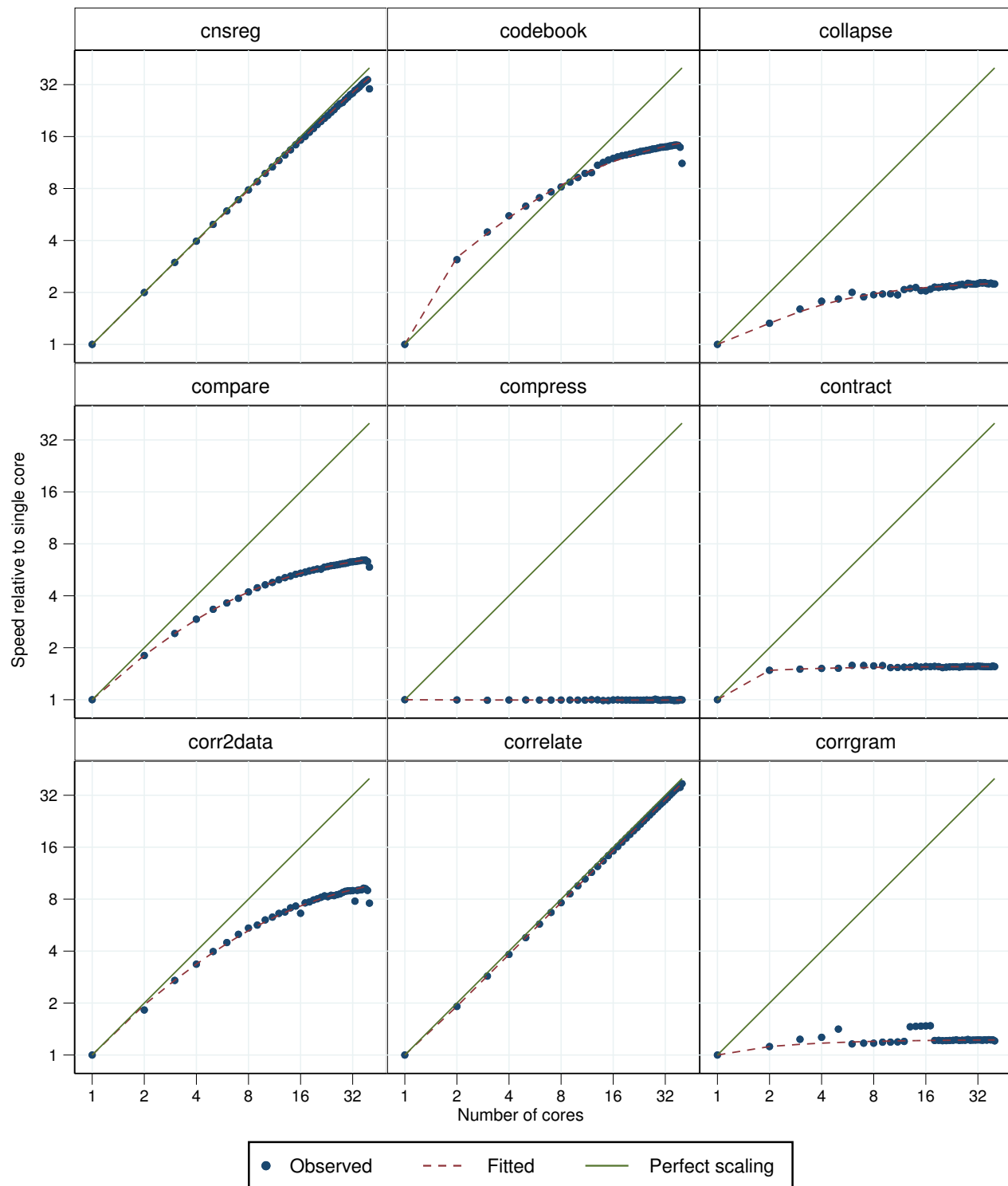


Figure 617. Parallelization performance plots.

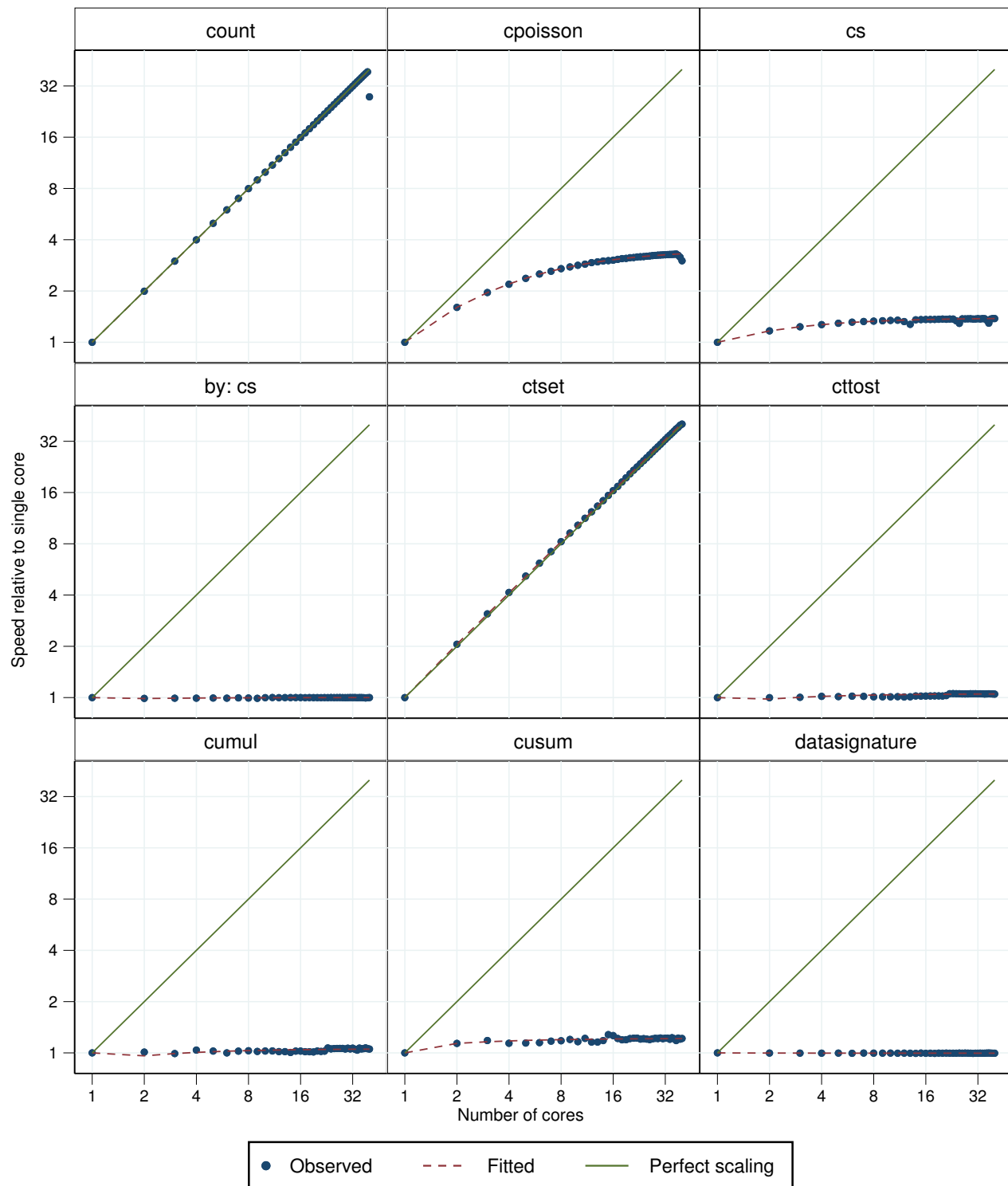


Figure 618. Parallelization performance plots.

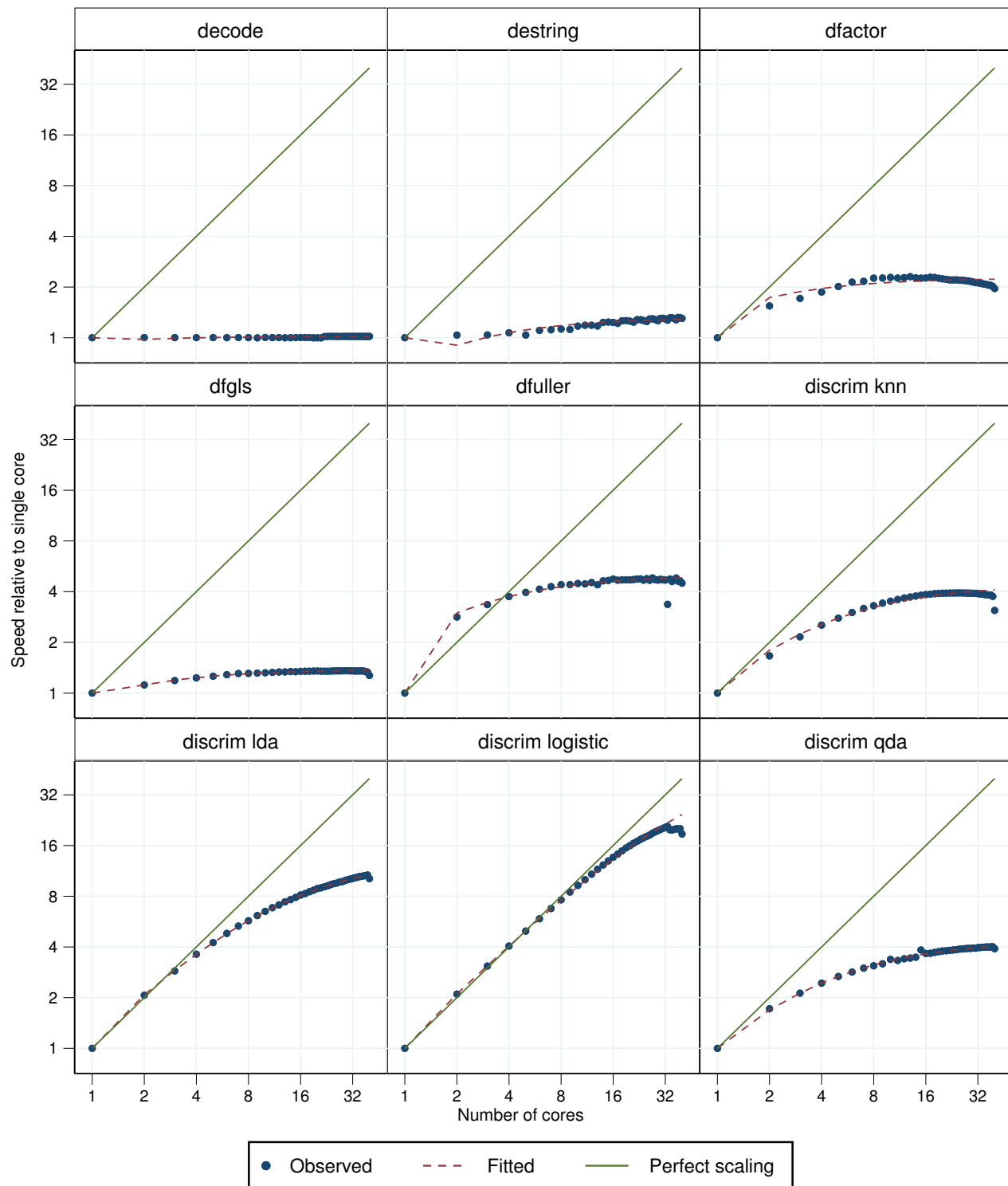


Figure 619. Parallelization performance plots.

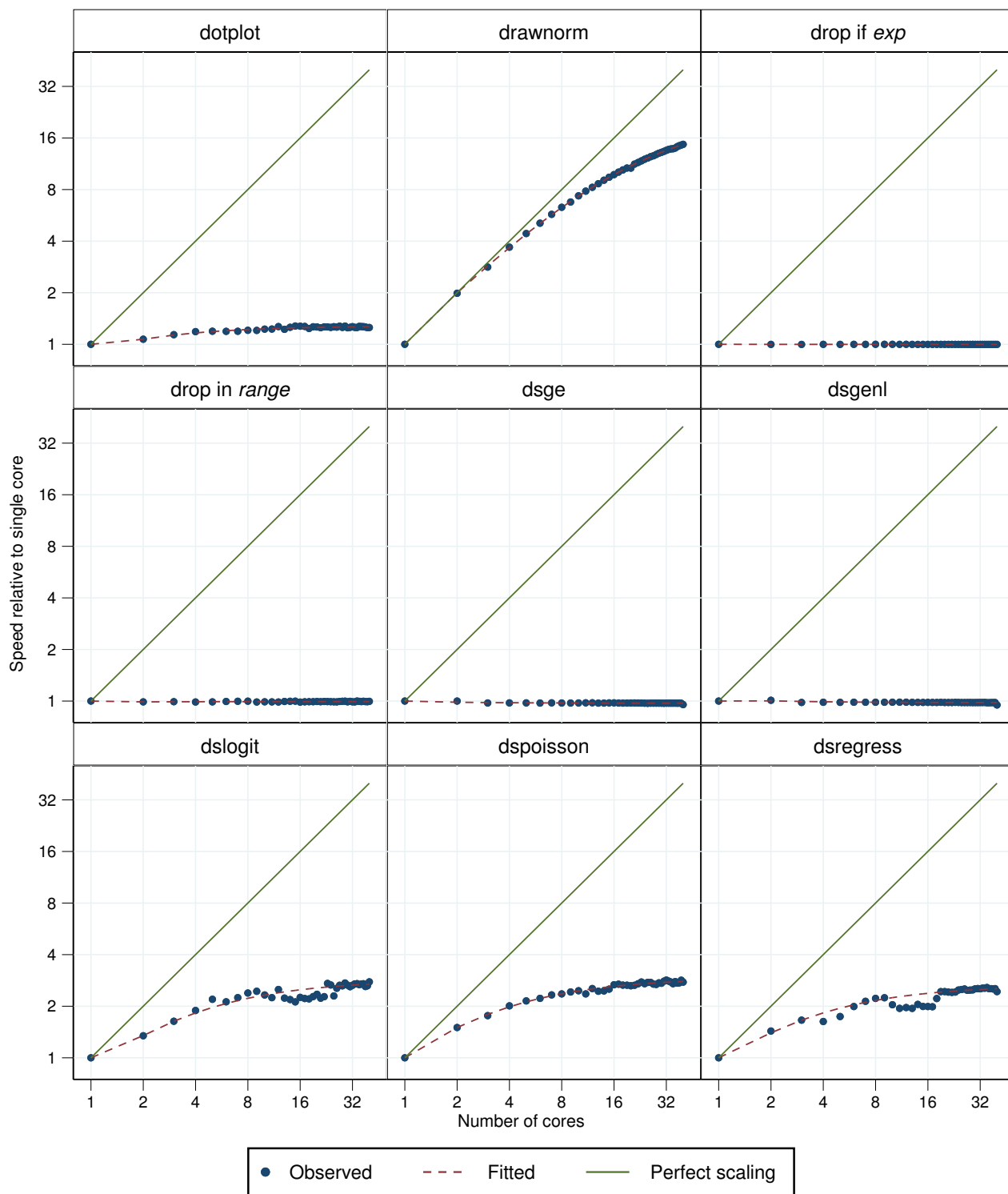


Figure 620. Parallelization performance plots.

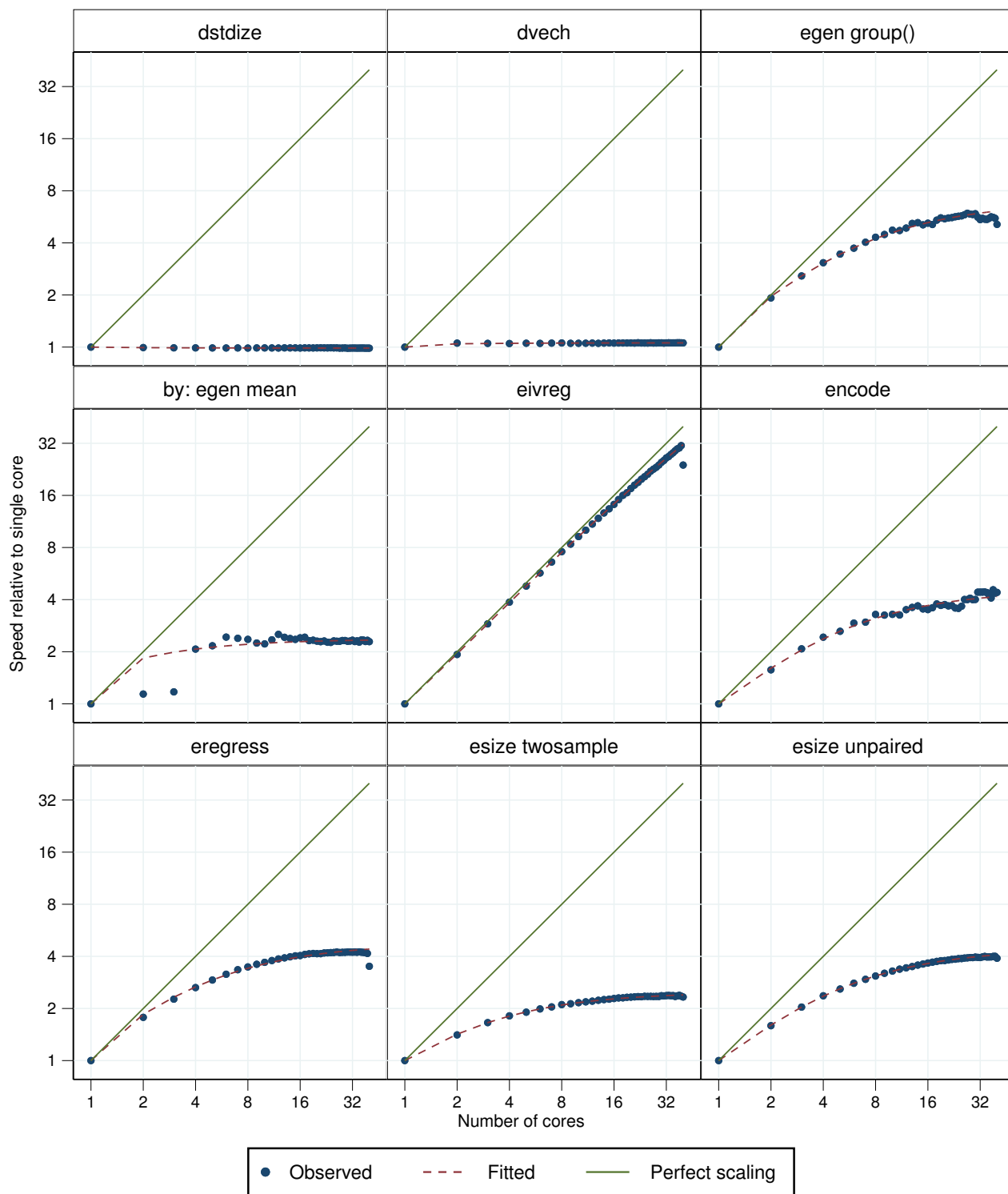


Figure 621. Parallelization performance plots.

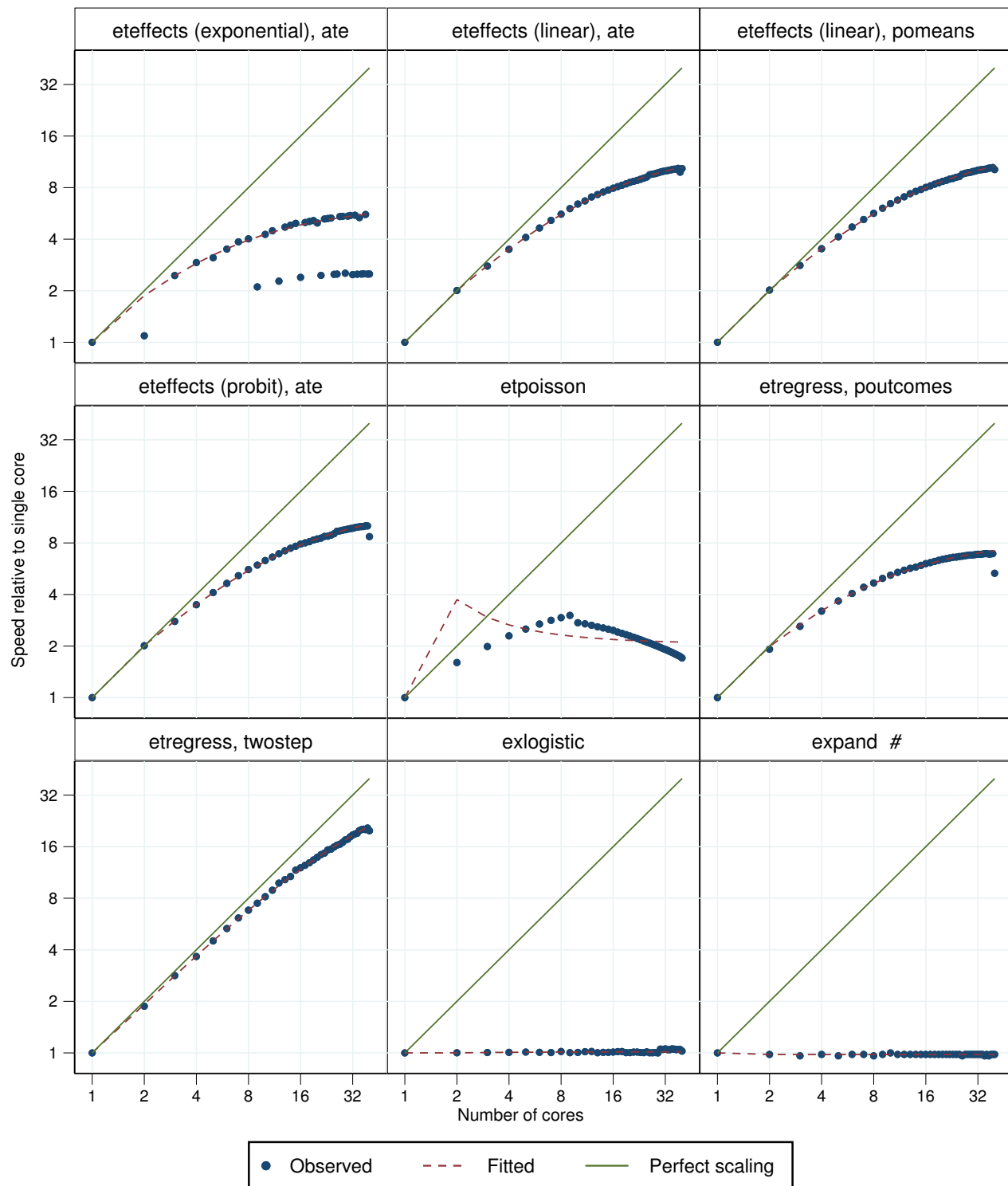


Figure 622. Parallelization performance plots.



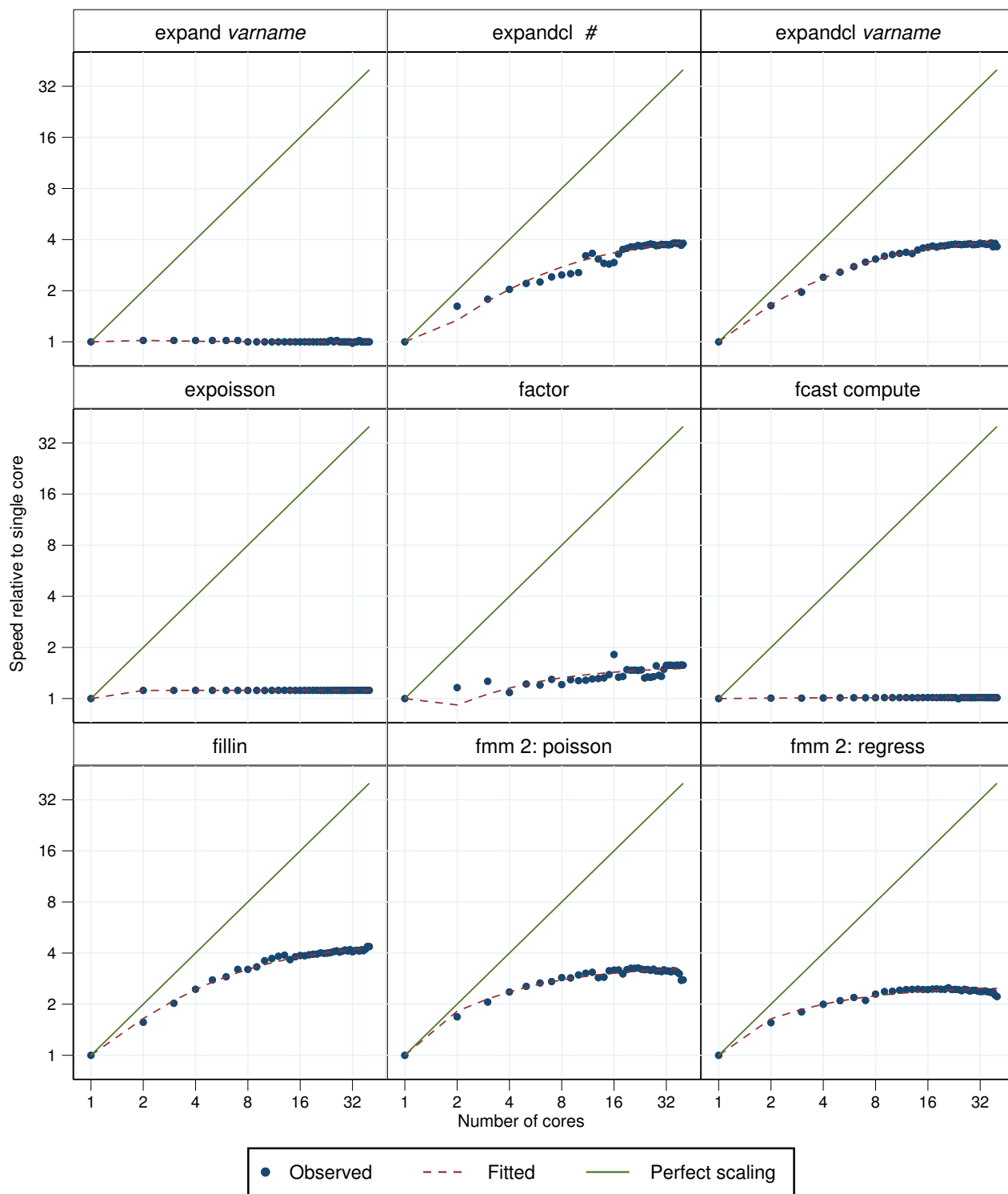


Figure 623. Parallelization performance plots.

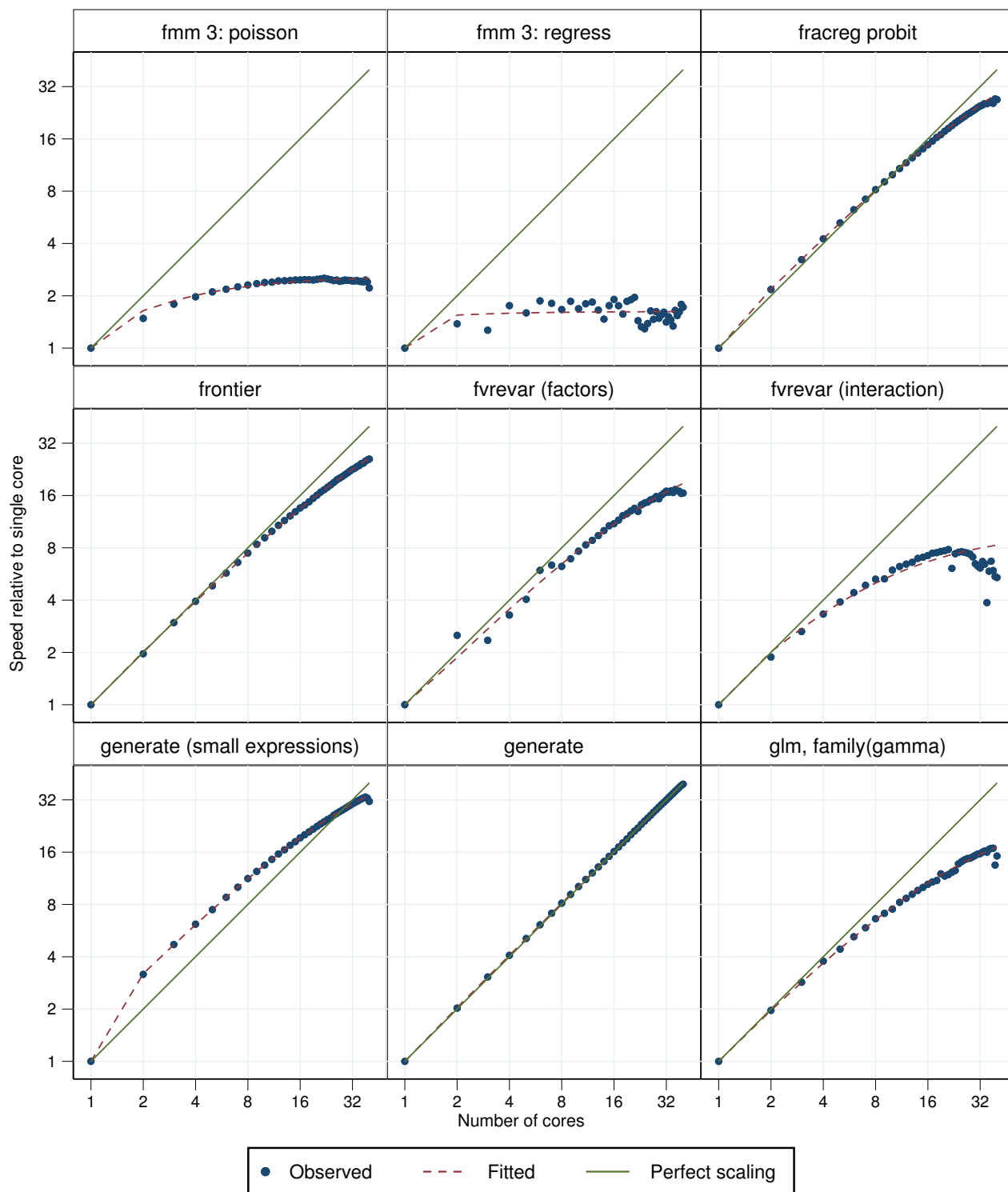


Figure 624. Parallelization performance plots.

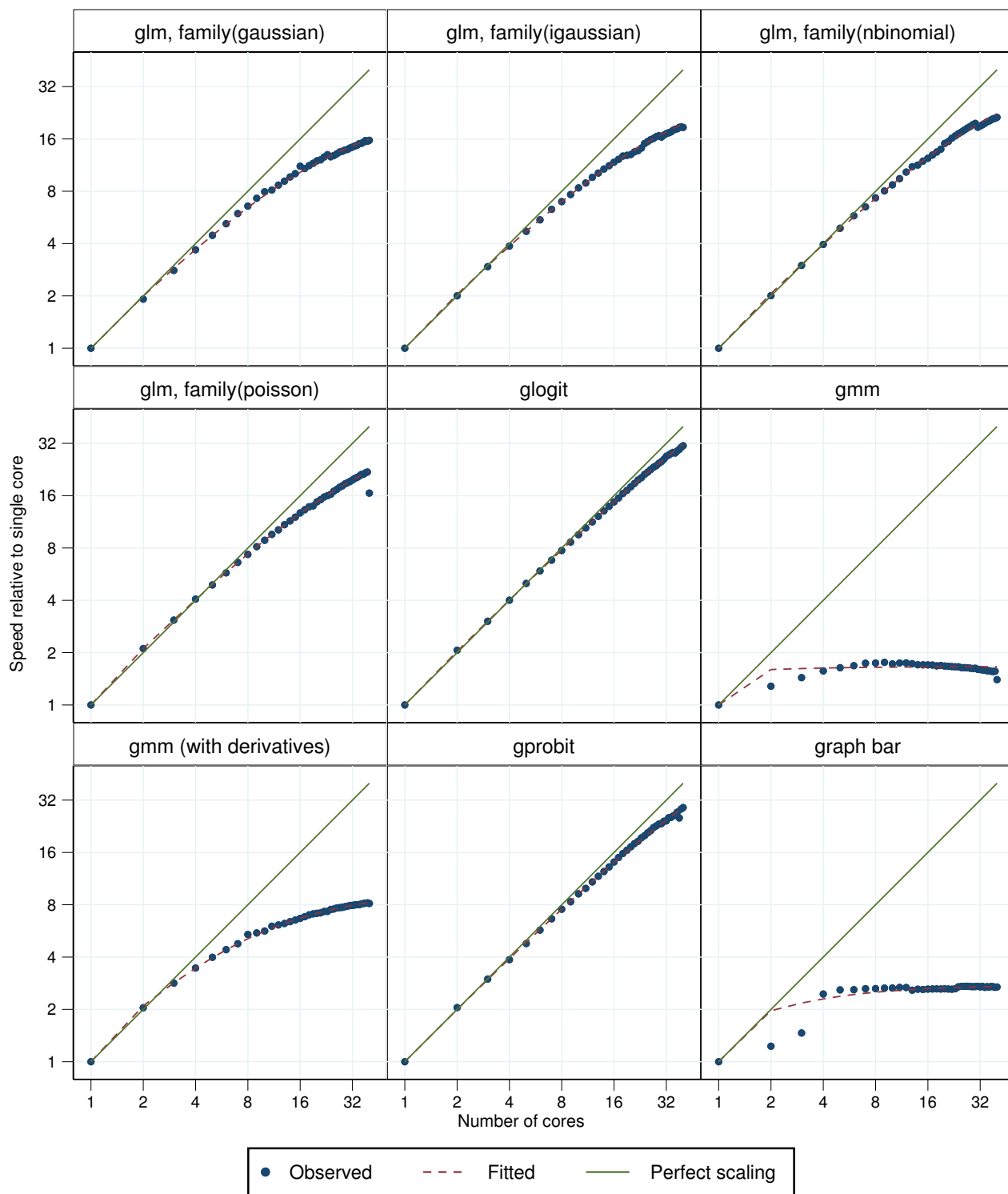


Figure 625. Parallelization performance plots.

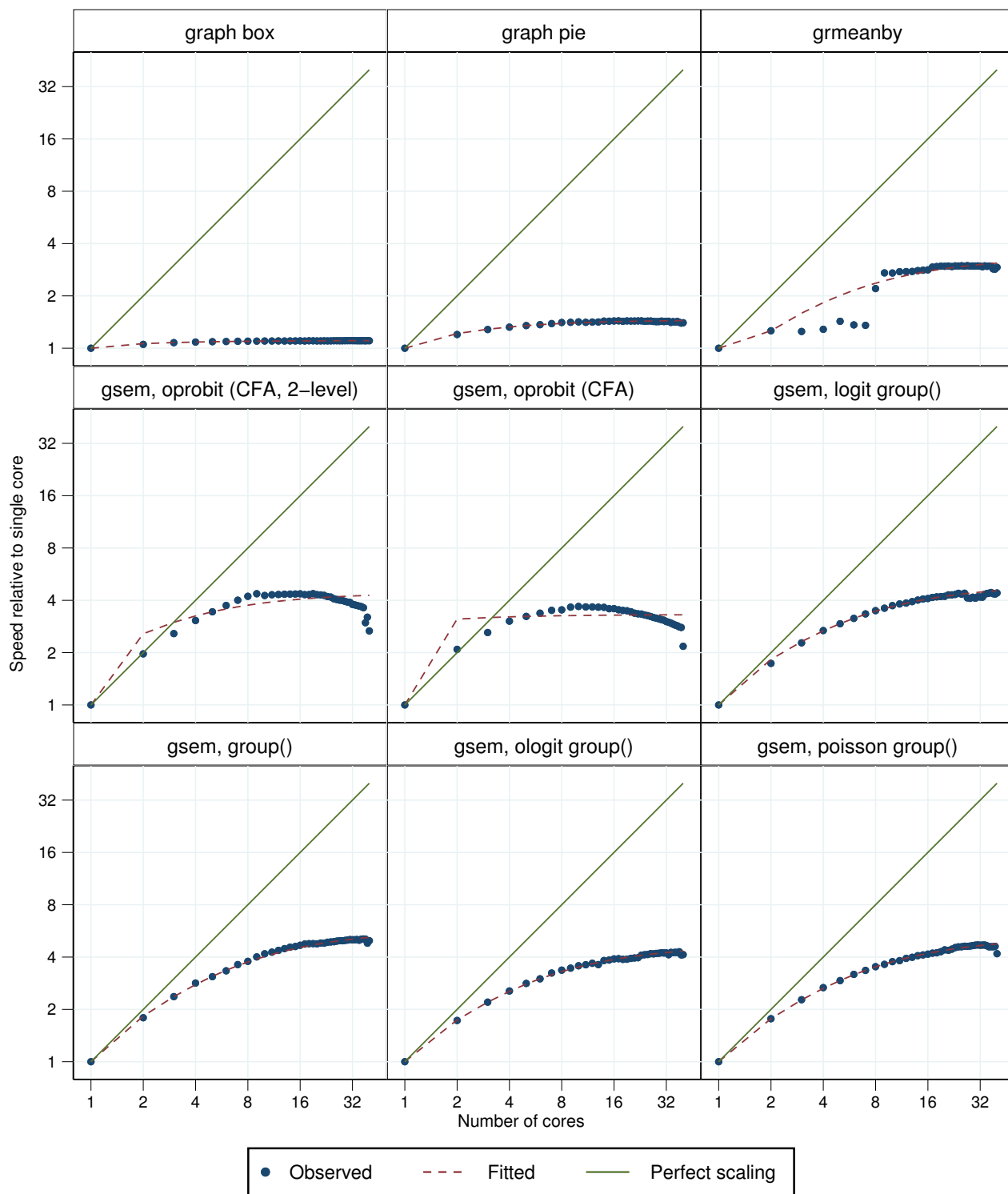


Figure 626. Parallelization performance plots.

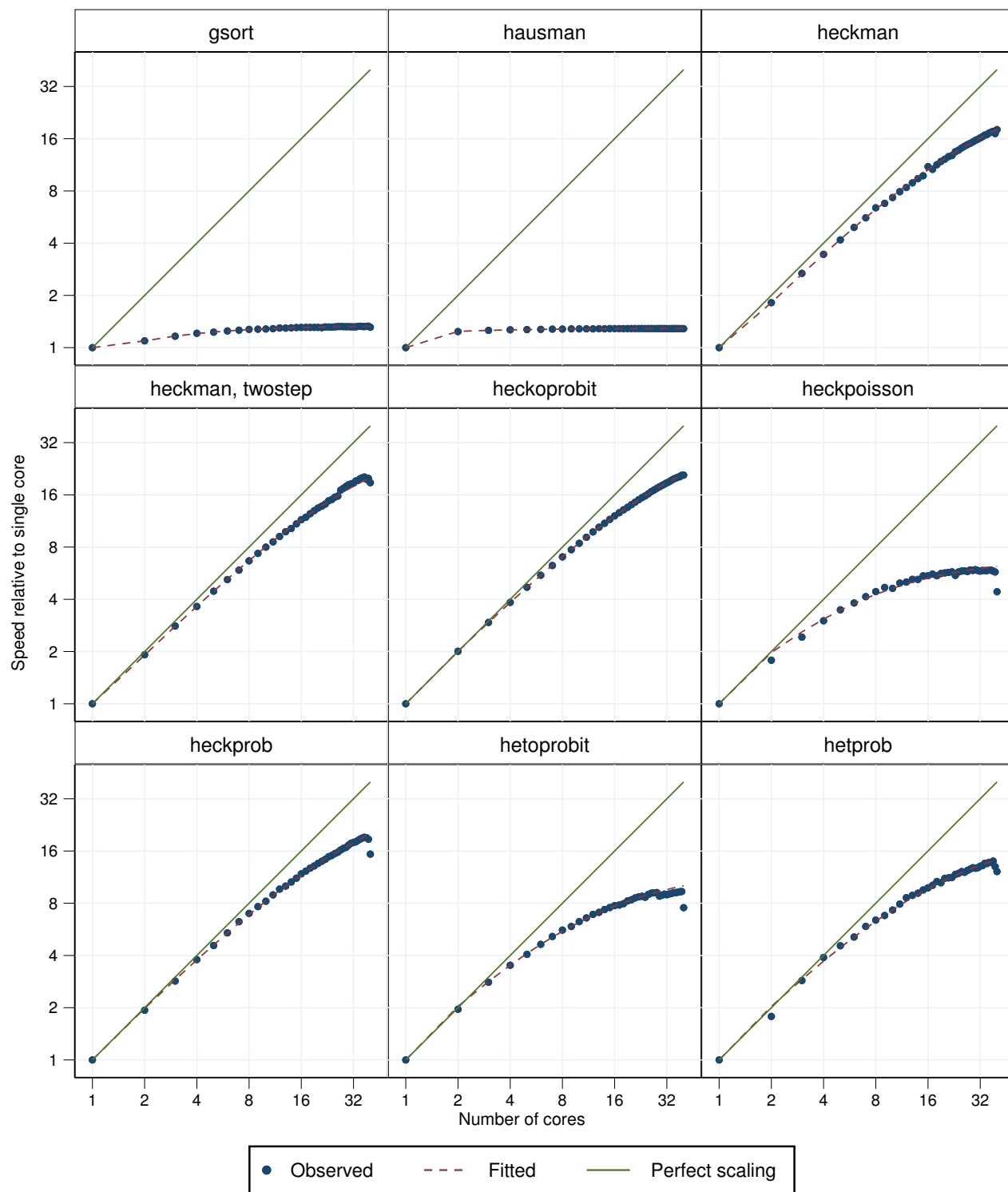


Figure 627. Parallelization performance plots.

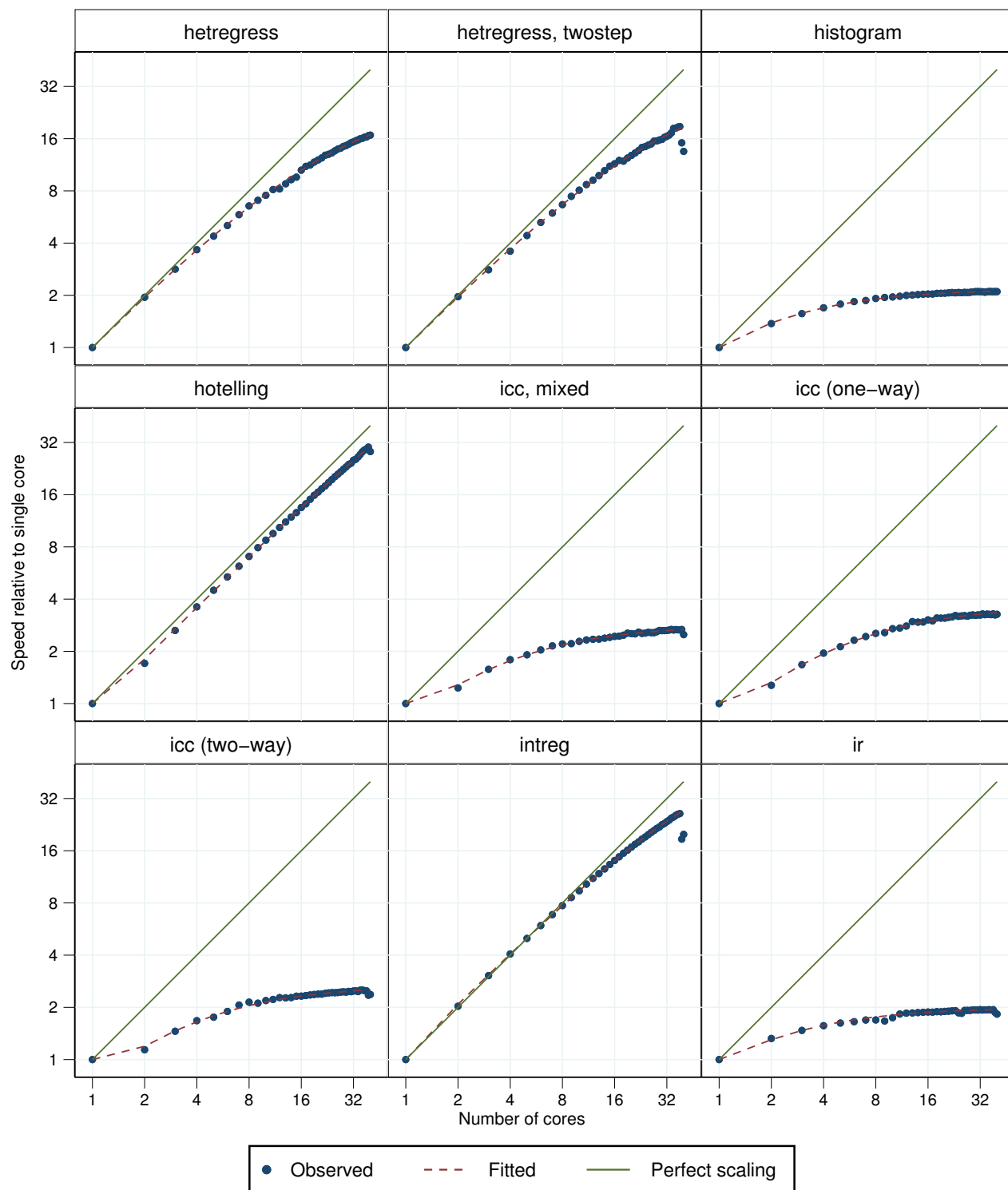


Figure 628. Parallelization performance plots.

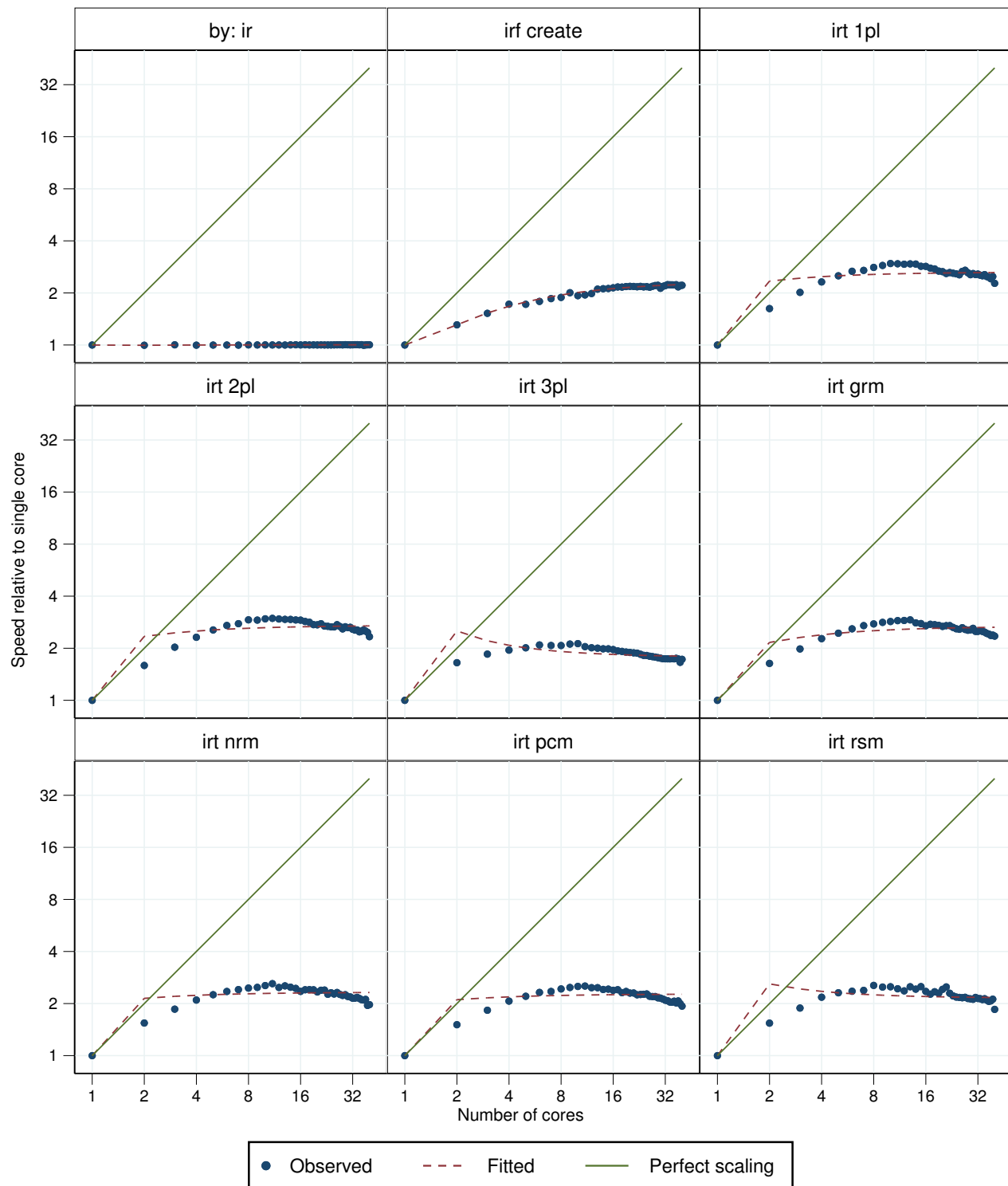


Figure 629. Parallelization performance plots.

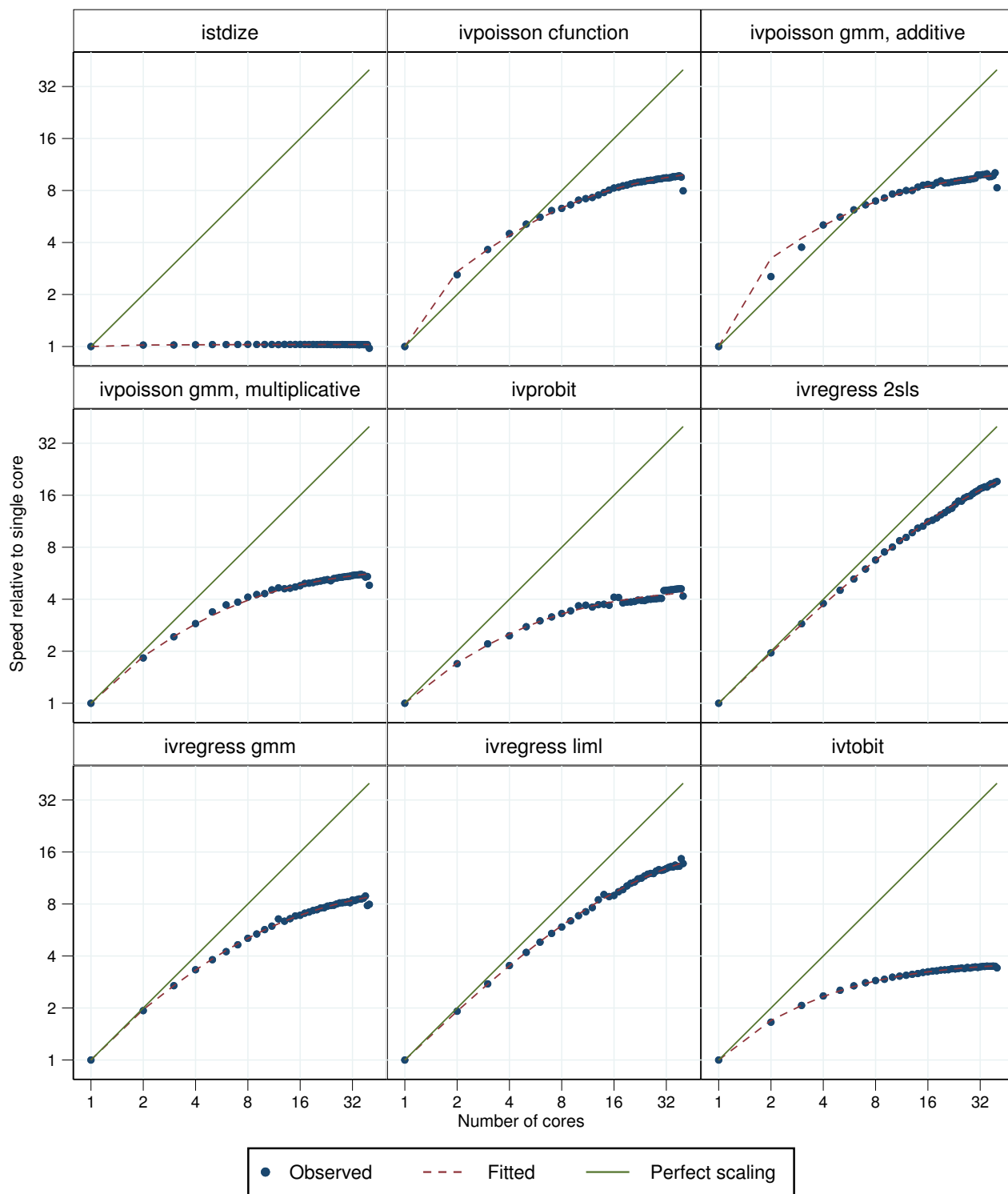


Figure 630. Parallelization performance plots.



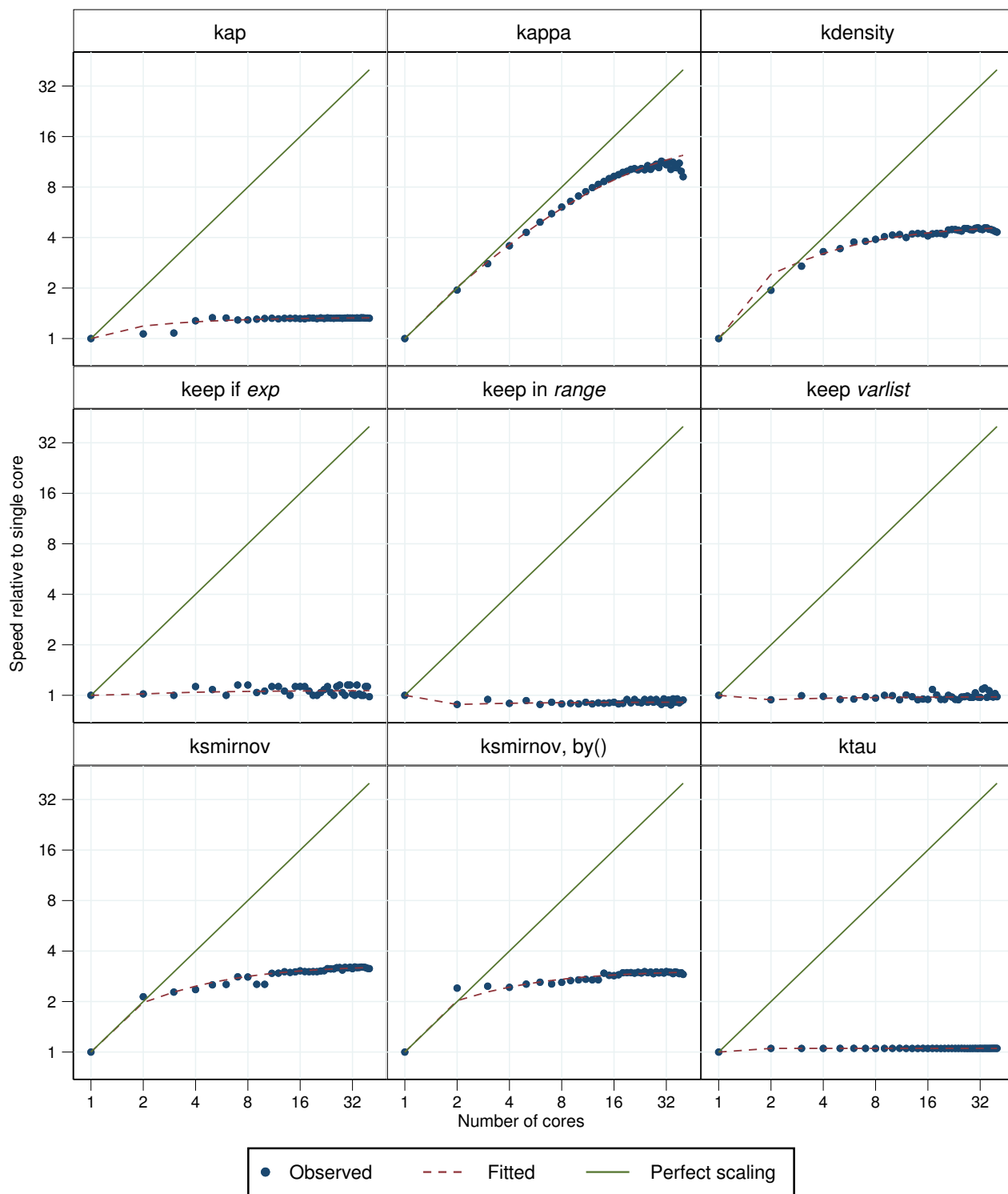


Figure 631. Parallelization performance plots.

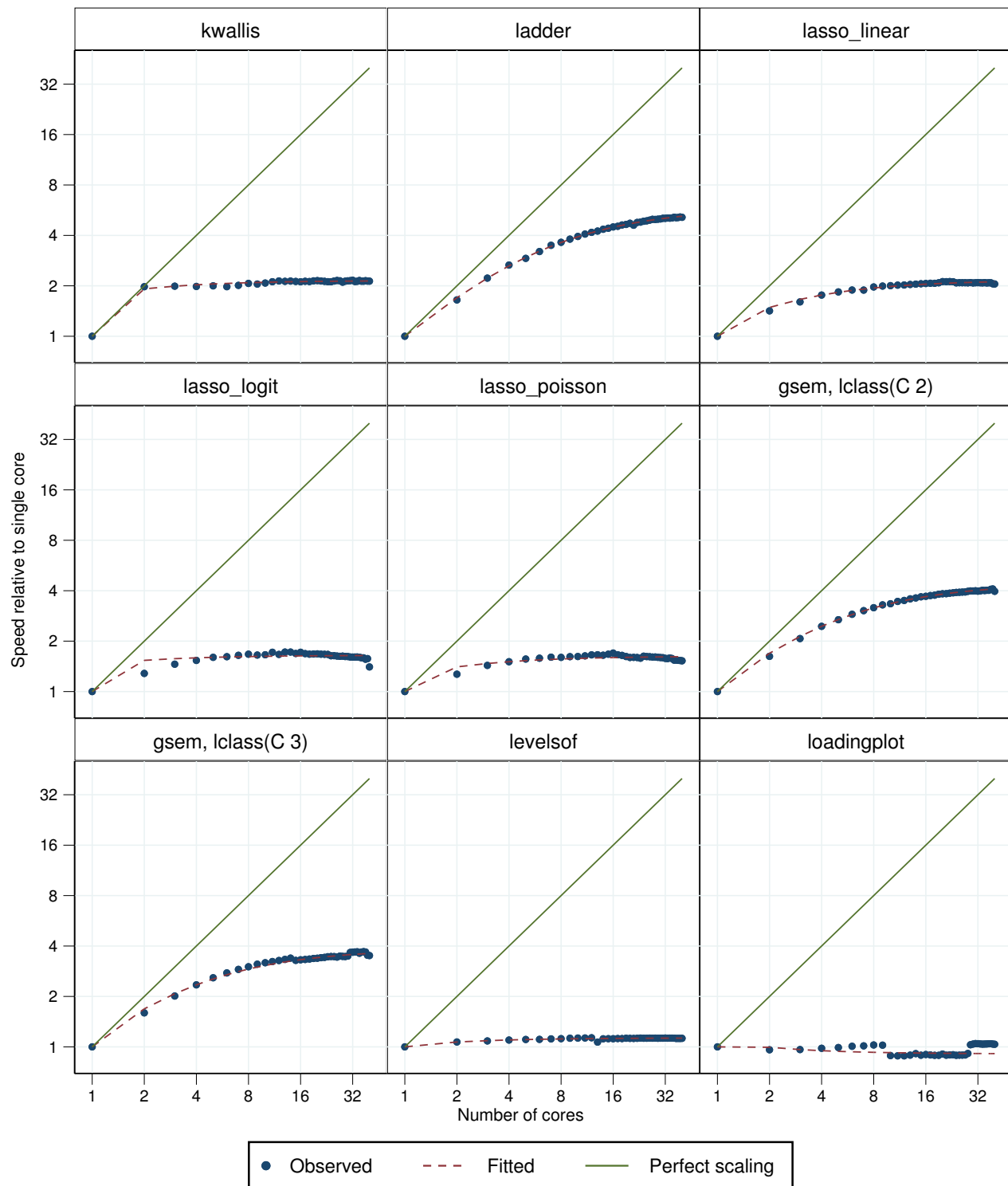


Figure 632. Parallelization performance plots.

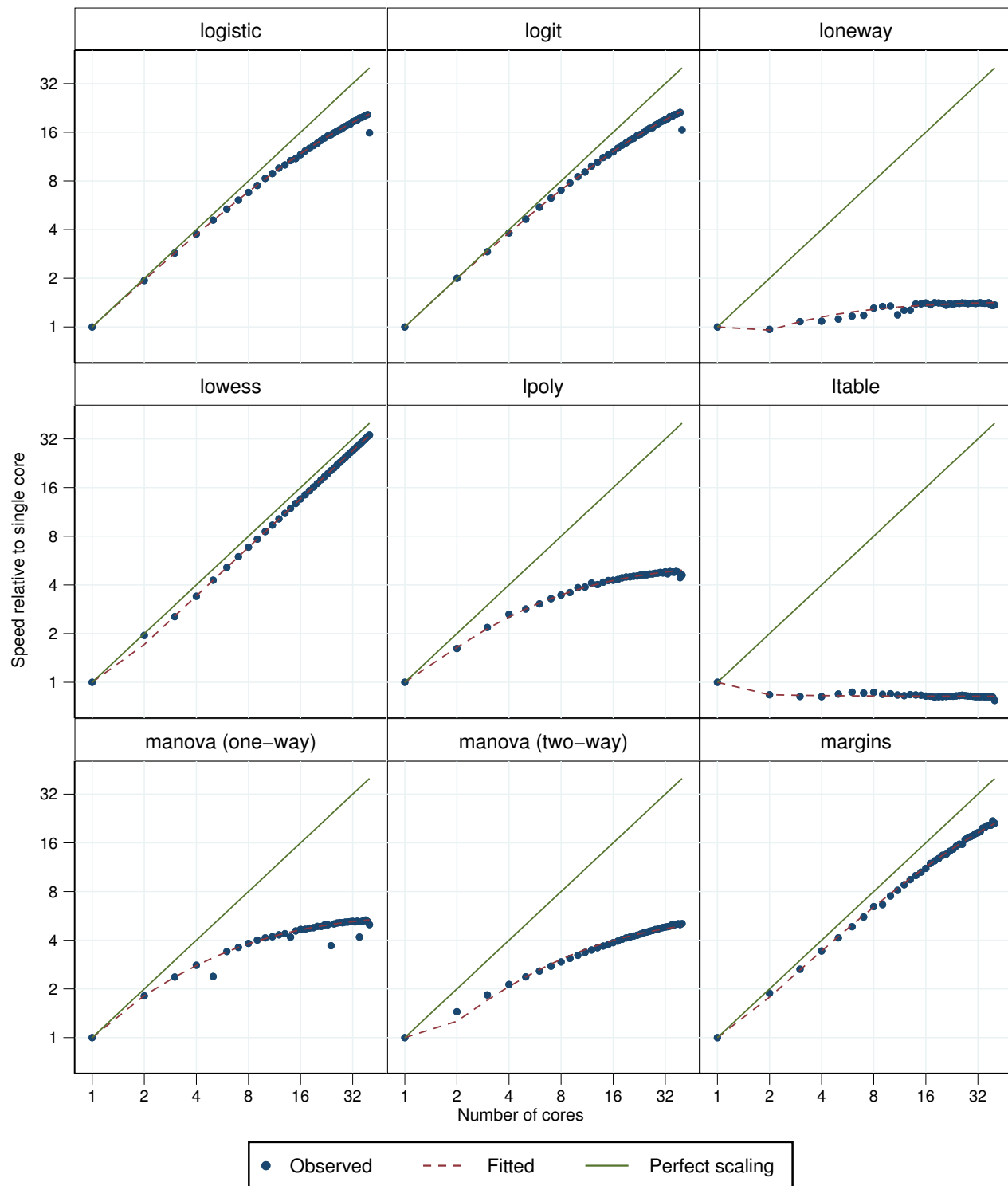


Figure 633. Parallelization performance plots.

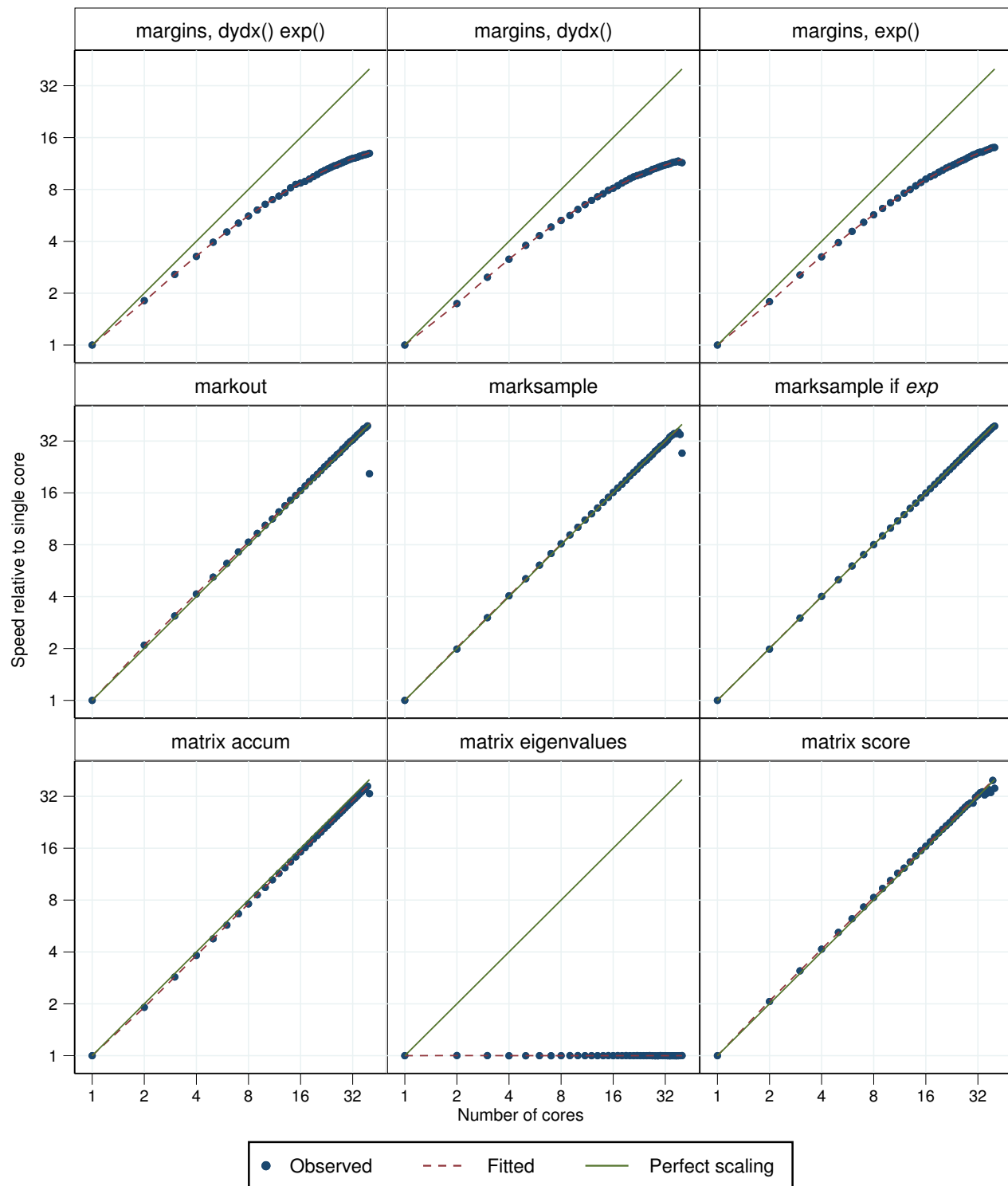


Figure 634. Parallelization performance plots.

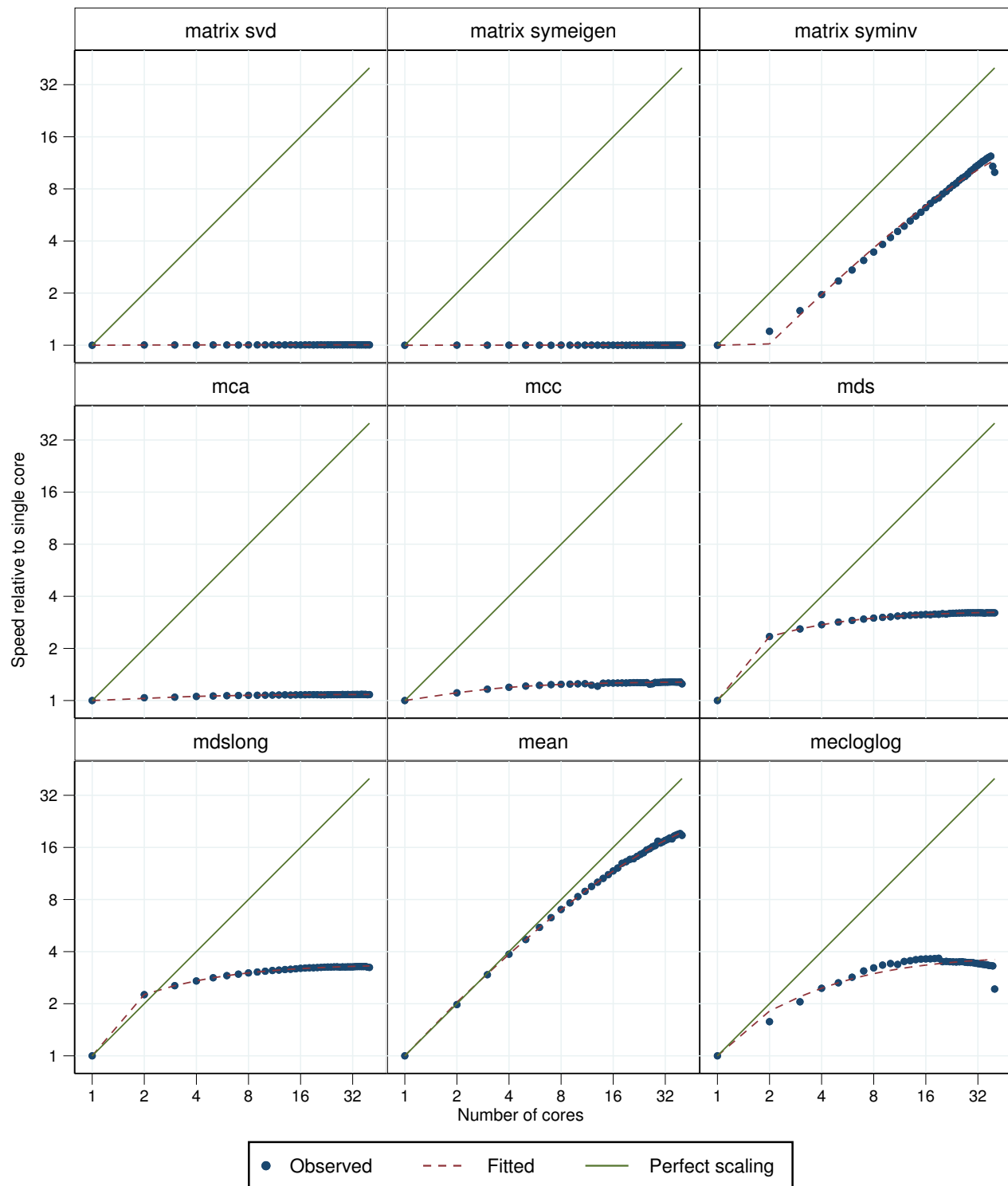


Figure 635. Parallelization performance plots.

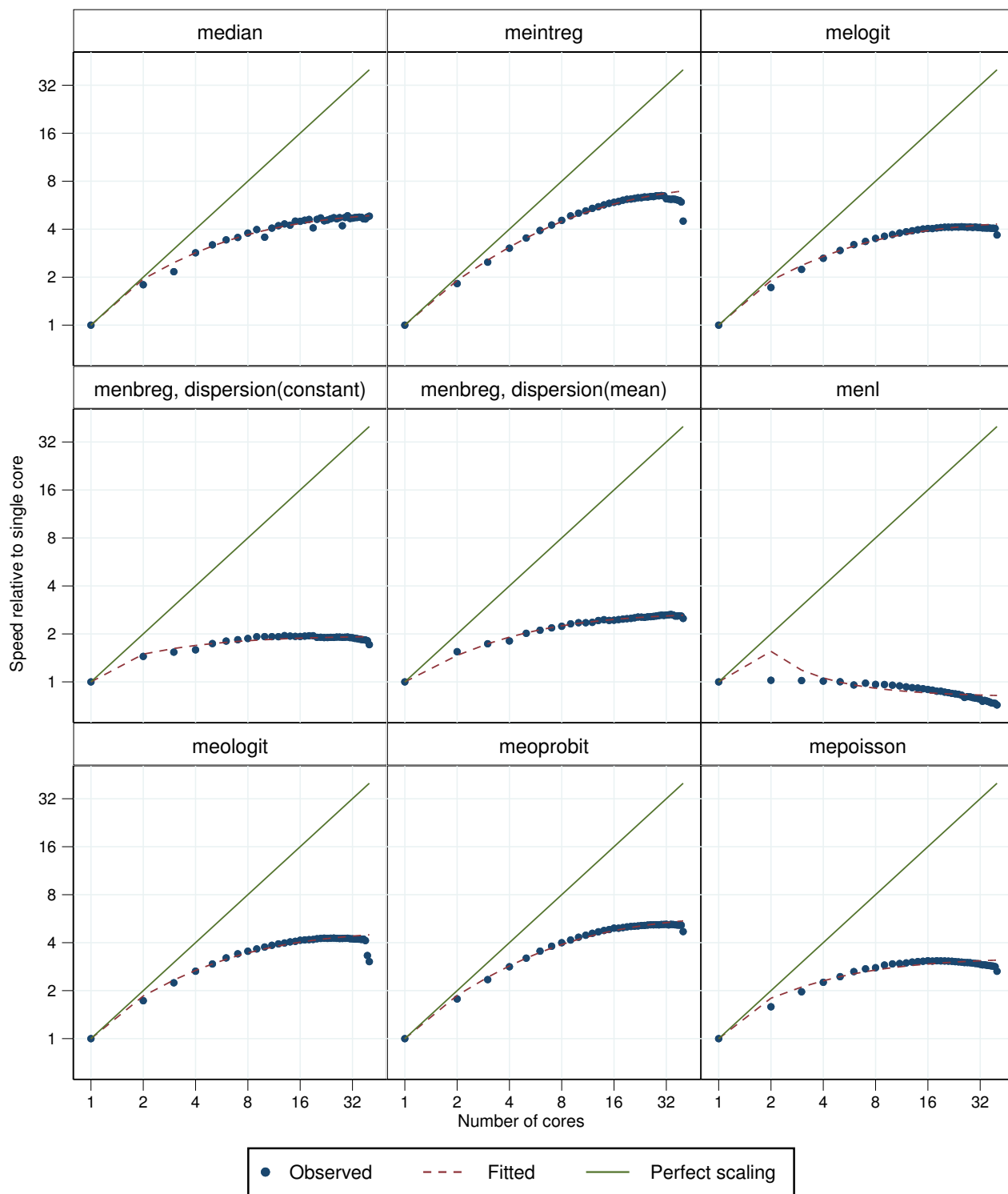


Figure 636. Parallelization performance plots.

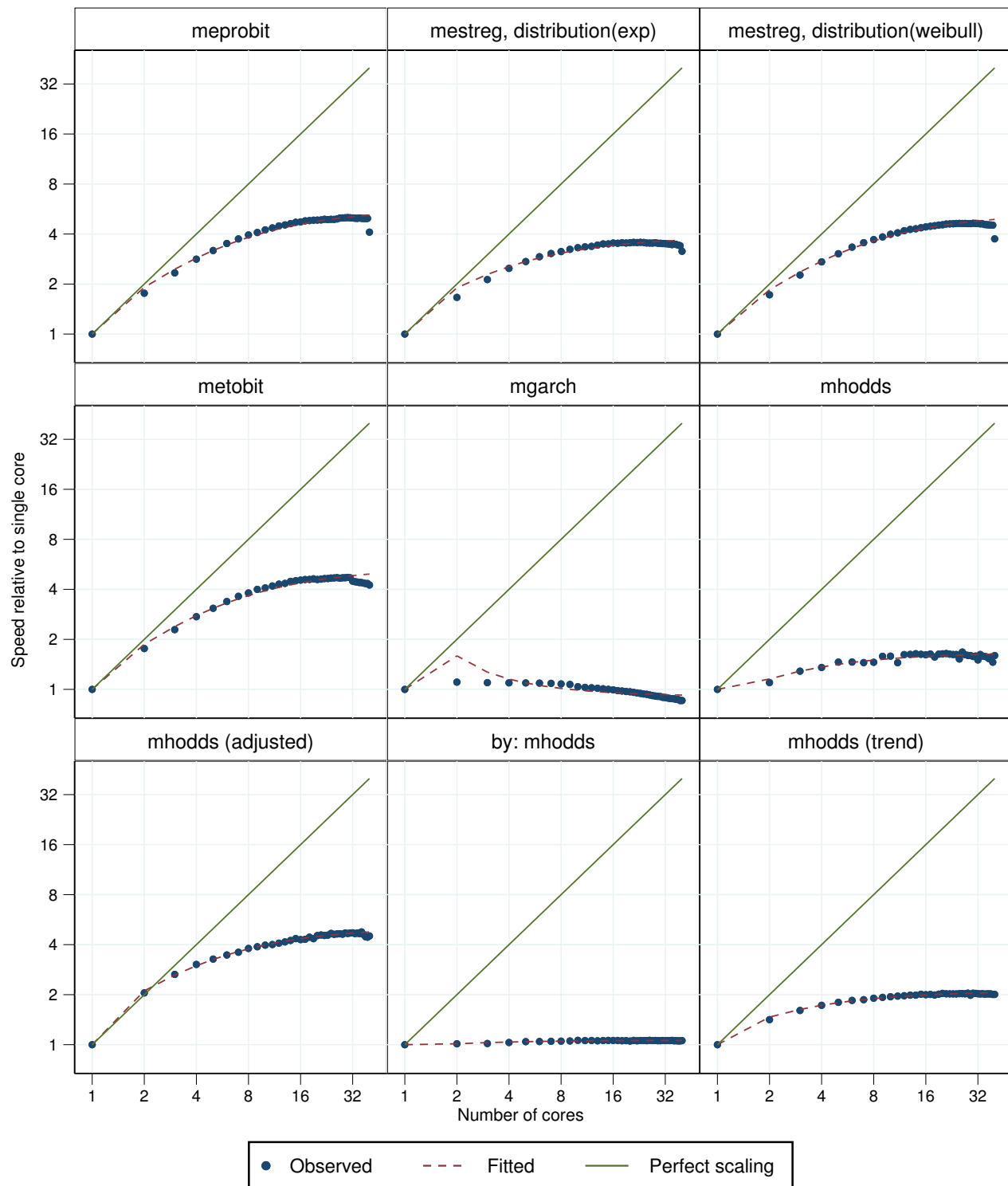


Figure 637. Parallelization performance plots.

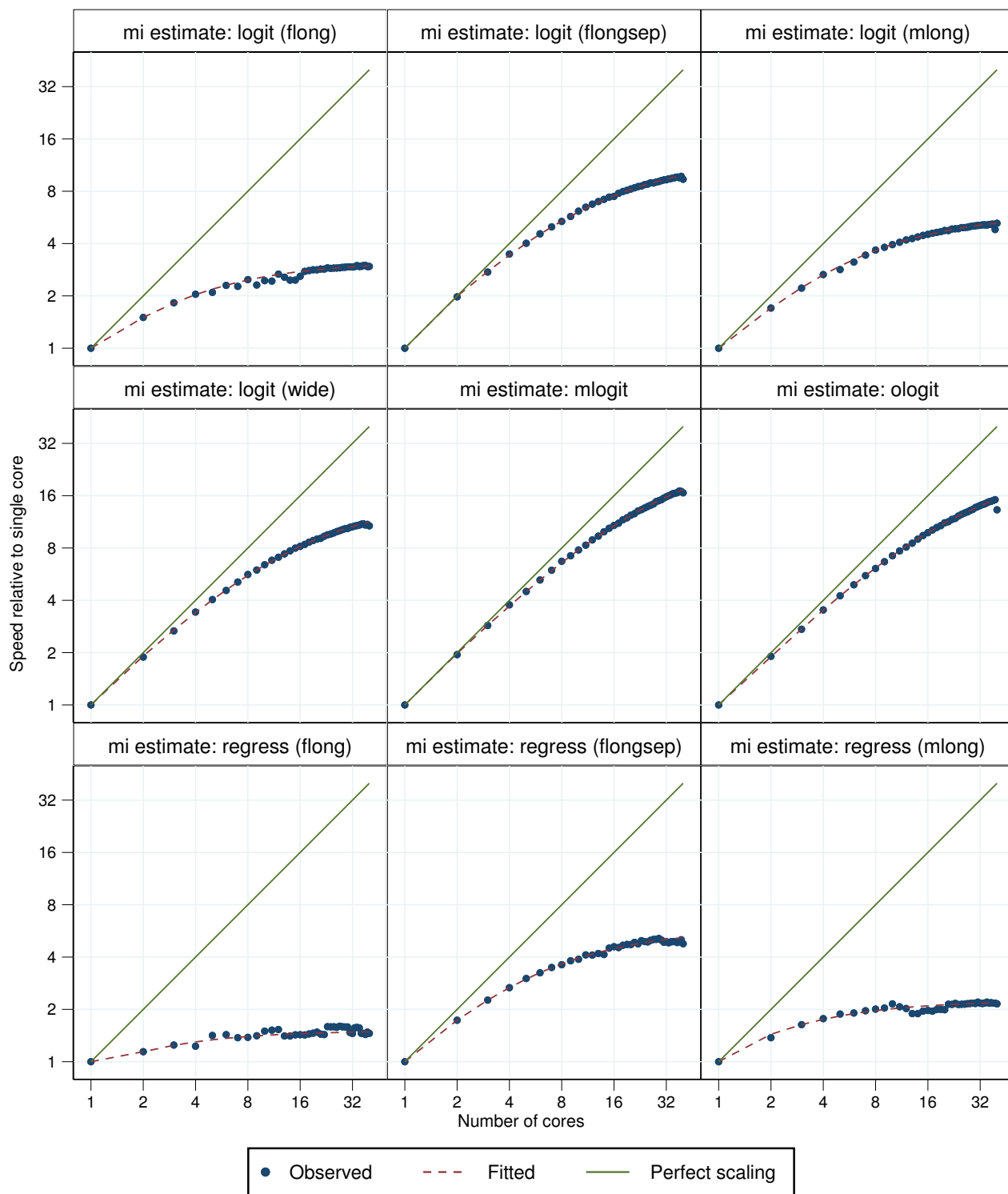


Figure 638. Parallelization performance plots.



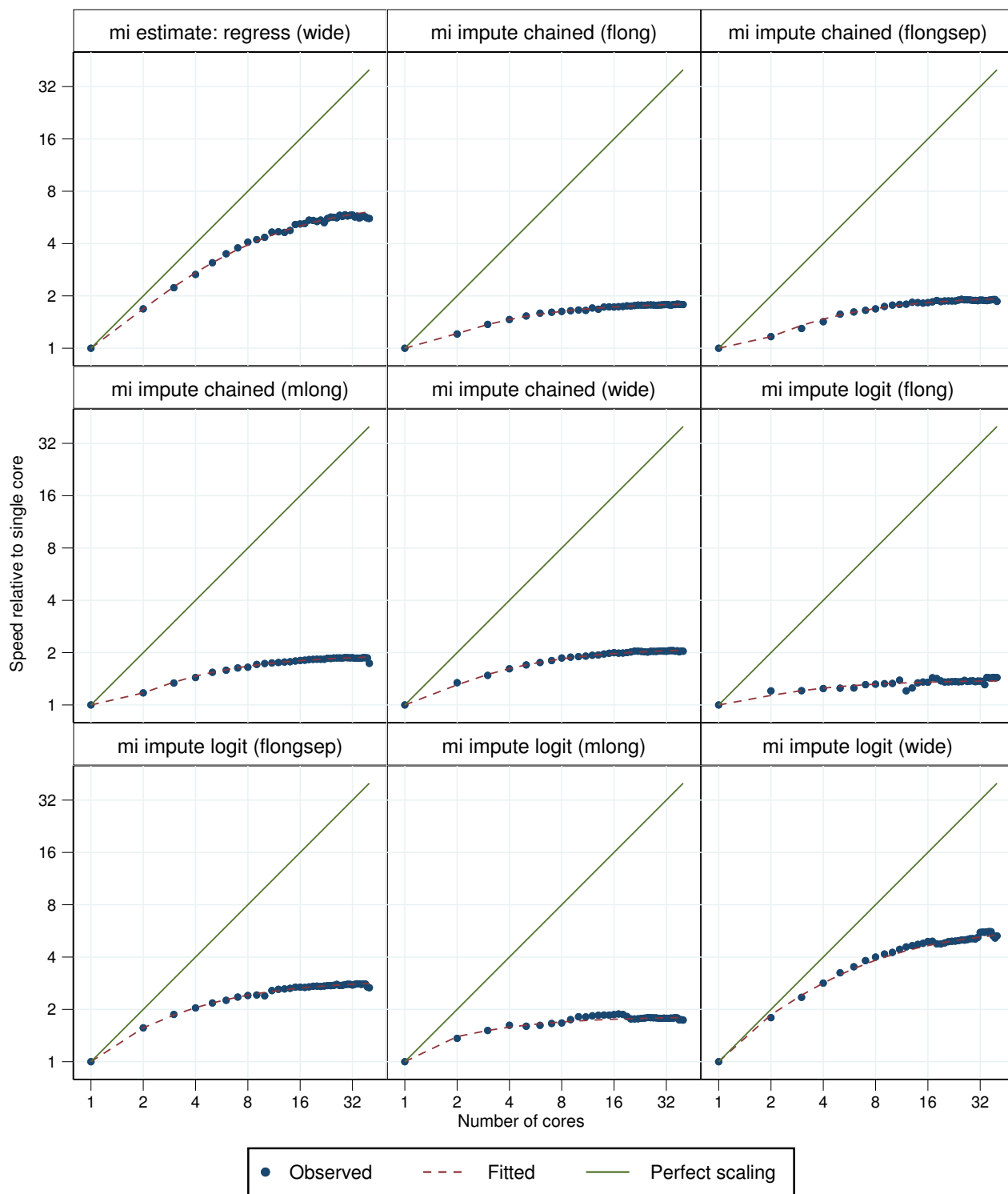


Figure 639. Parallelization performance plots.

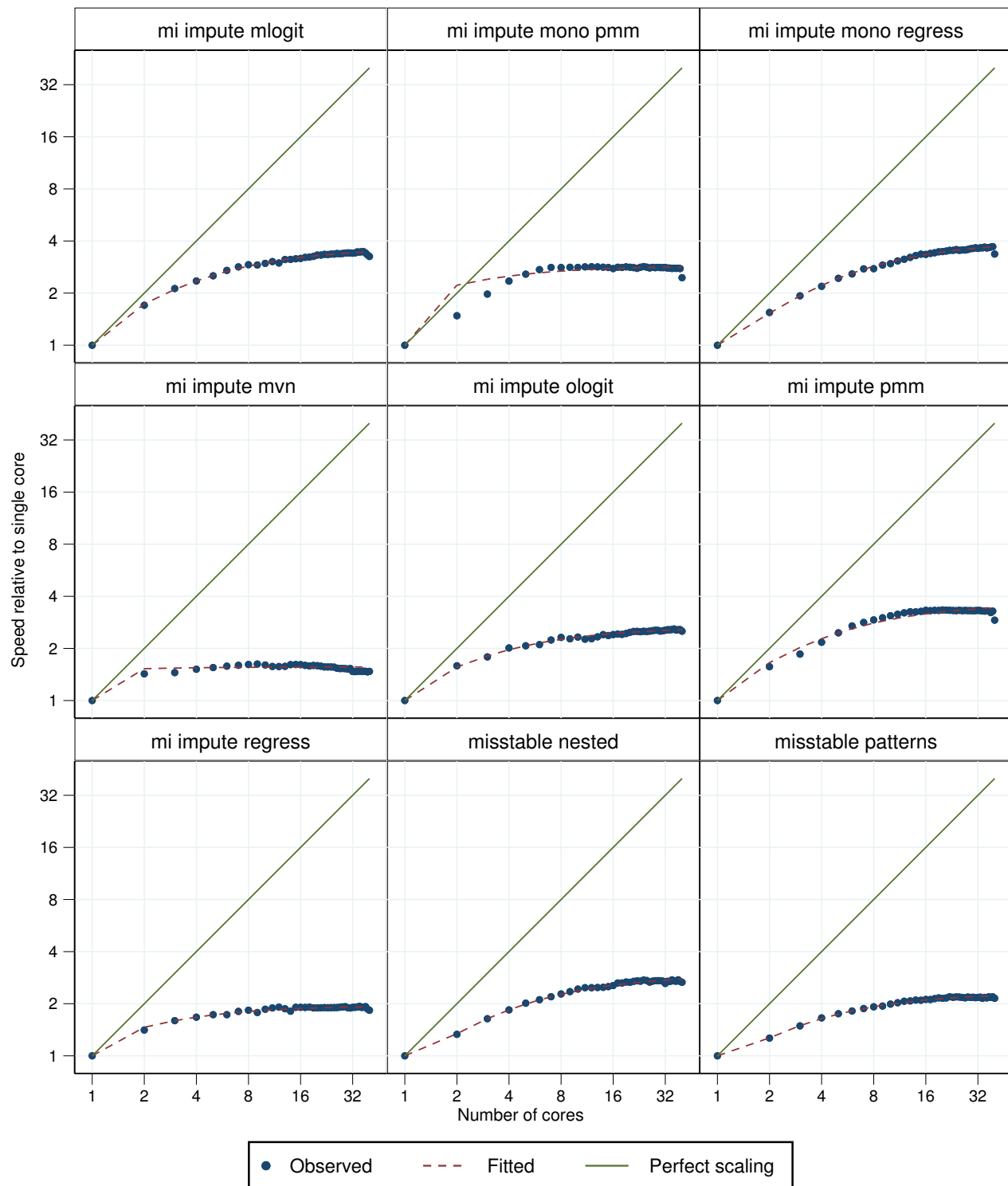


Figure 640. Parallelization performance plots.

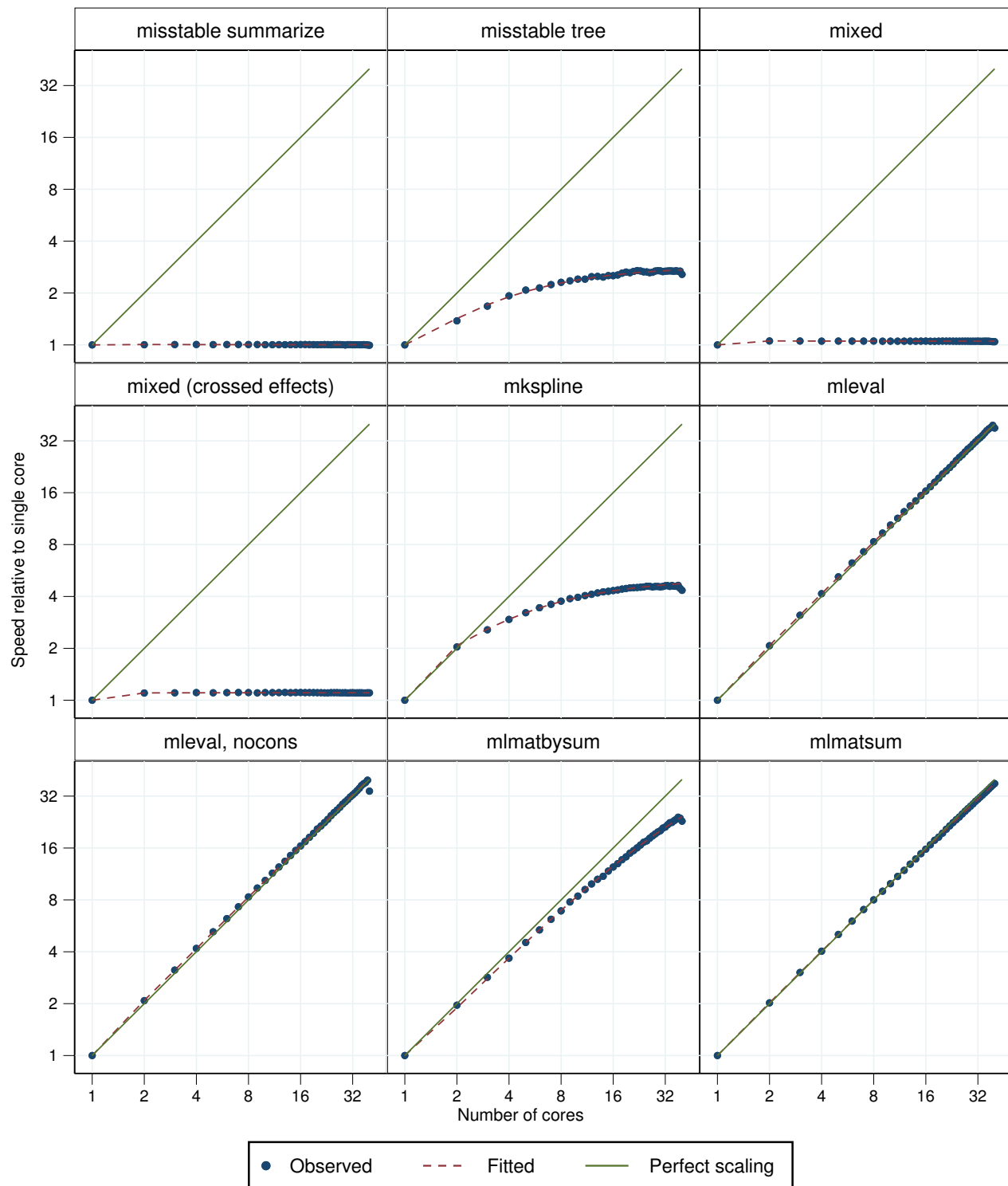


Figure 641. Parallelization performance plots.

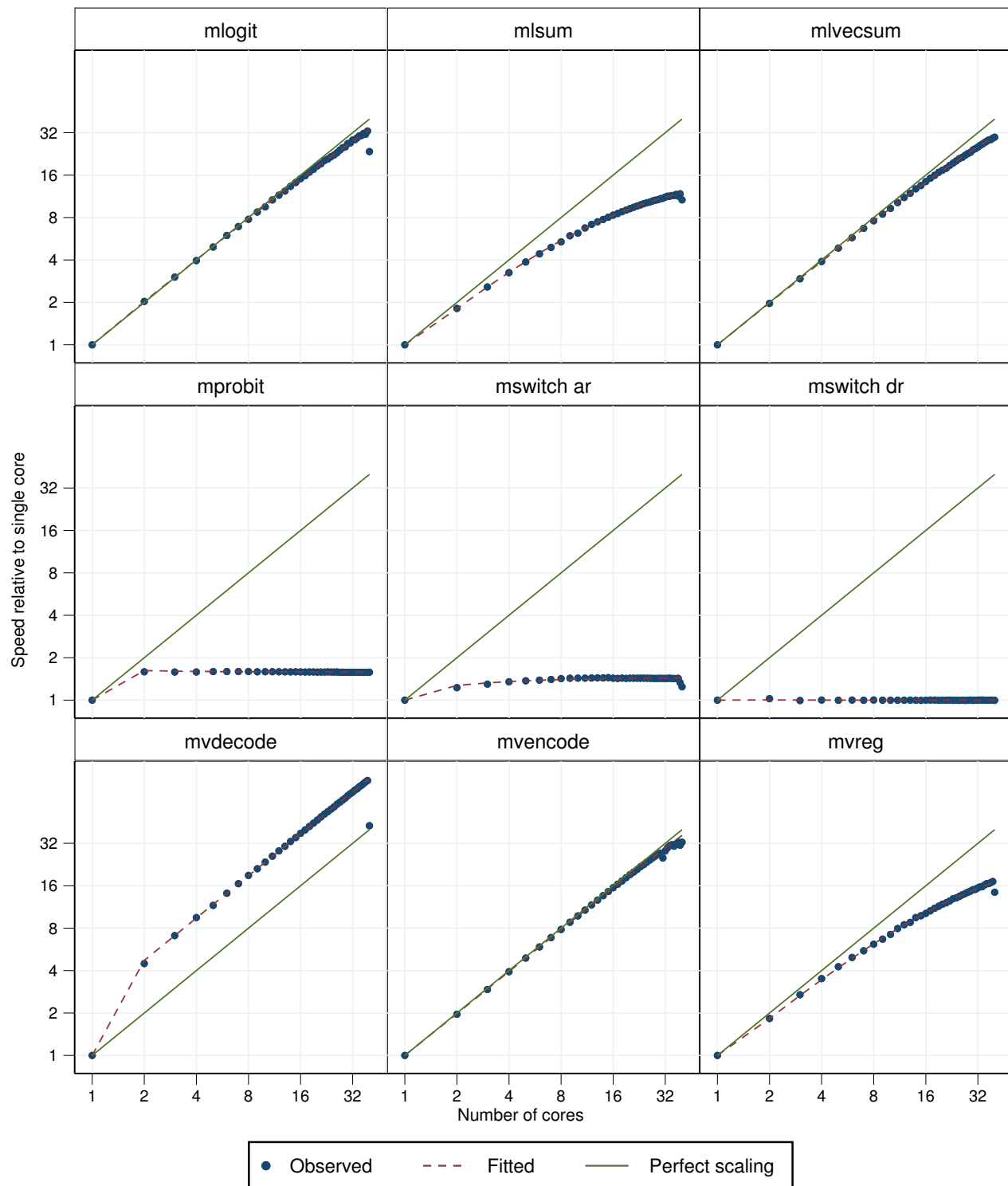


Figure 642. Parallelization performance plots.

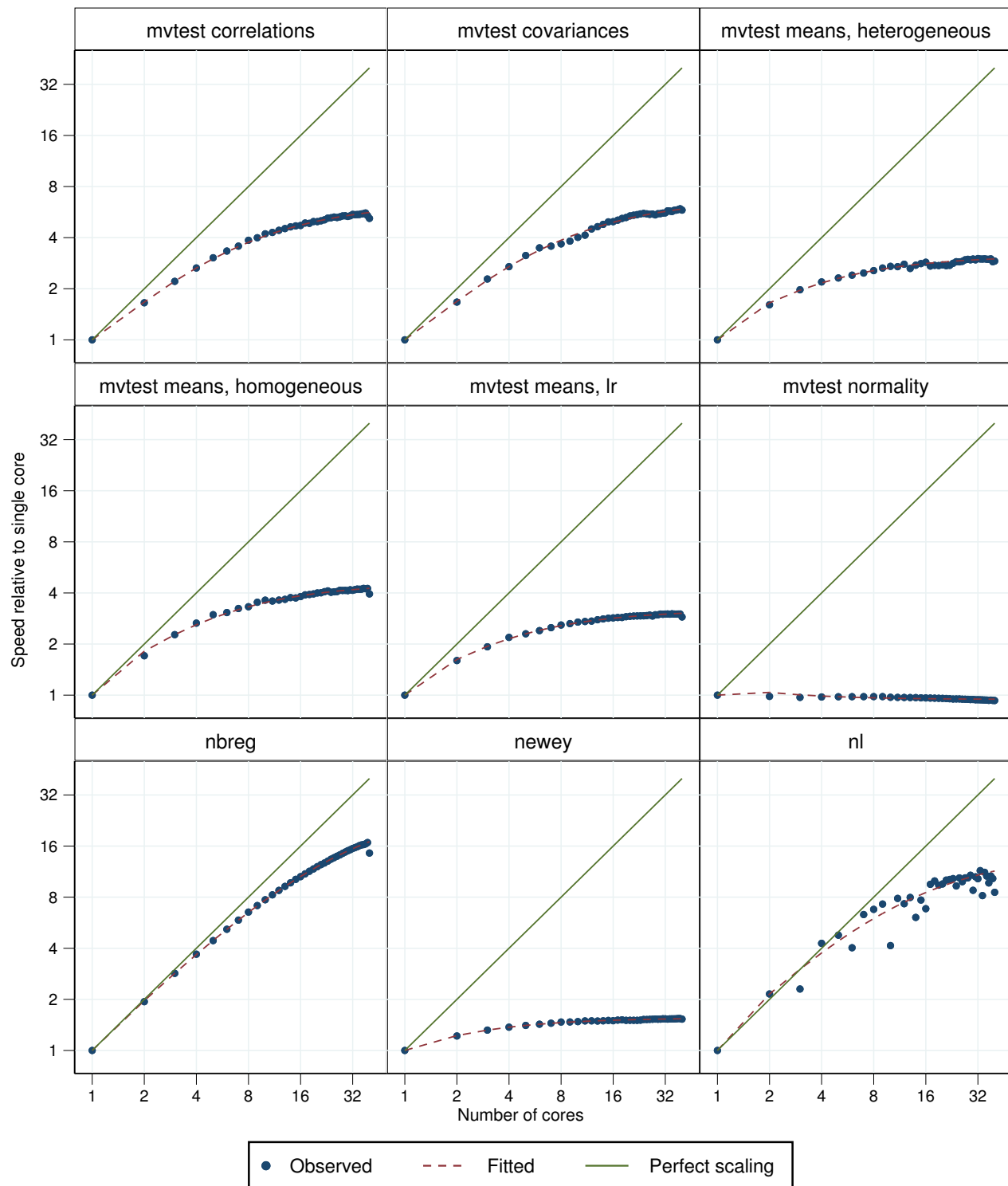


Figure 643. Parallelization performance plots.

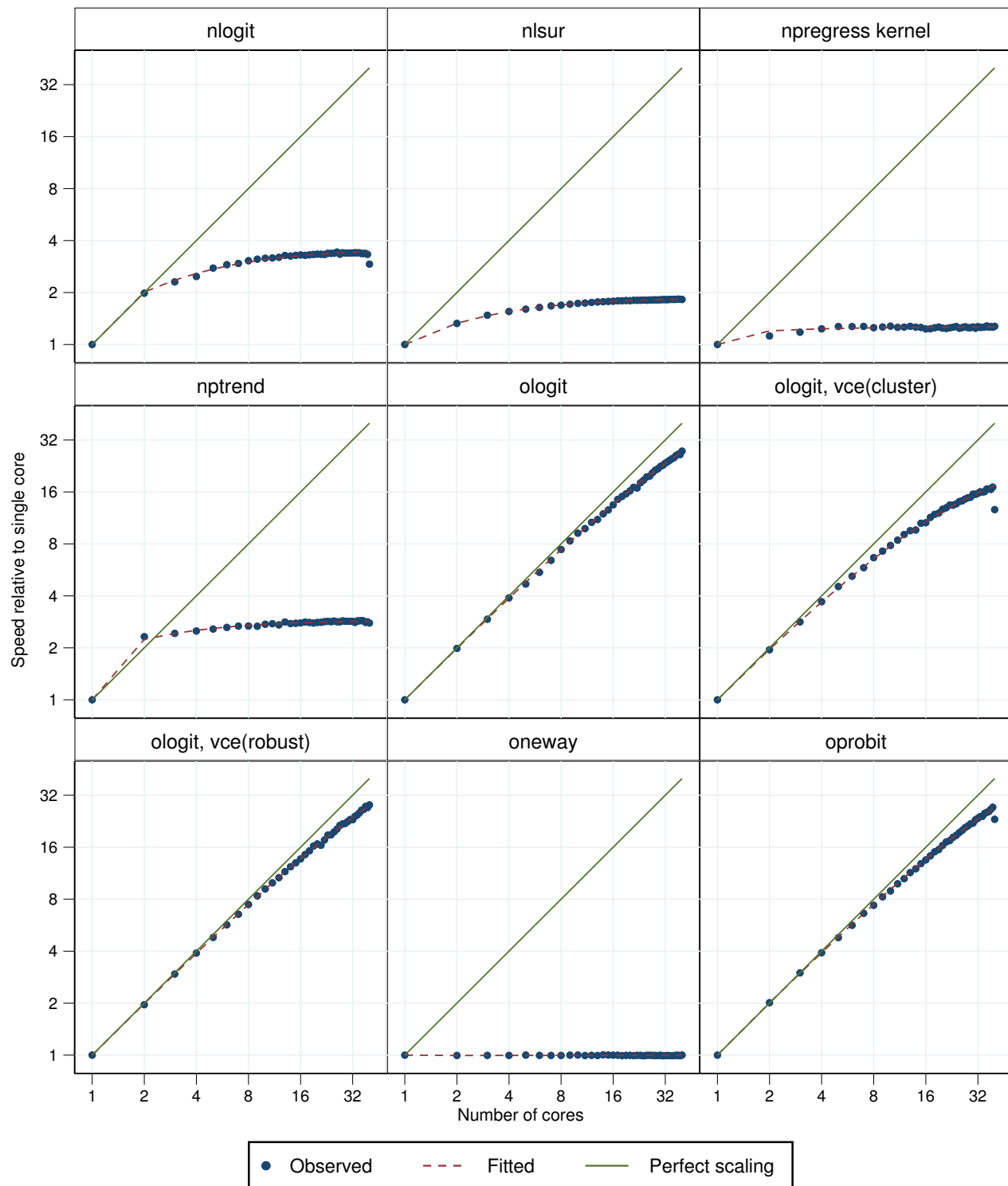


Figure 644. Parallelization performance plots.

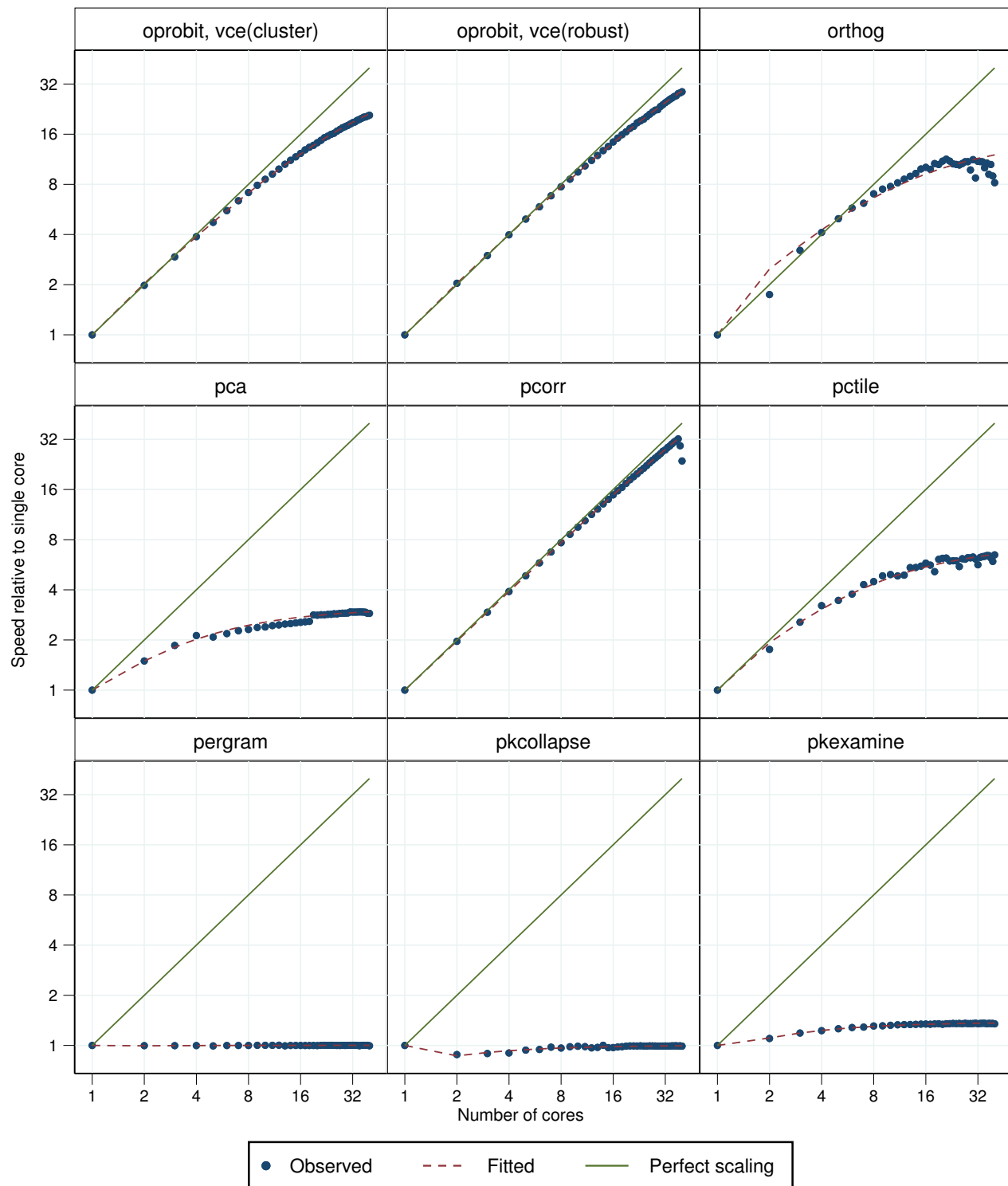


Figure 645. Parallelization performance plots.

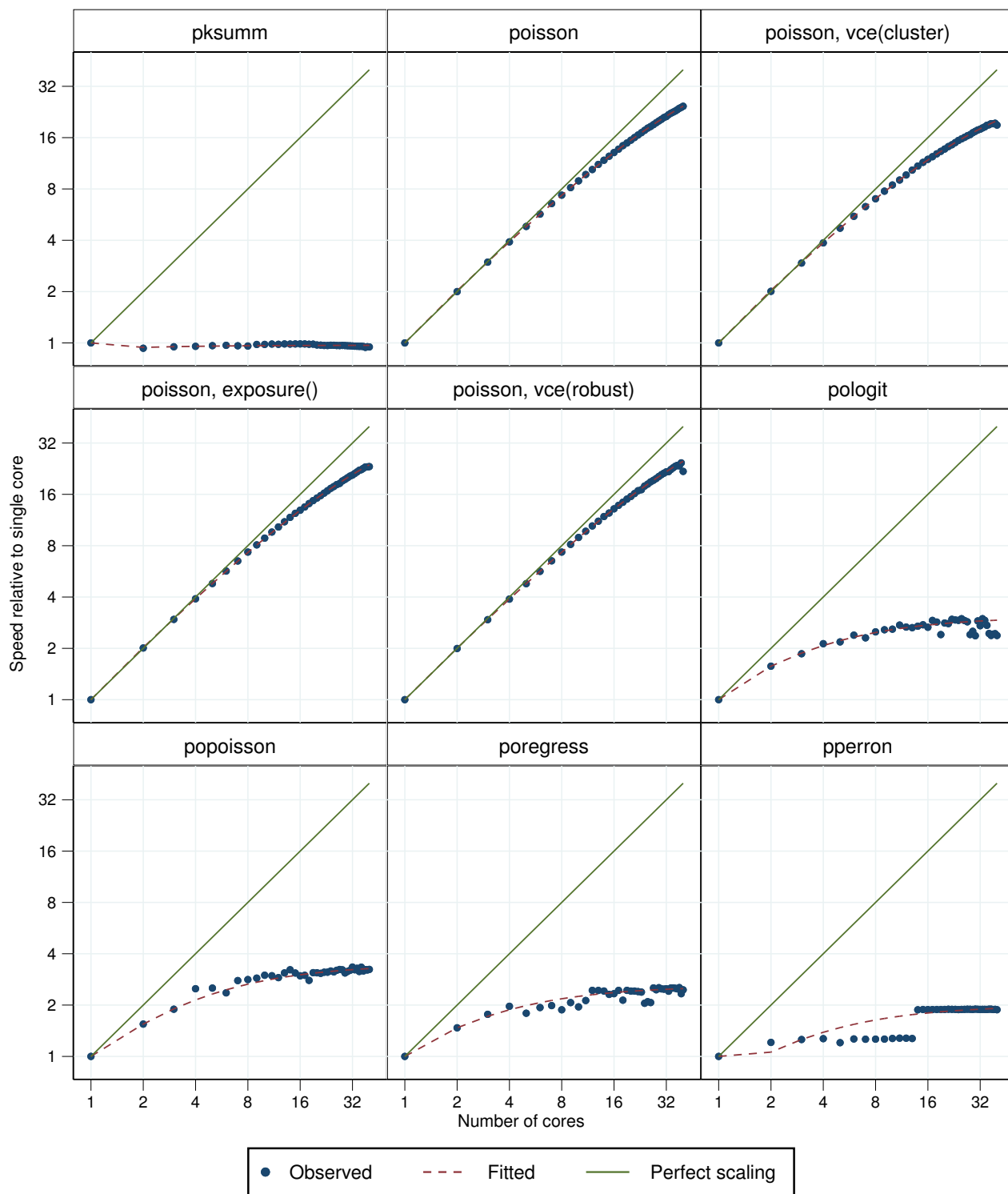


Figure 646. Parallelization performance plots.



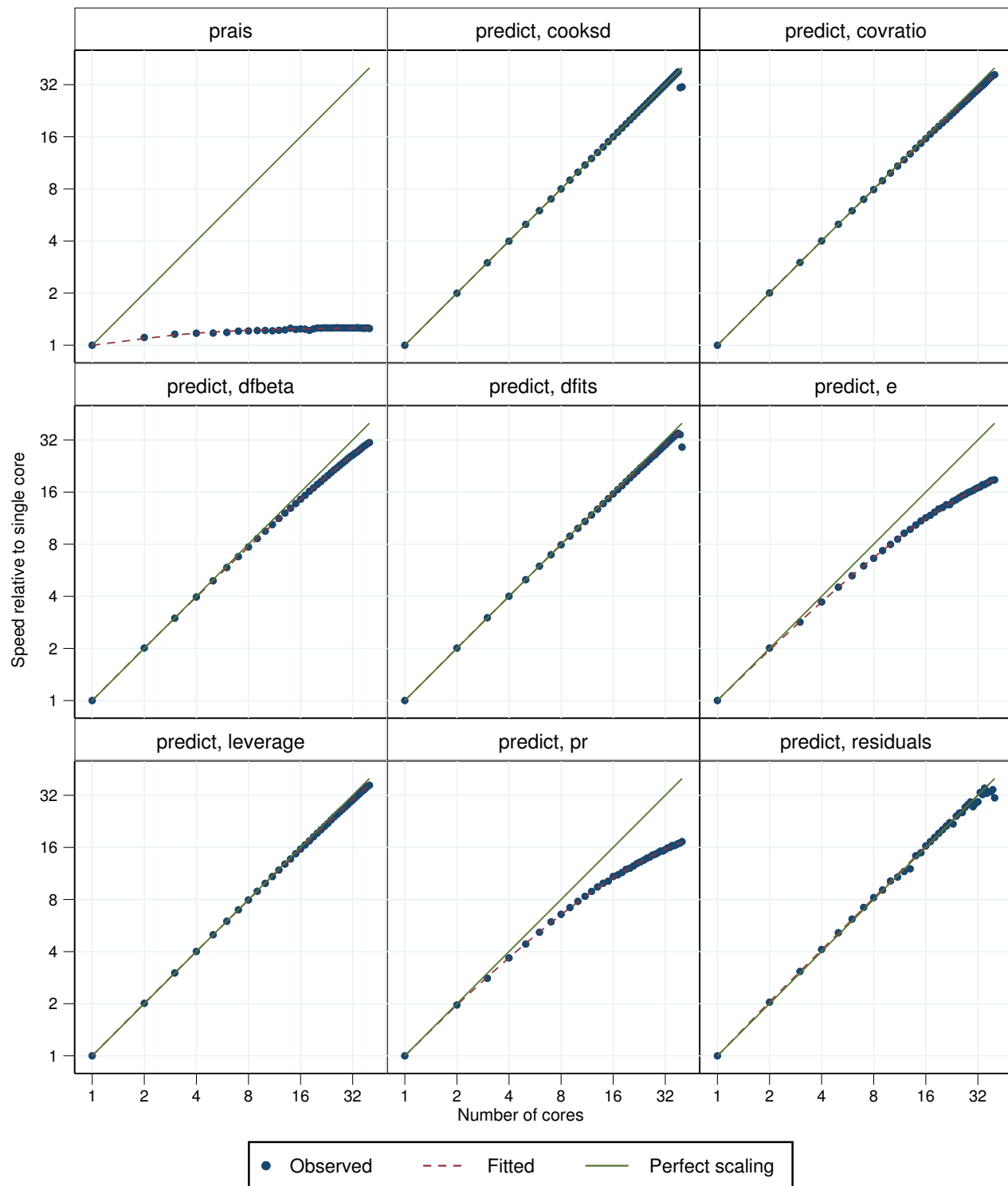


Figure 647. Parallelization performance plots.

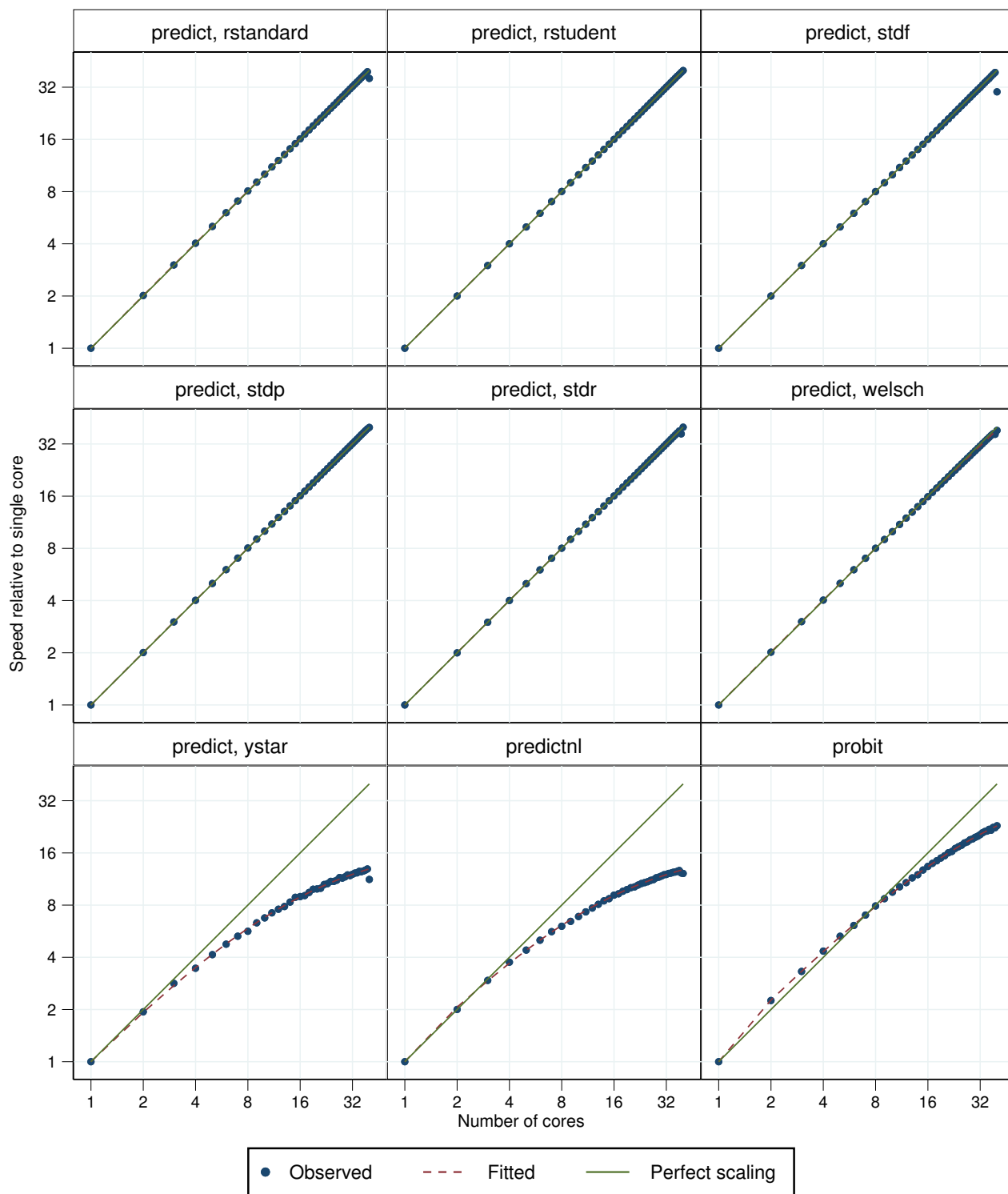


Figure 648. Parallelization performance plots.

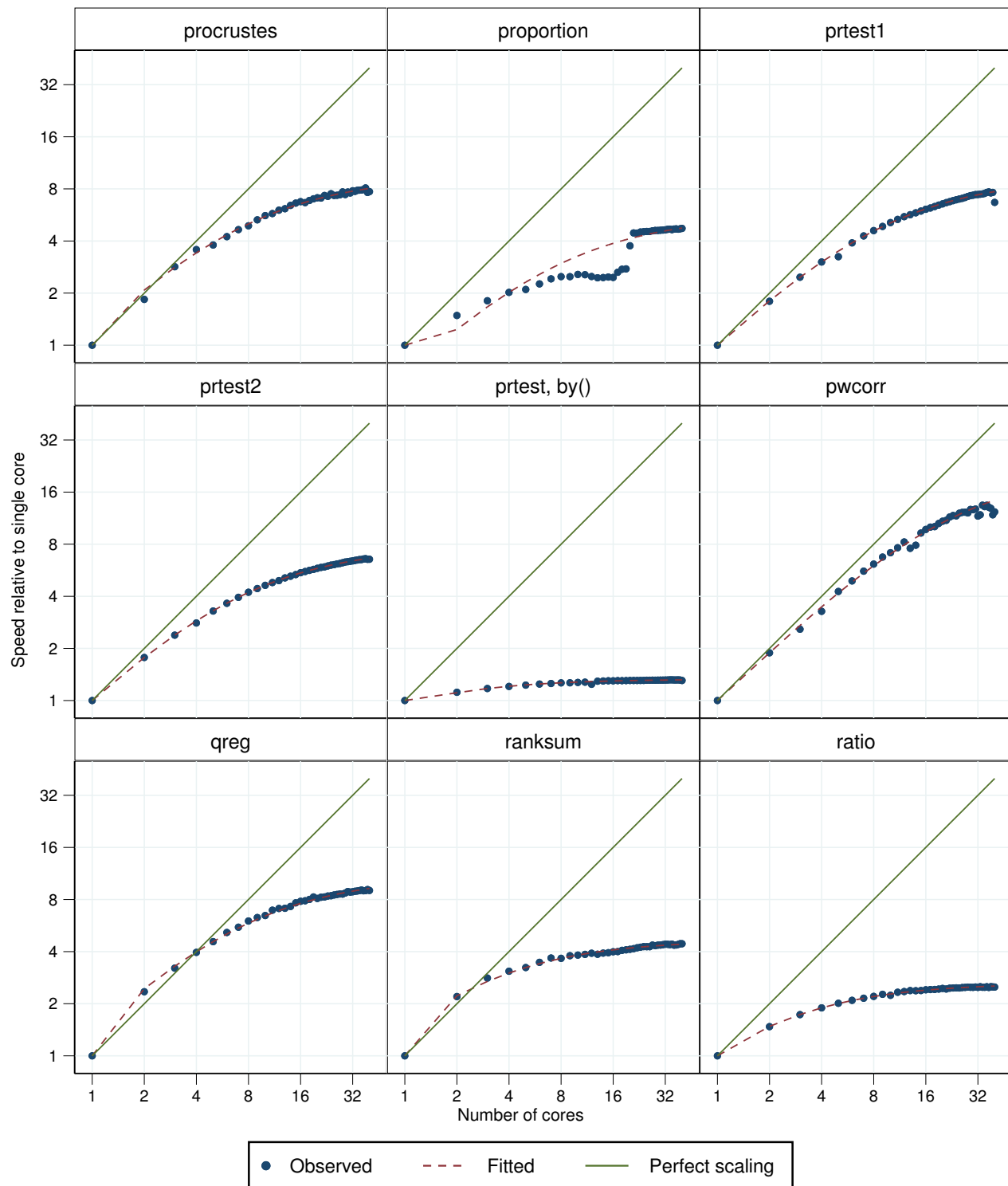


Figure 649. Parallelization performance plots.

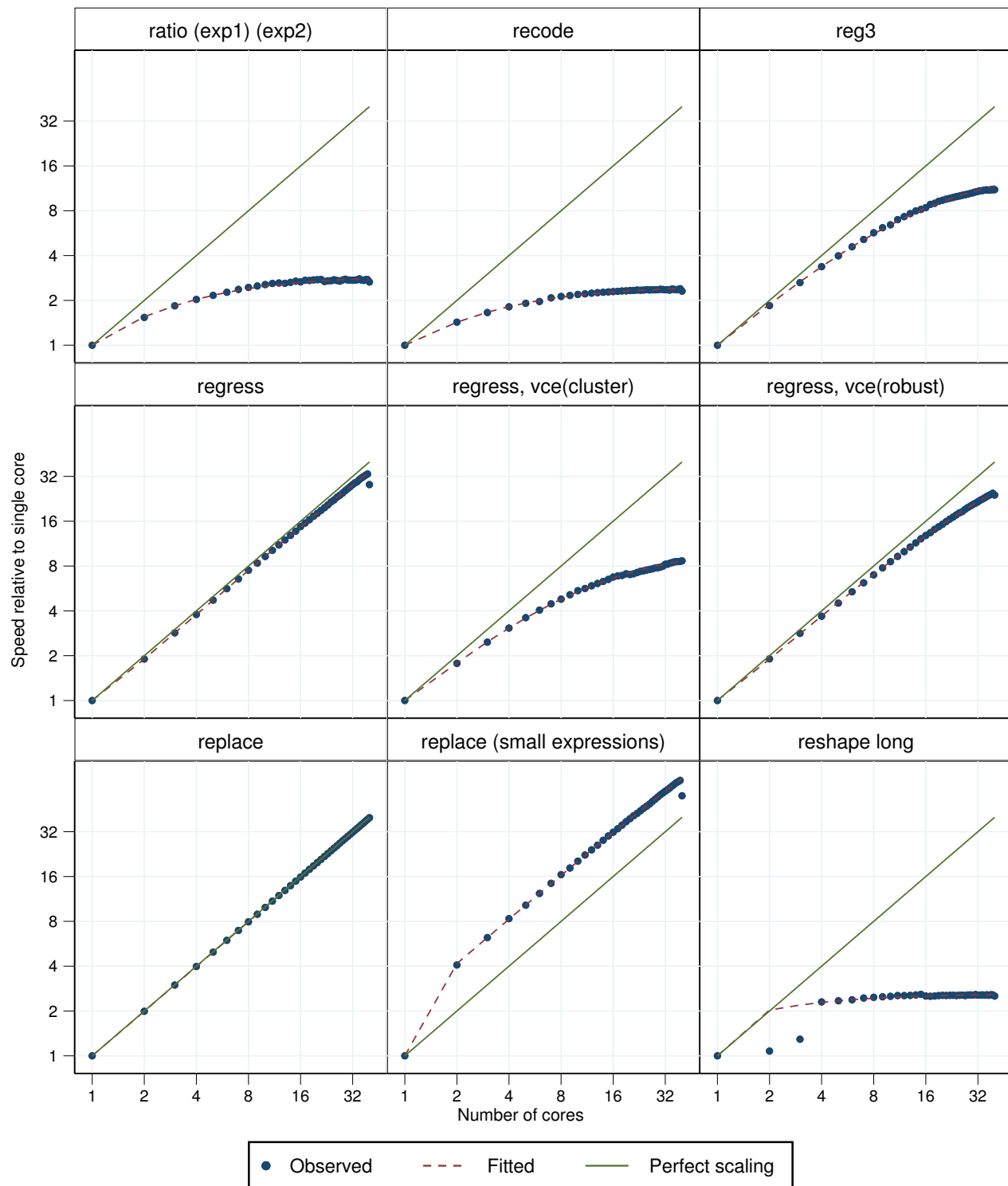


Figure 650. Parallelization performance plots.

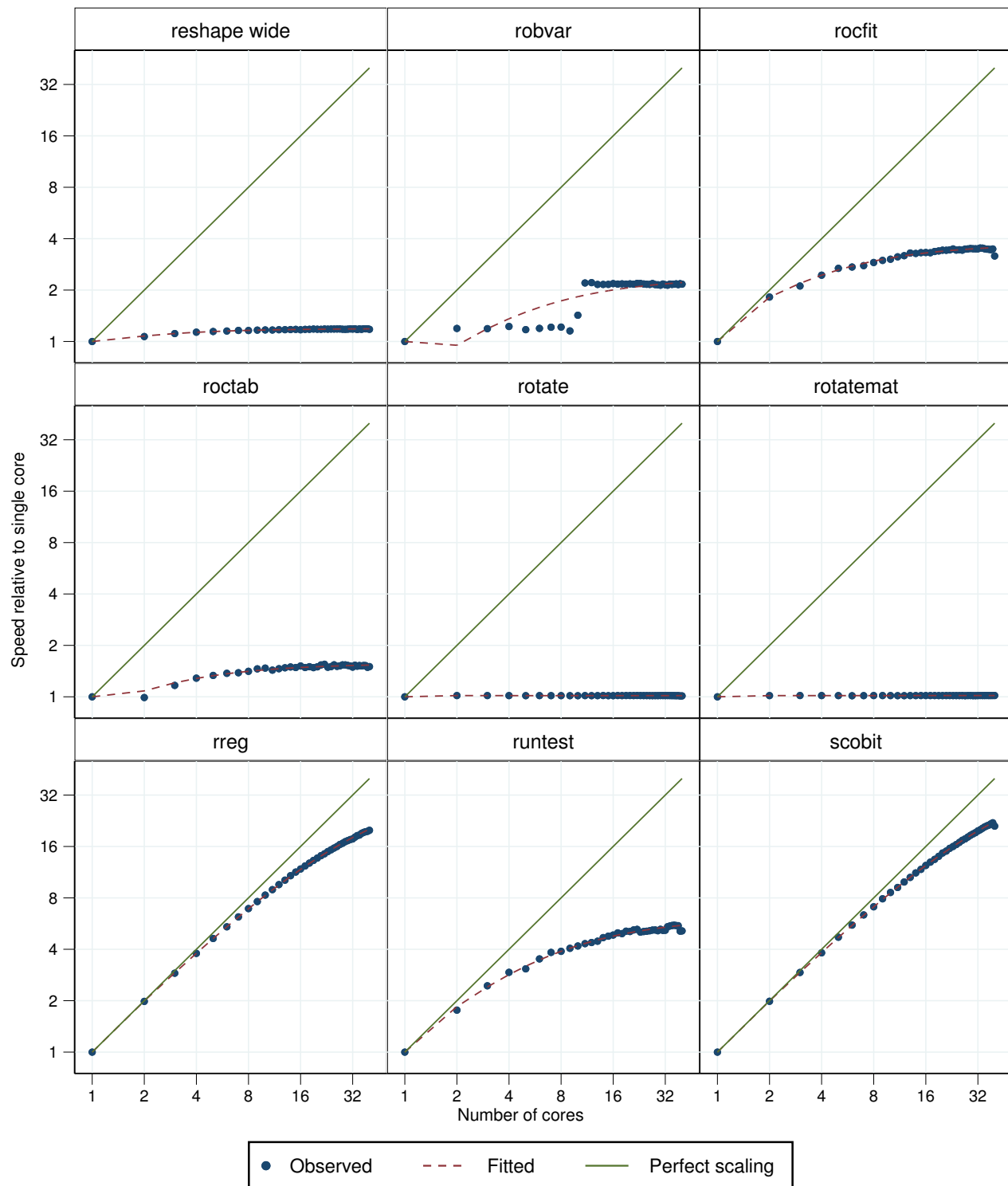


Figure 651. Parallelization performance plots.

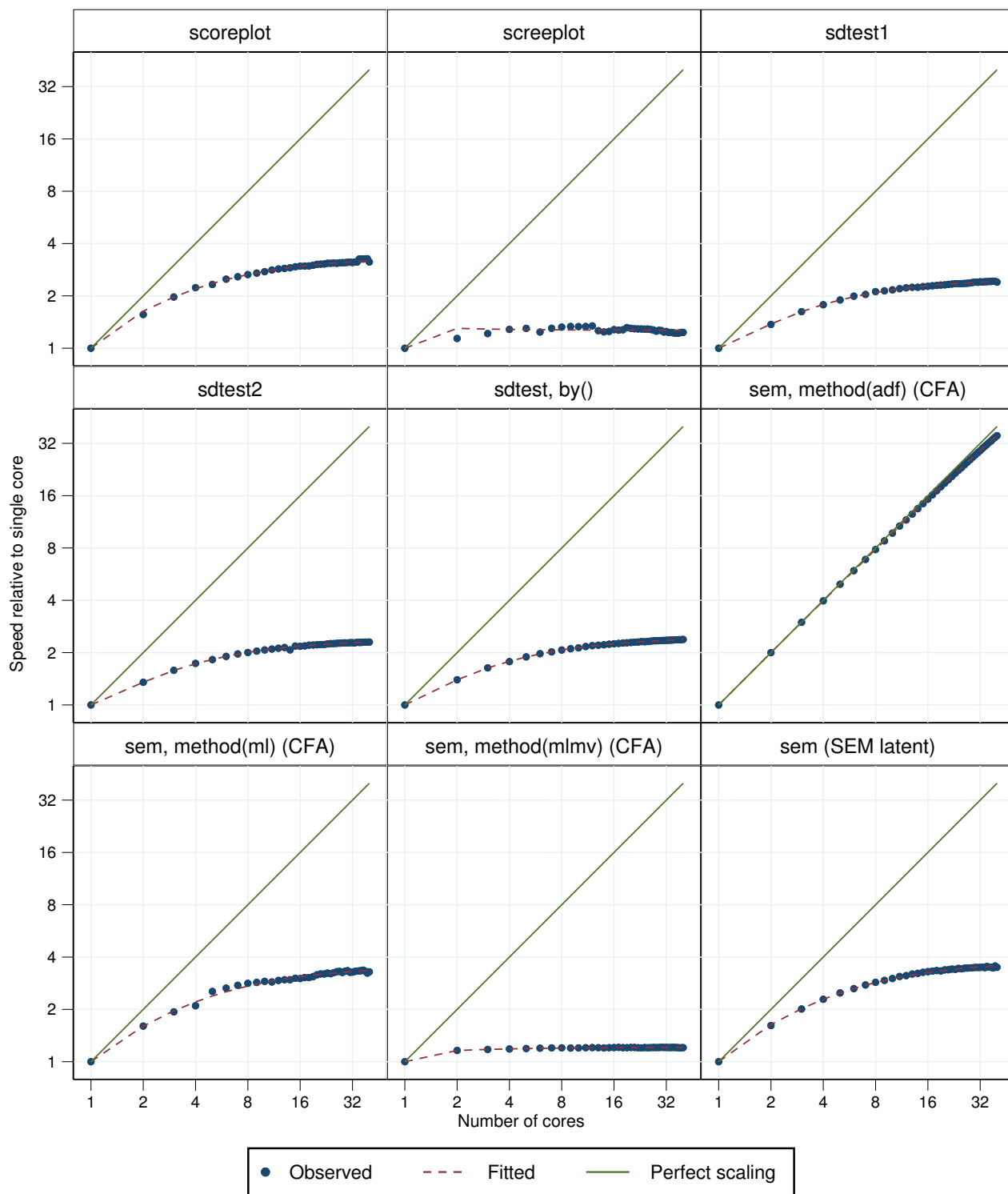


Figure 652. Parallelization performance plots.

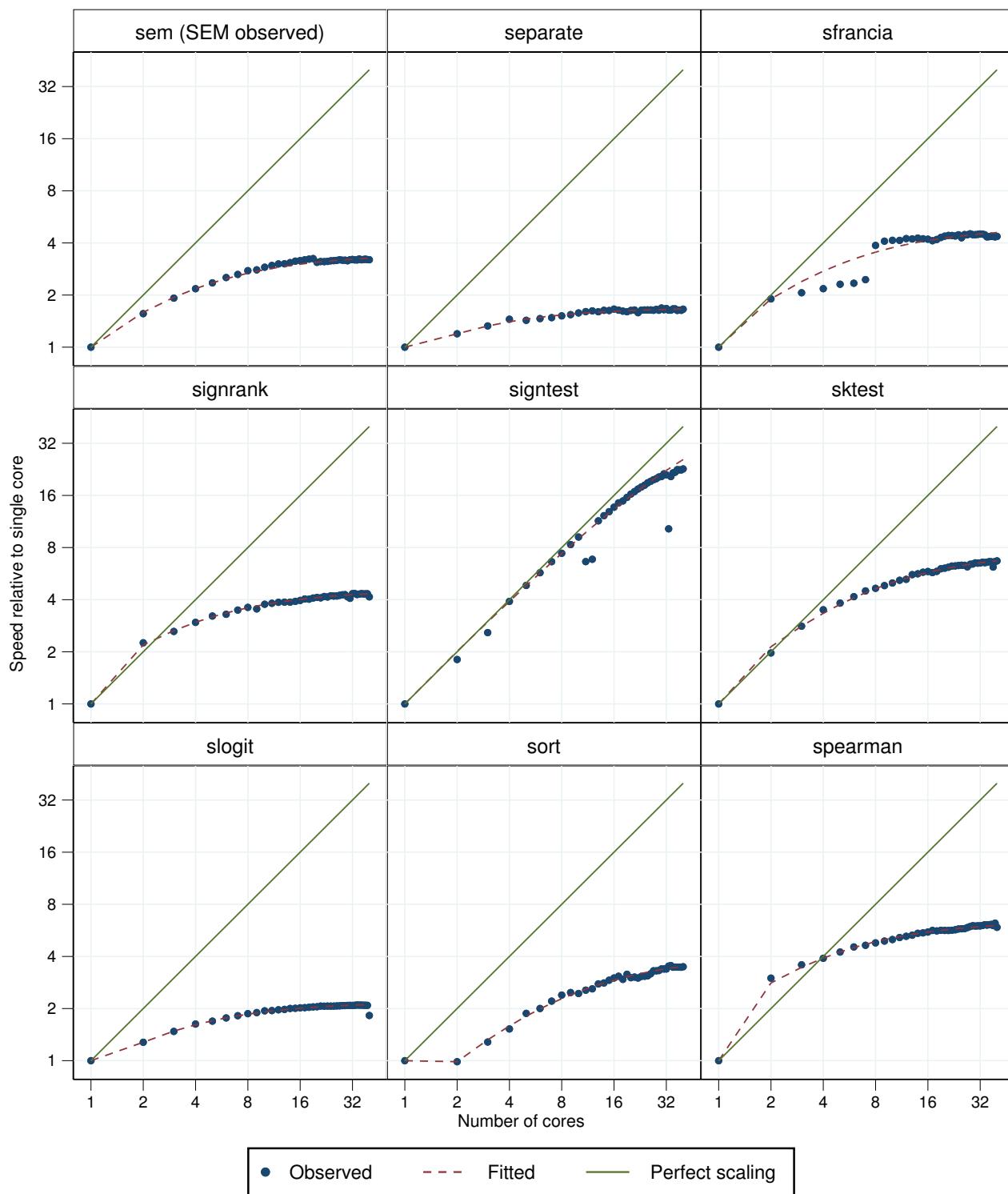


Figure 653. Parallelization performance plots.

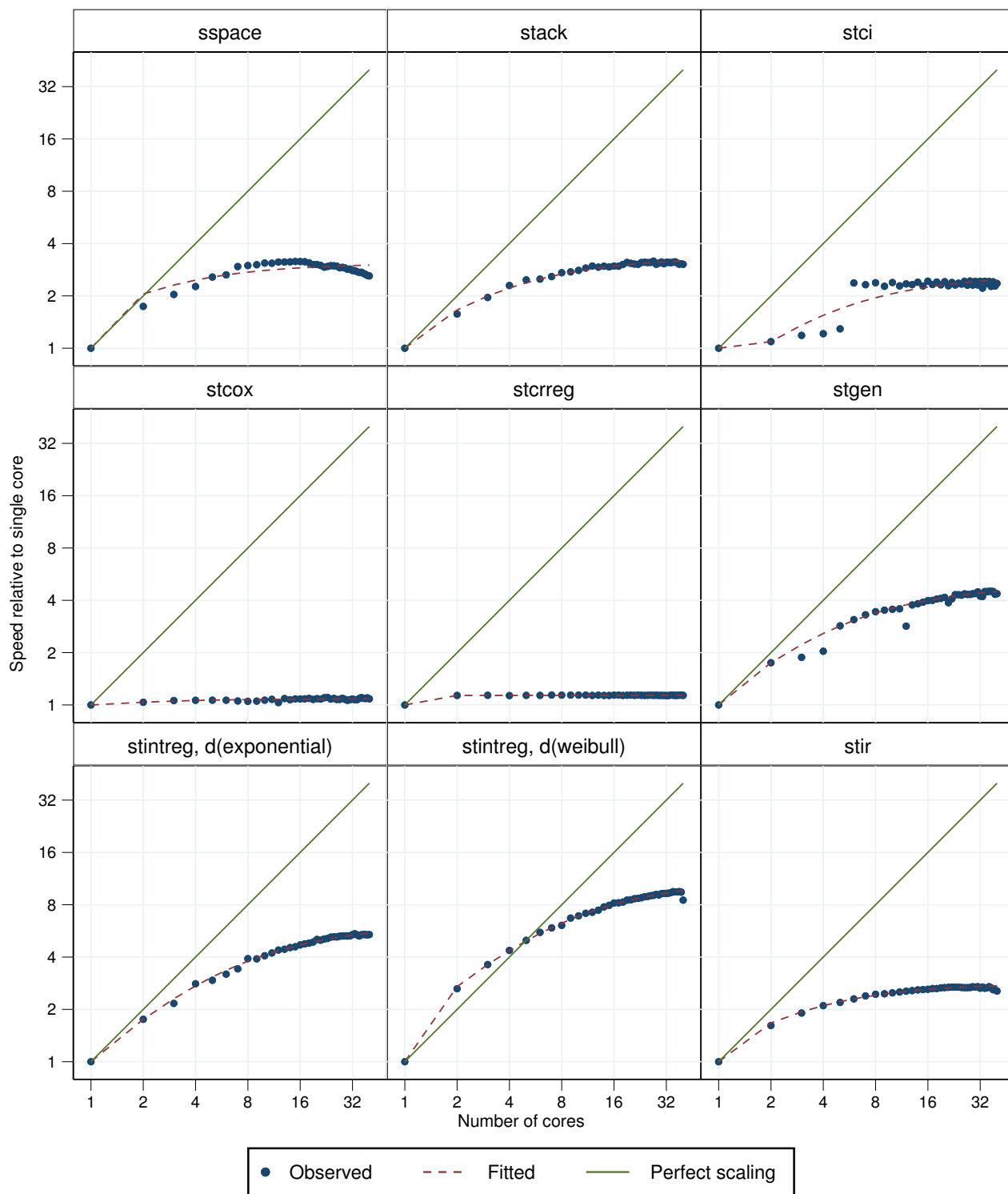


Figure 654. Parallelization performance plots.



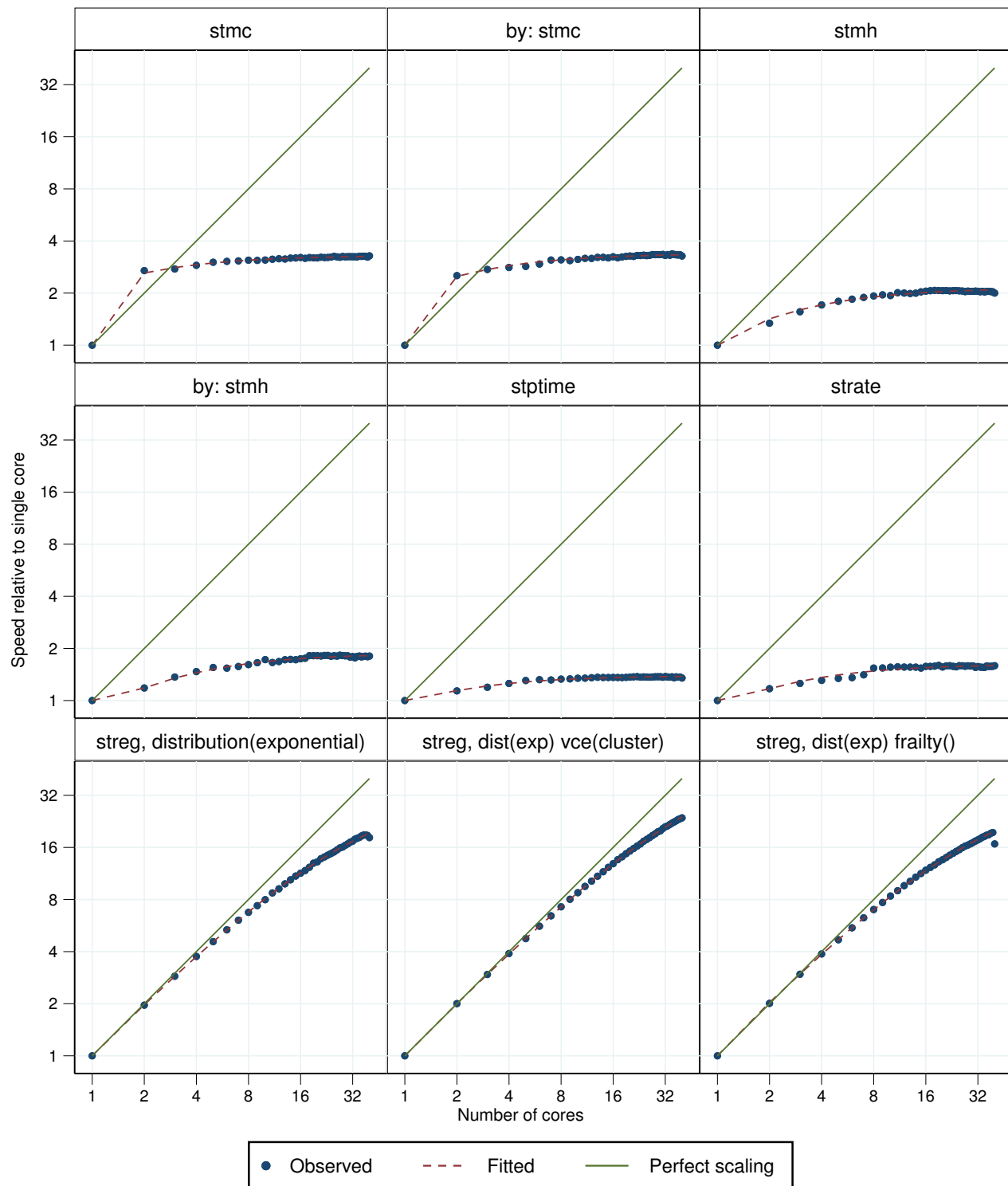


Figure 655. Parallelization performance plots.

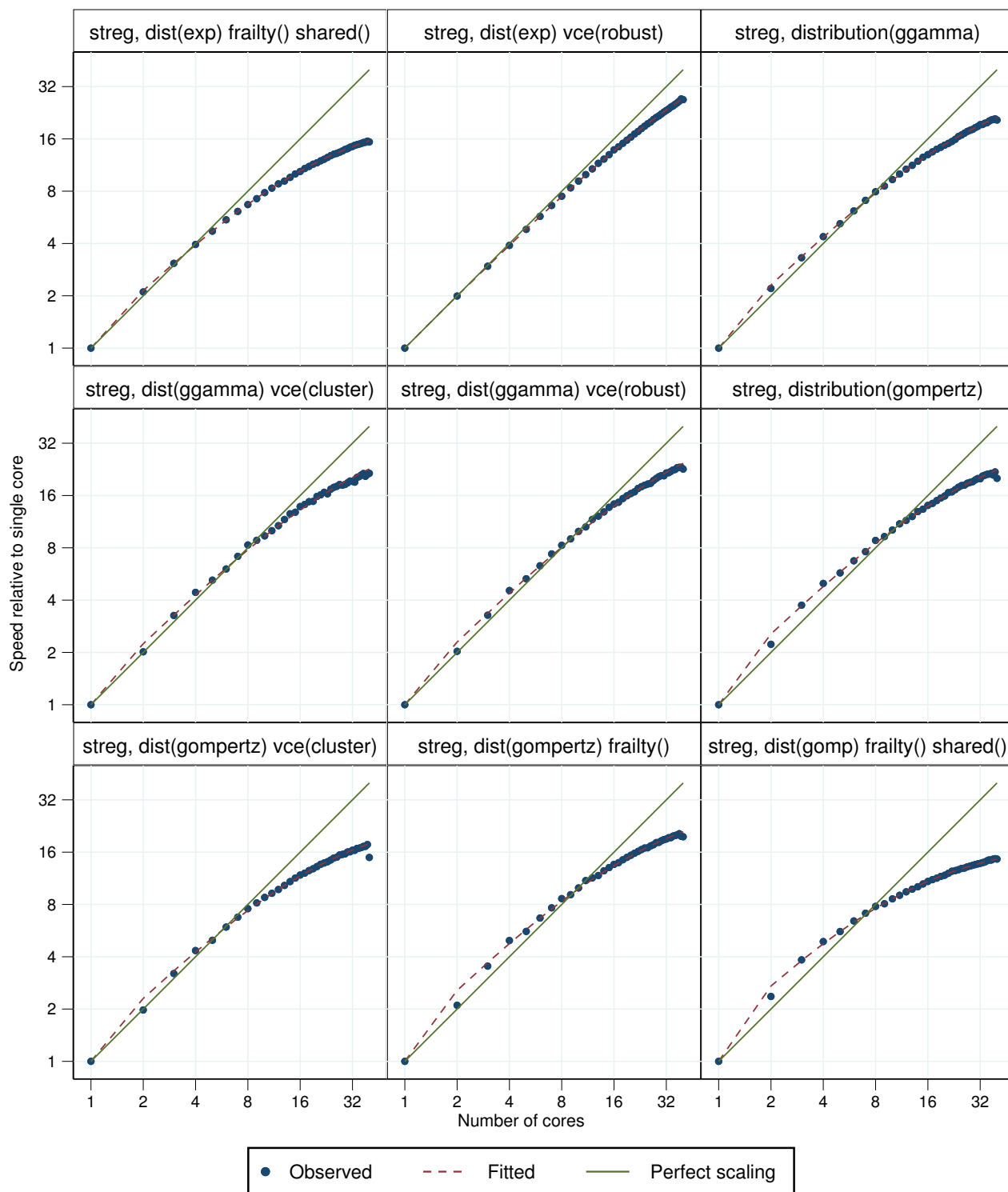


Figure 656. Parallelization performance plots.

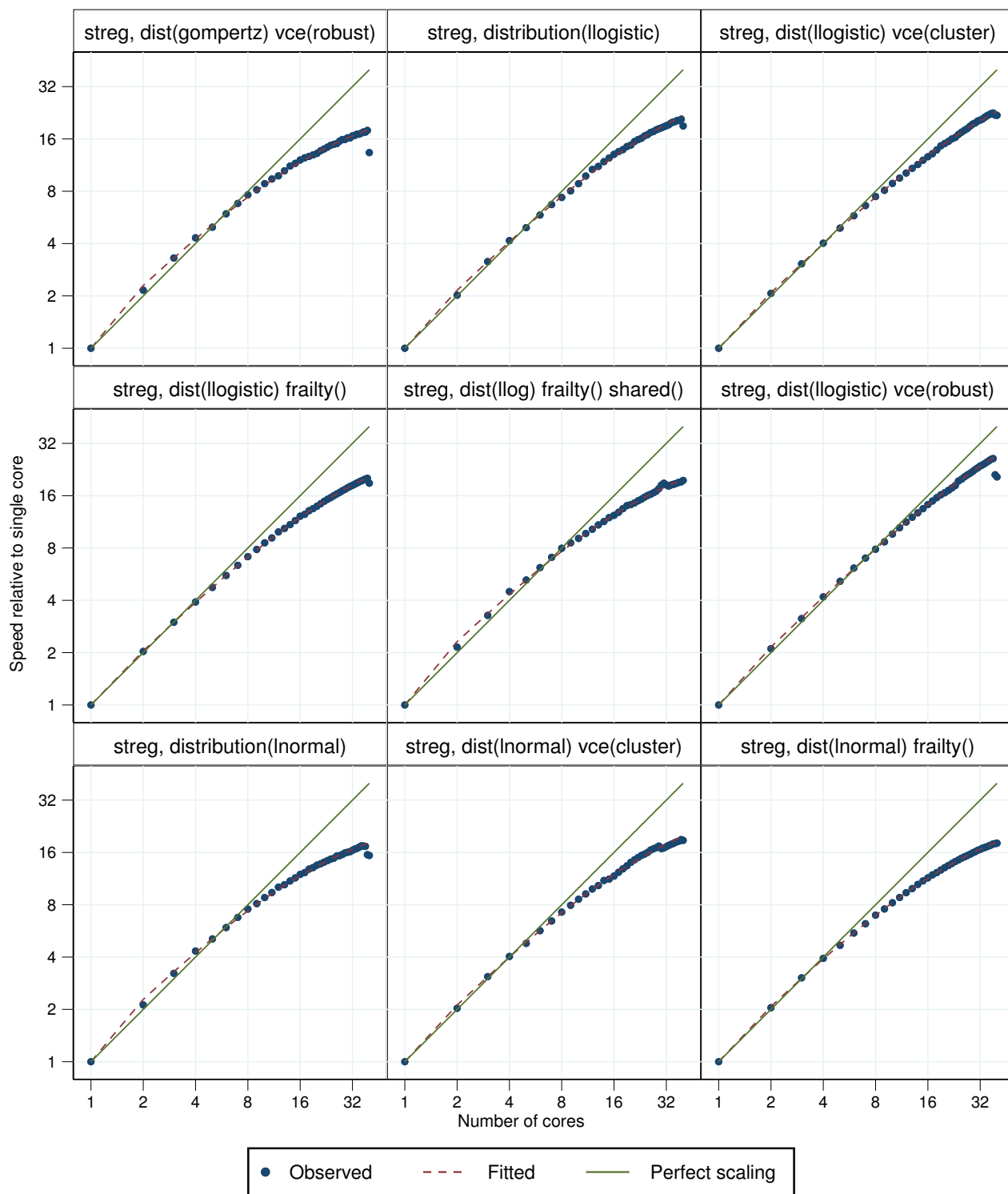


Figure 657. Parallelization performance plots.

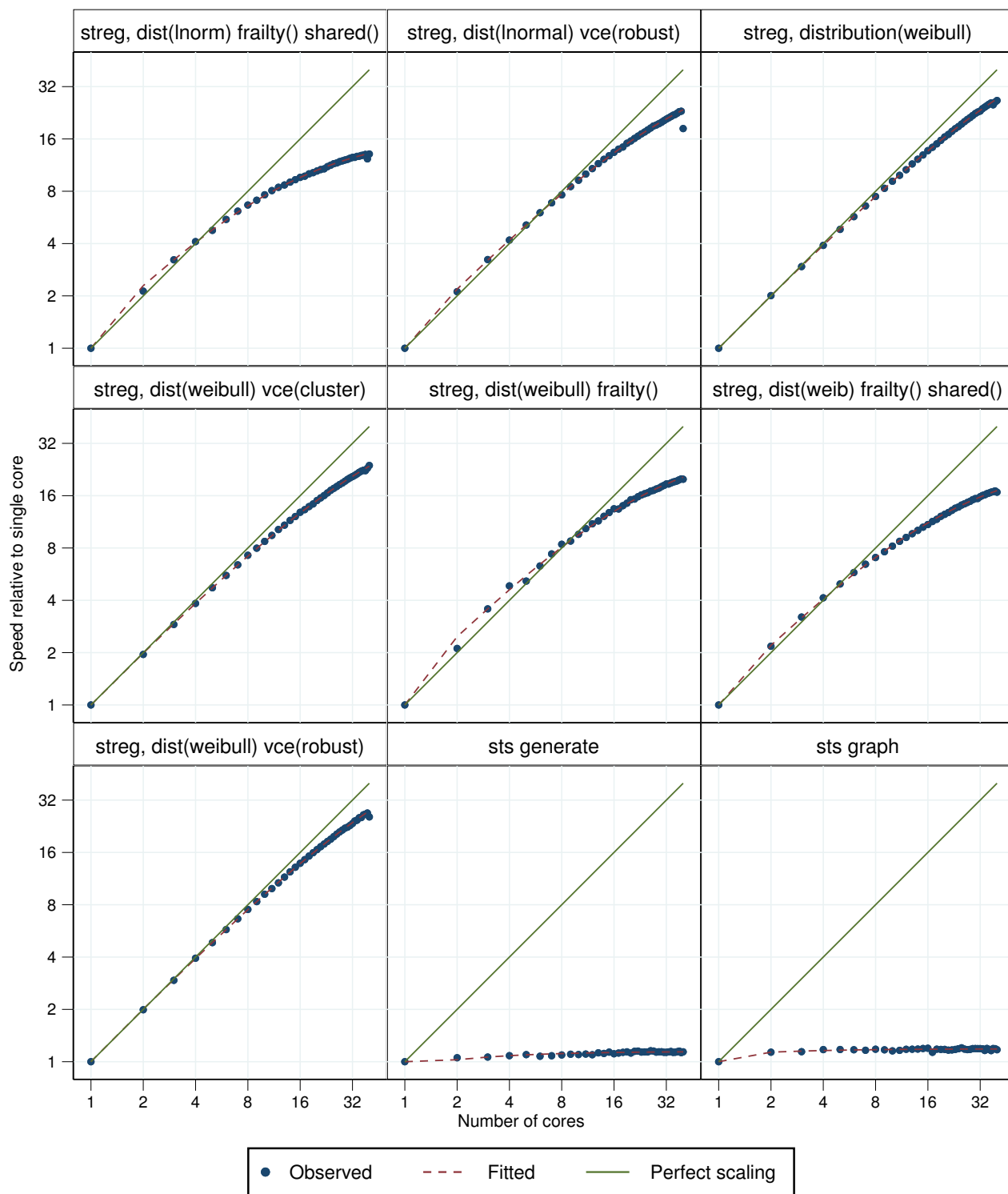


Figure 658. Parallelization performance plots.

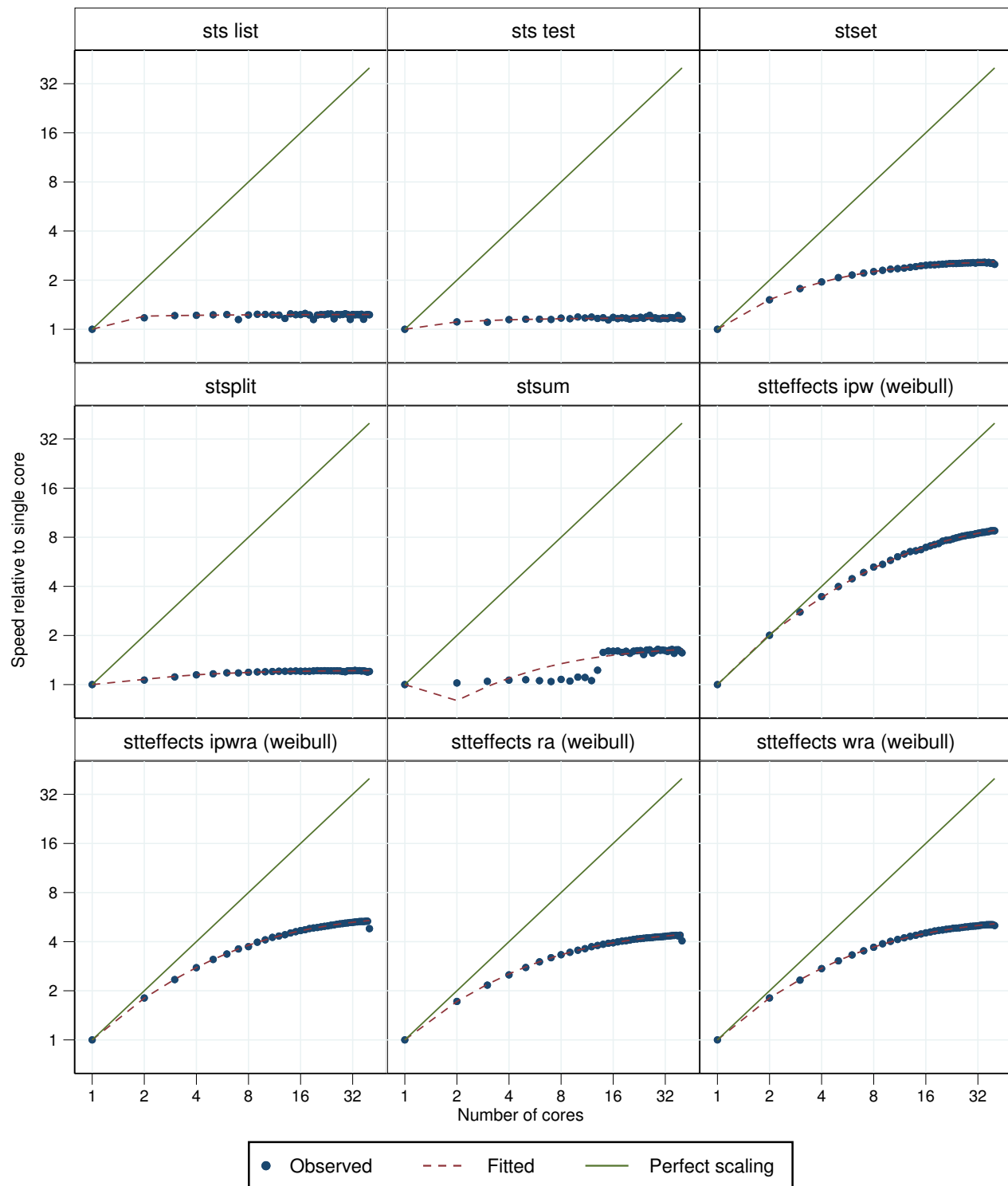


Figure 659. Parallelization performance plots.

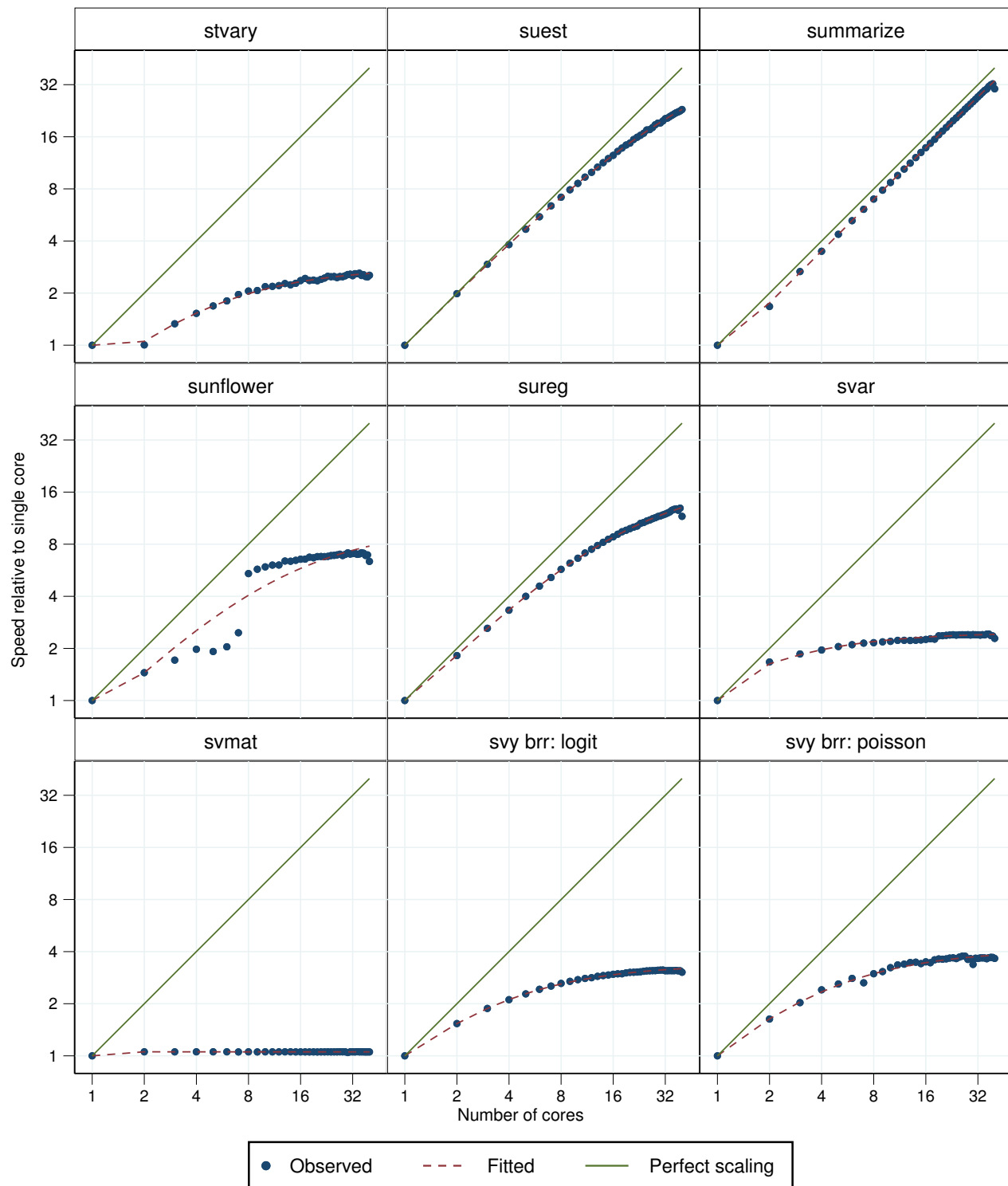


Figure 660. Parallelization performance plots.

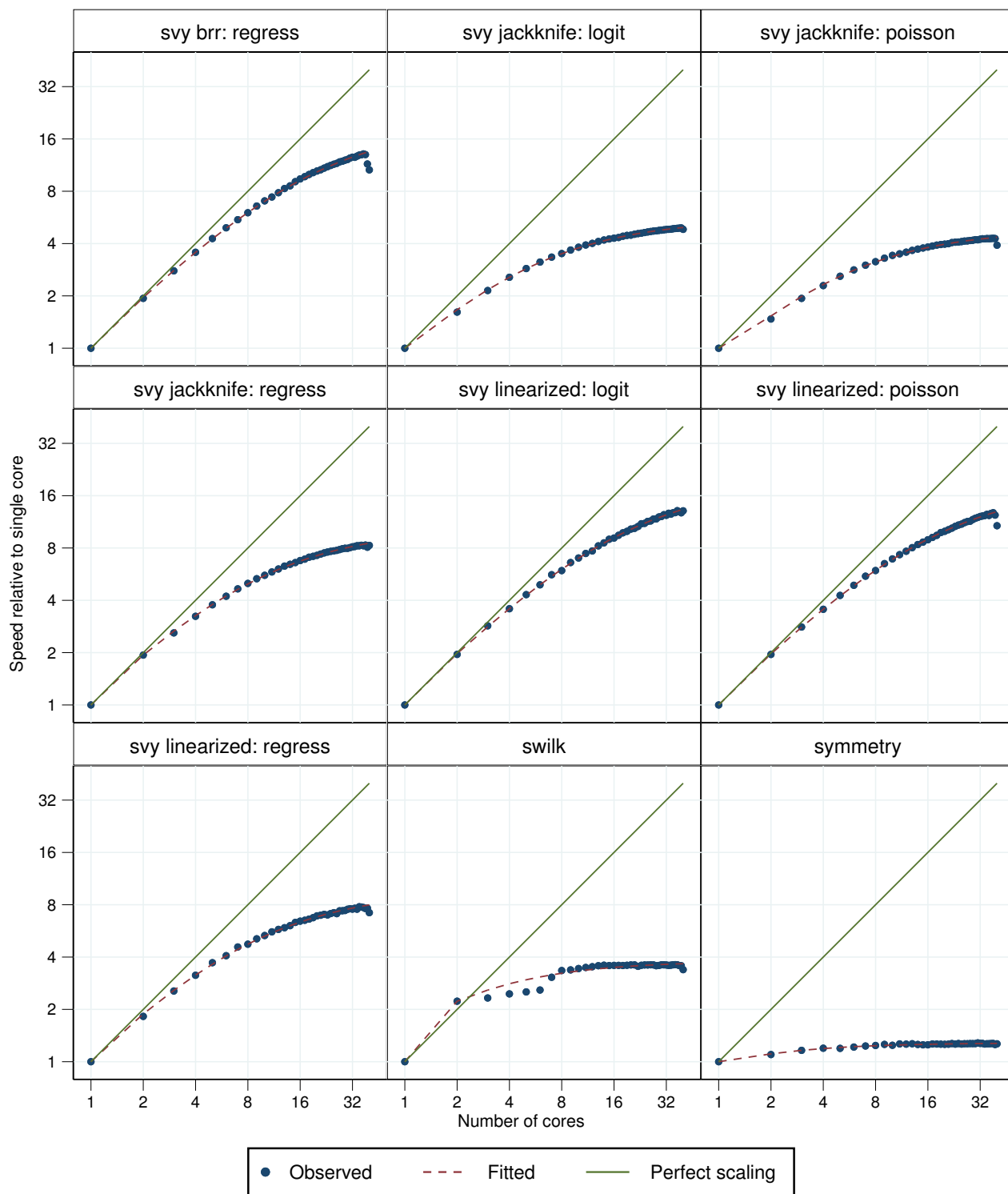


Figure 661. Parallelization performance plots.

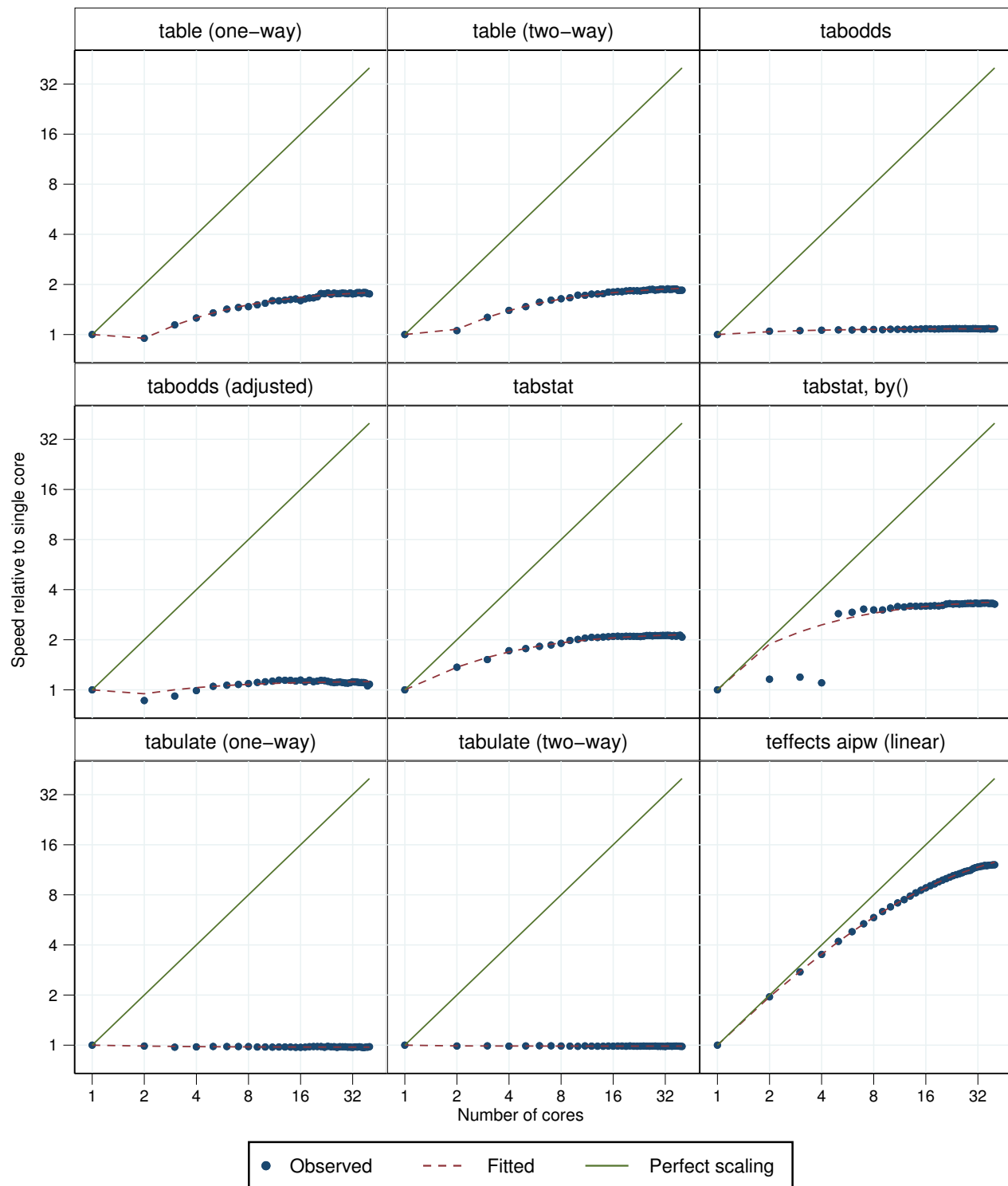


Figure 662. Parallelization performance plots.



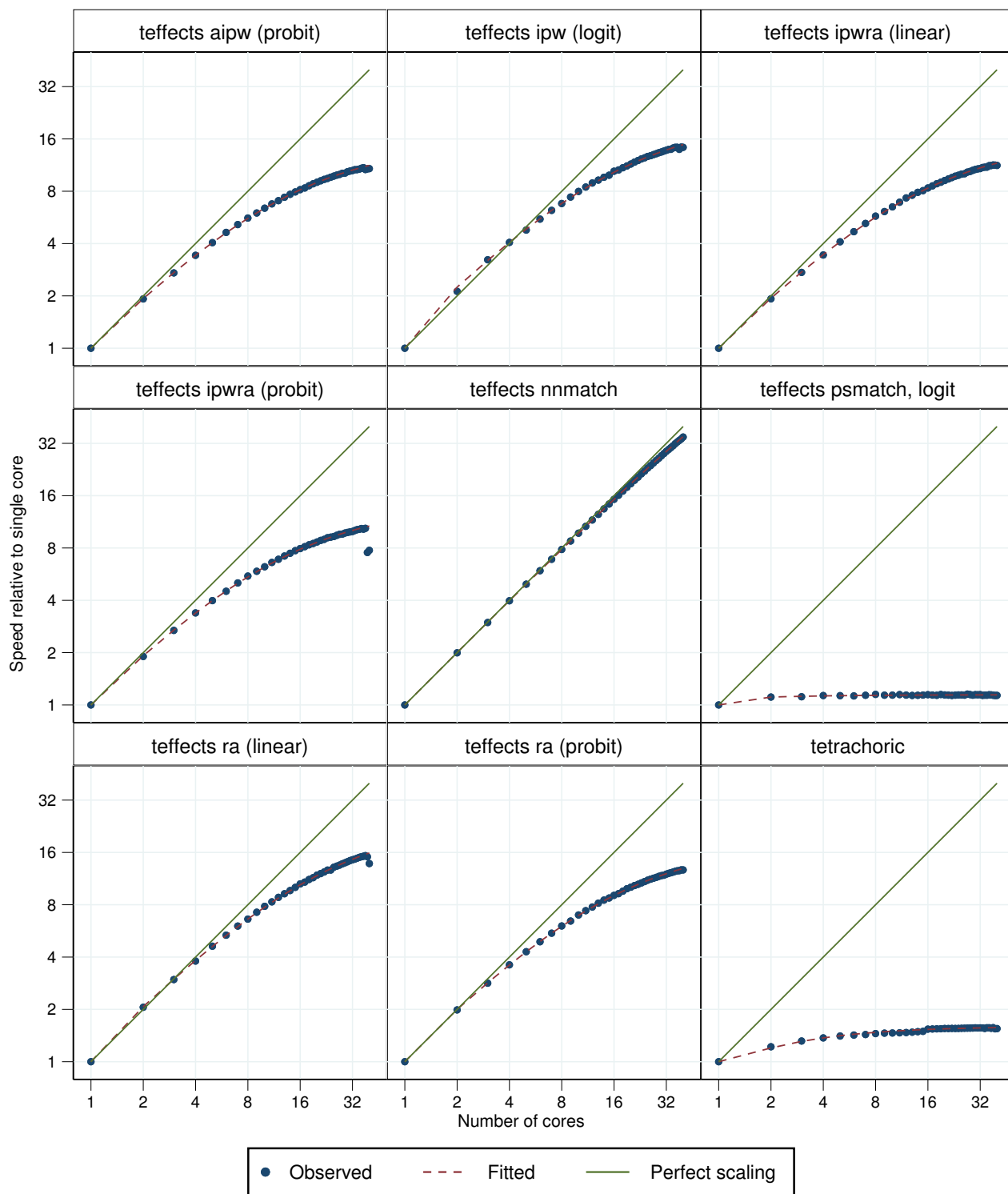


Figure 663. Parallelization performance plots.

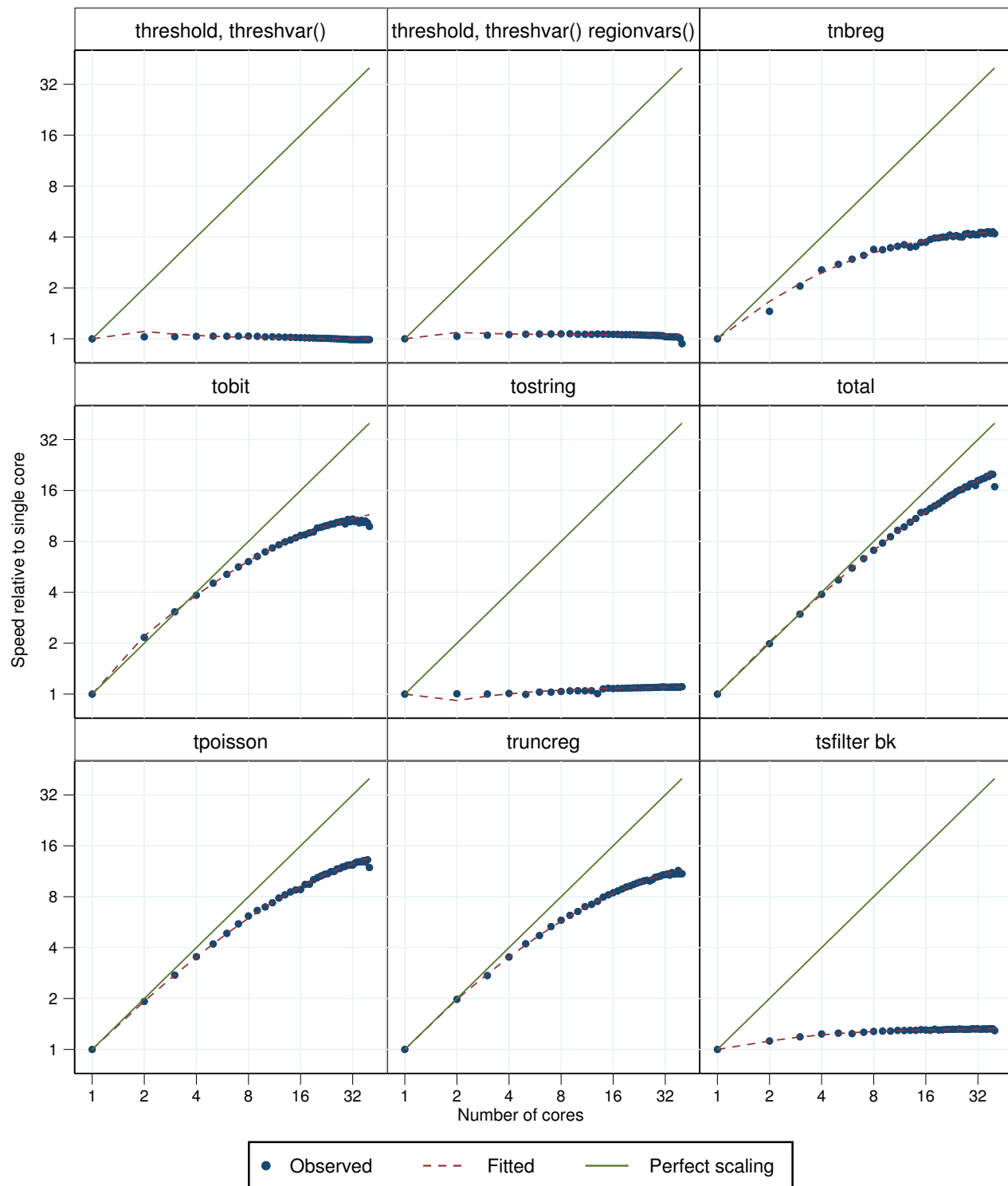


Figure 664. Parallelization performance plots.

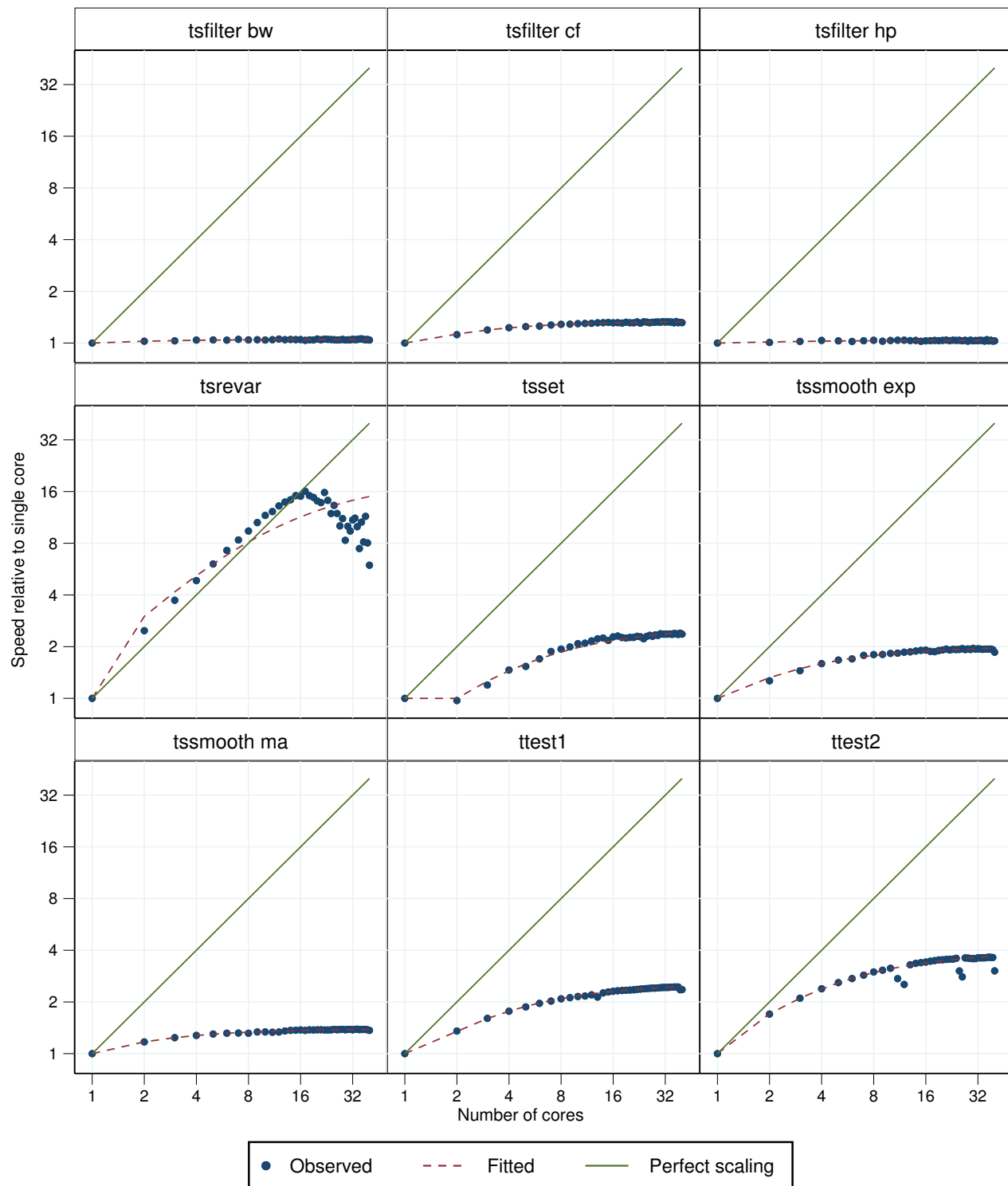


Figure 665. Parallelization performance plots.

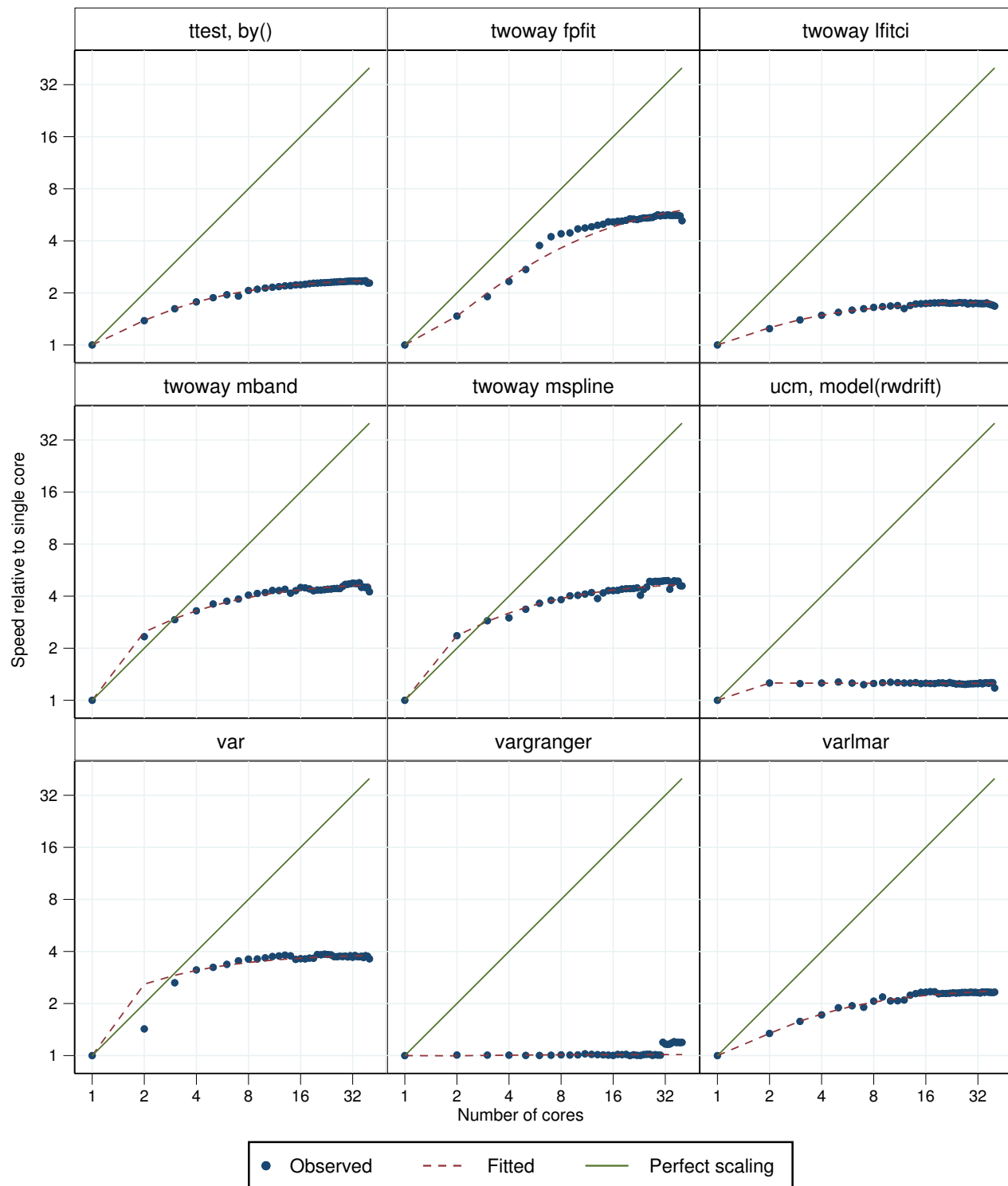


Figure 666. Parallelization performance plots.

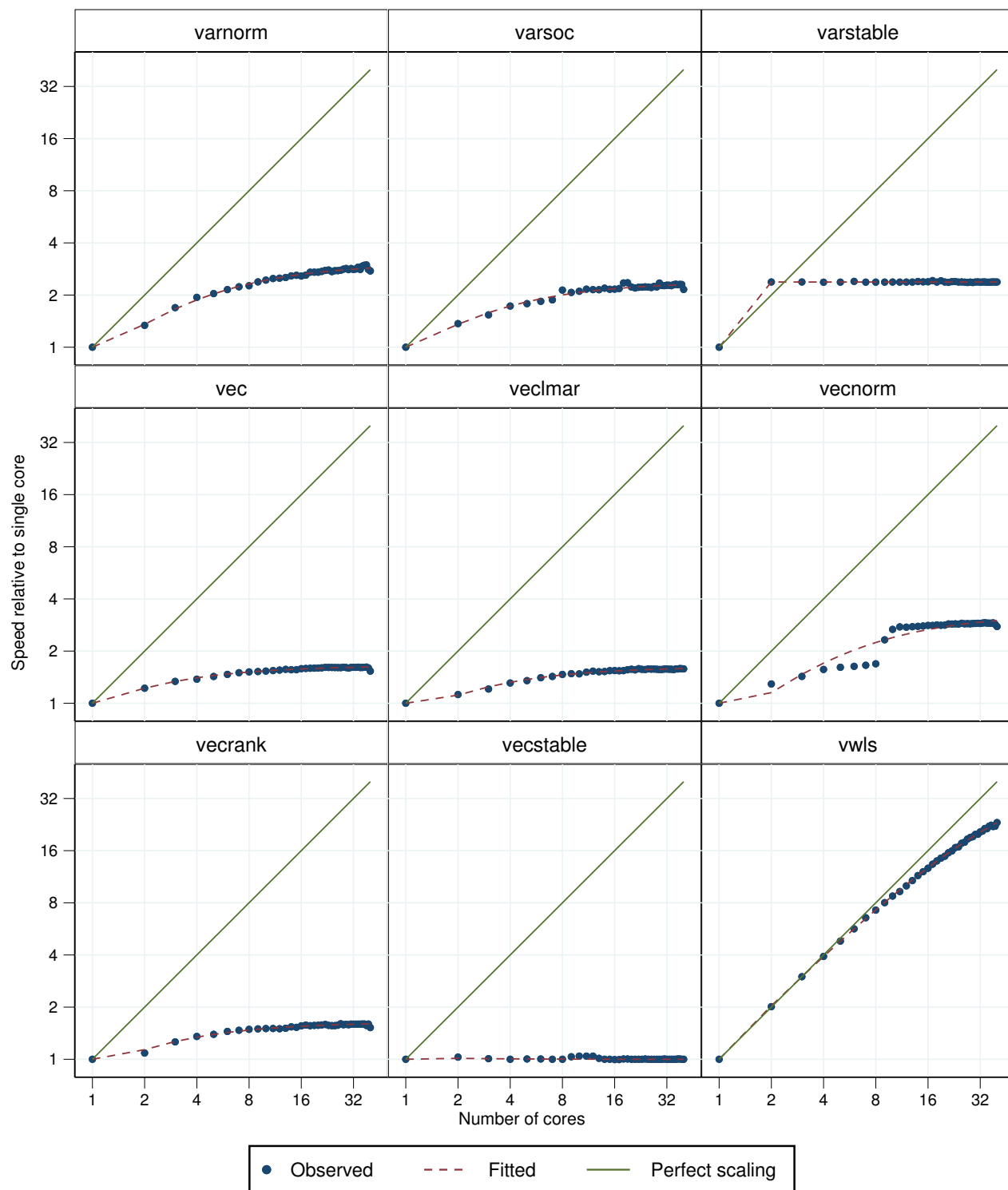


Figure 667. Parallelization performance plots.

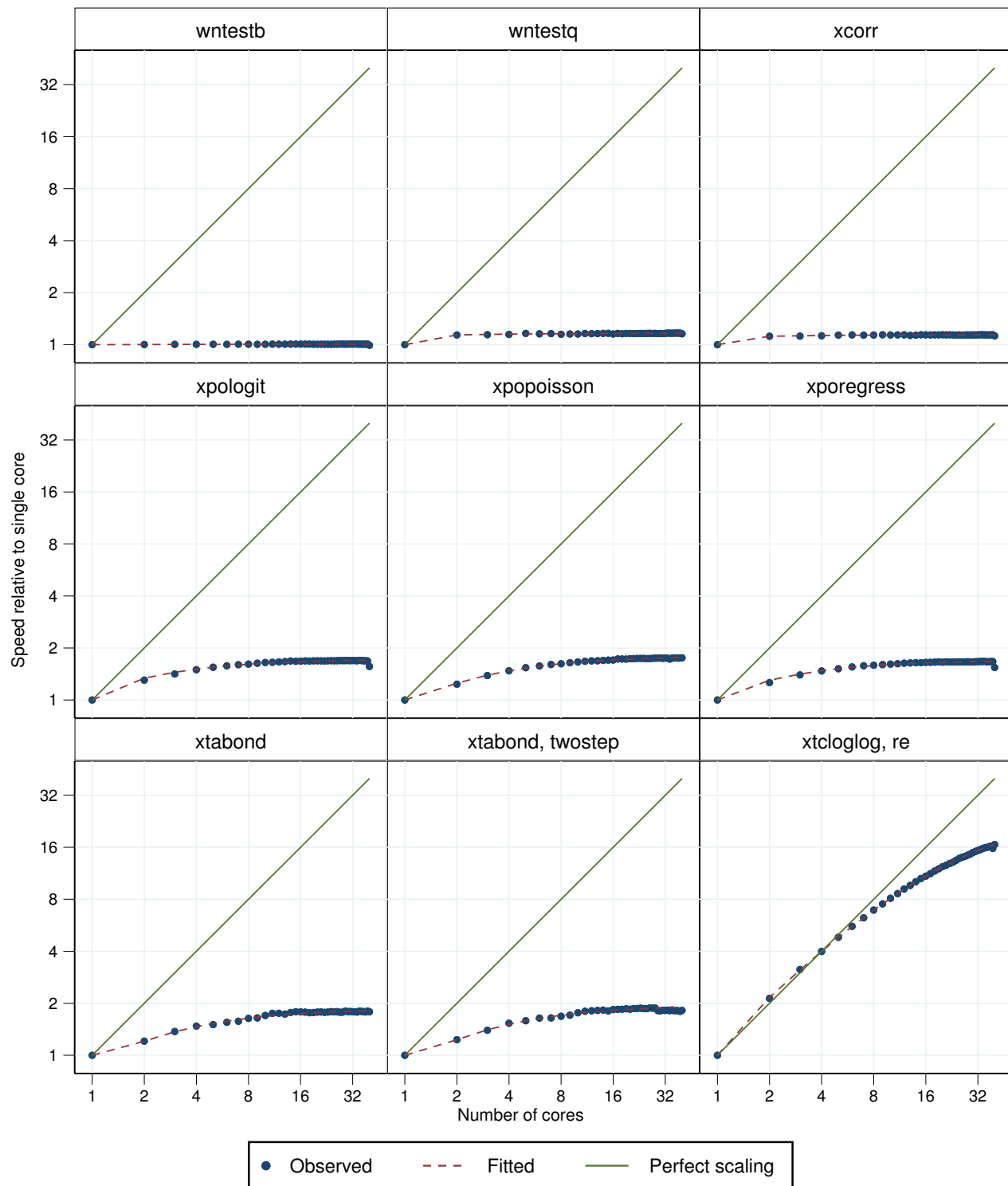


Figure 668. Parallelization performance plots.

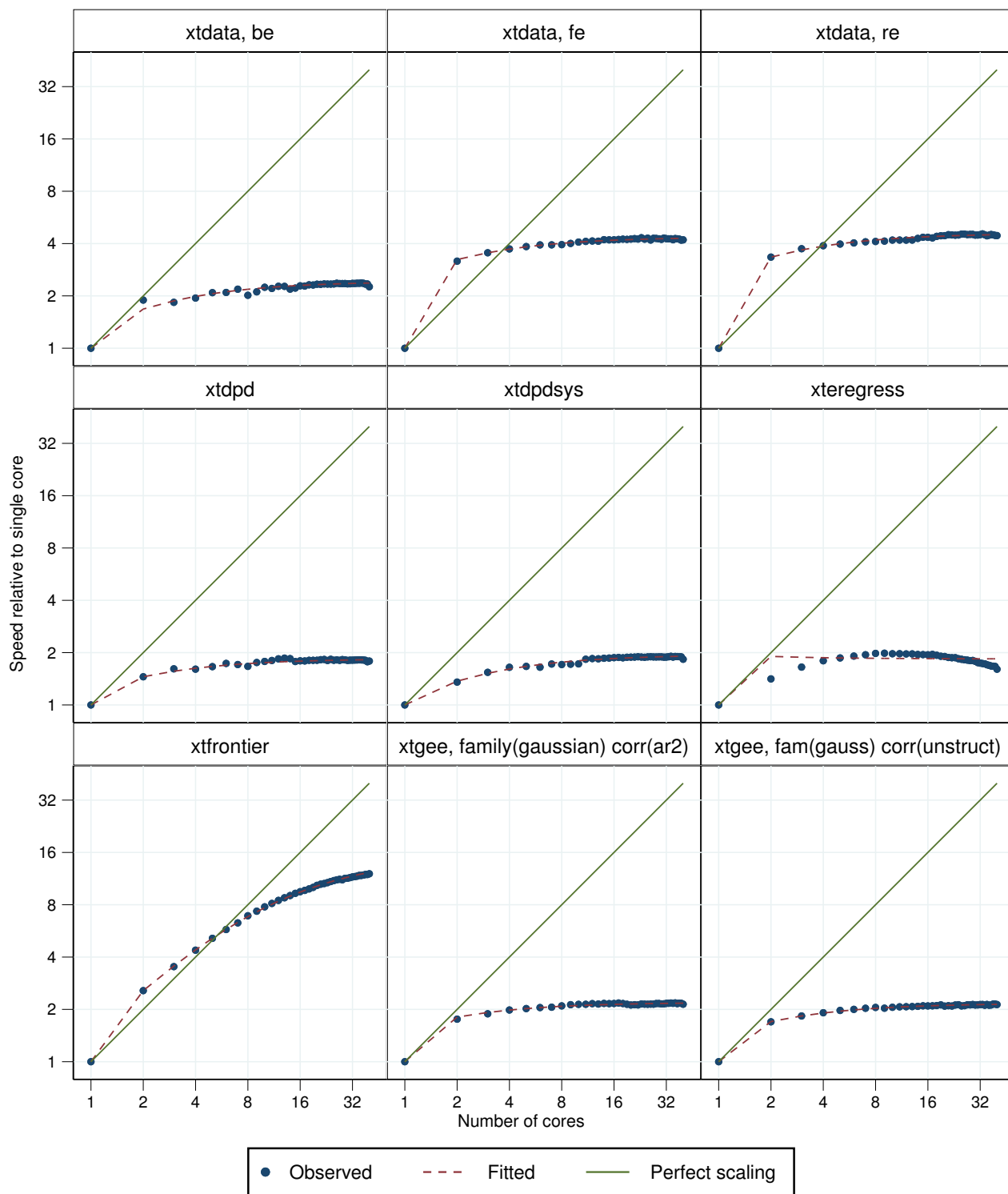


Figure 669. Parallelization performance plots.

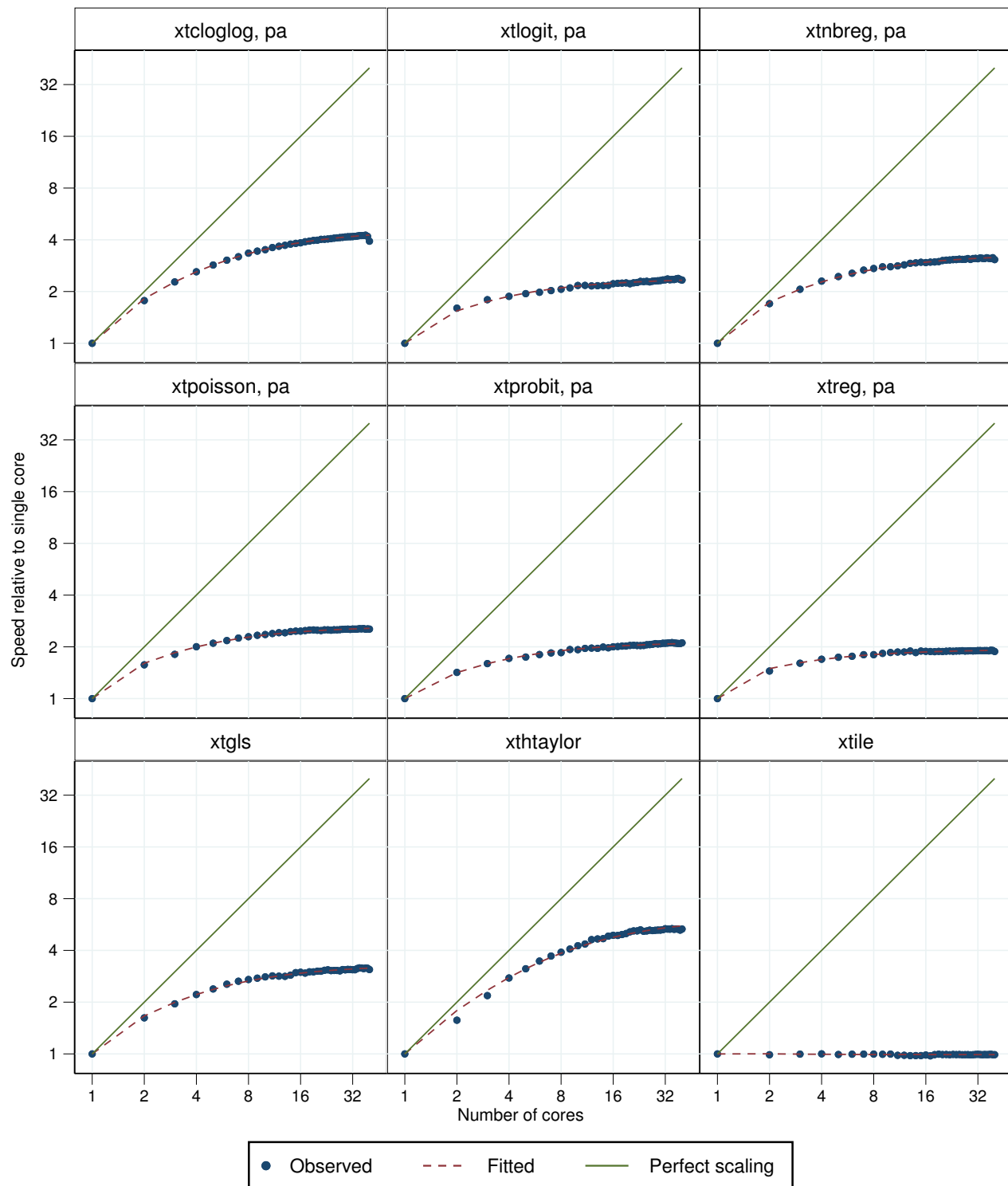


Figure 670. Parallelization performance plots.



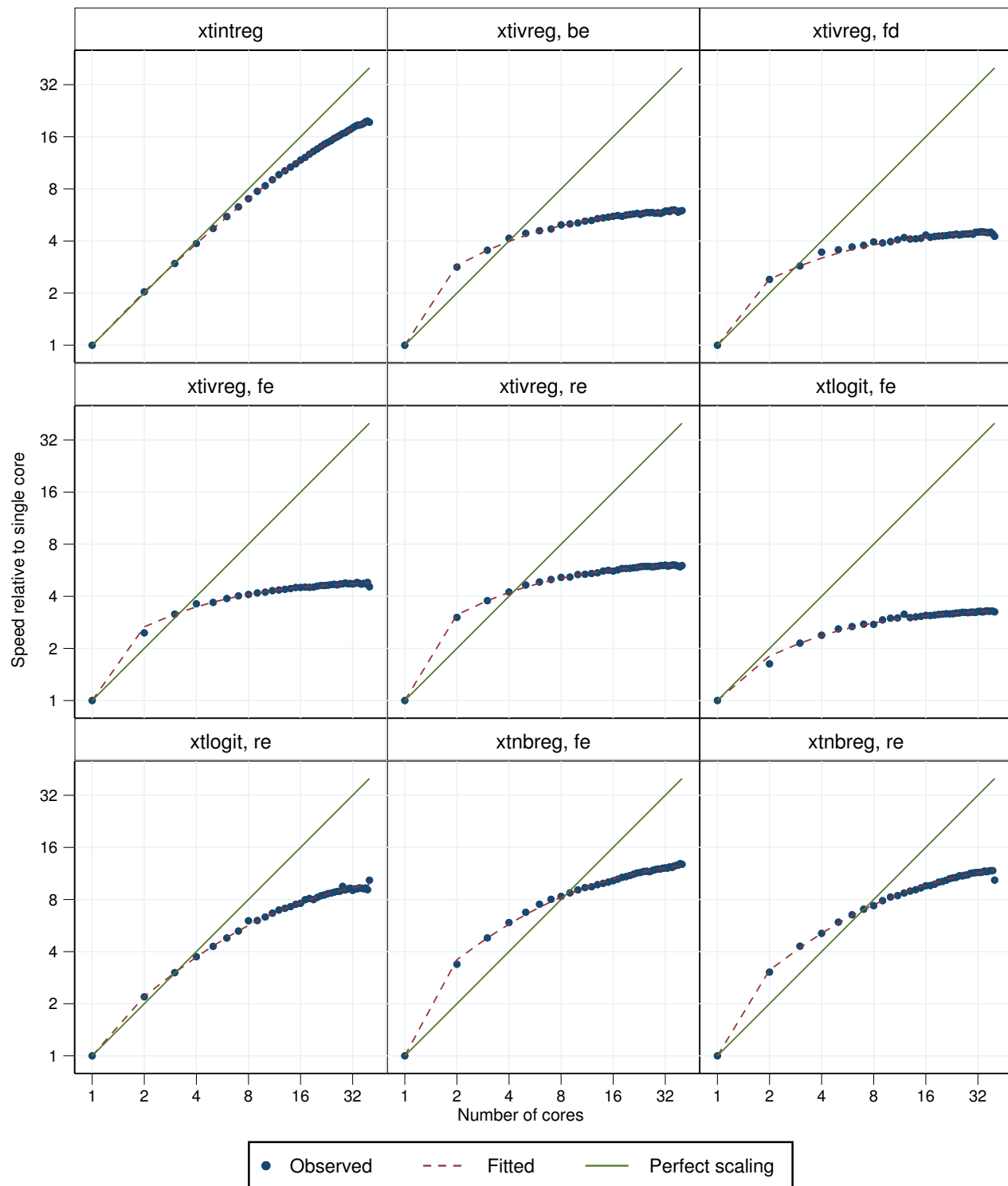


Figure 671. Parallelization performance plots.

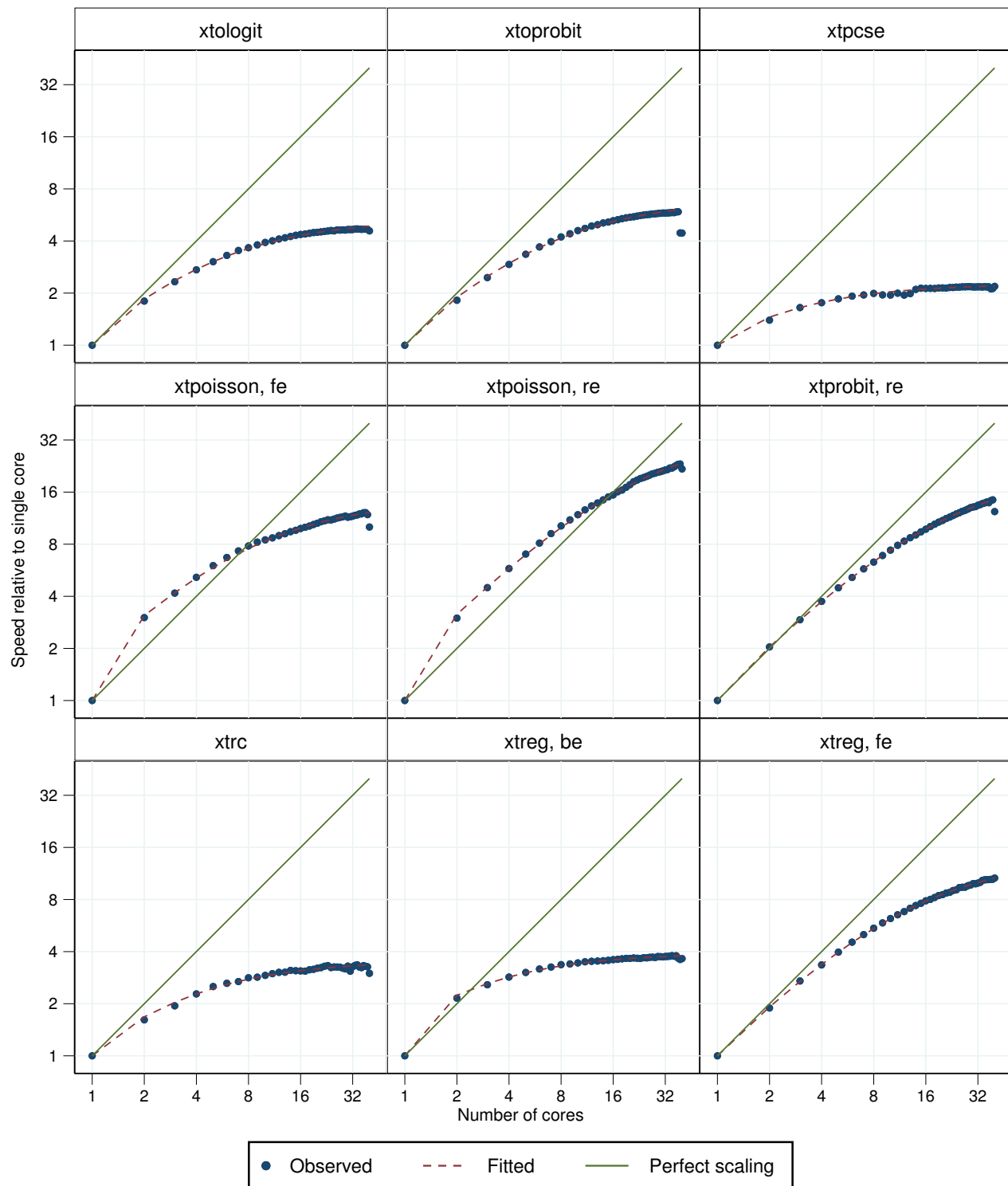


Figure 672. Parallelization performance plots.

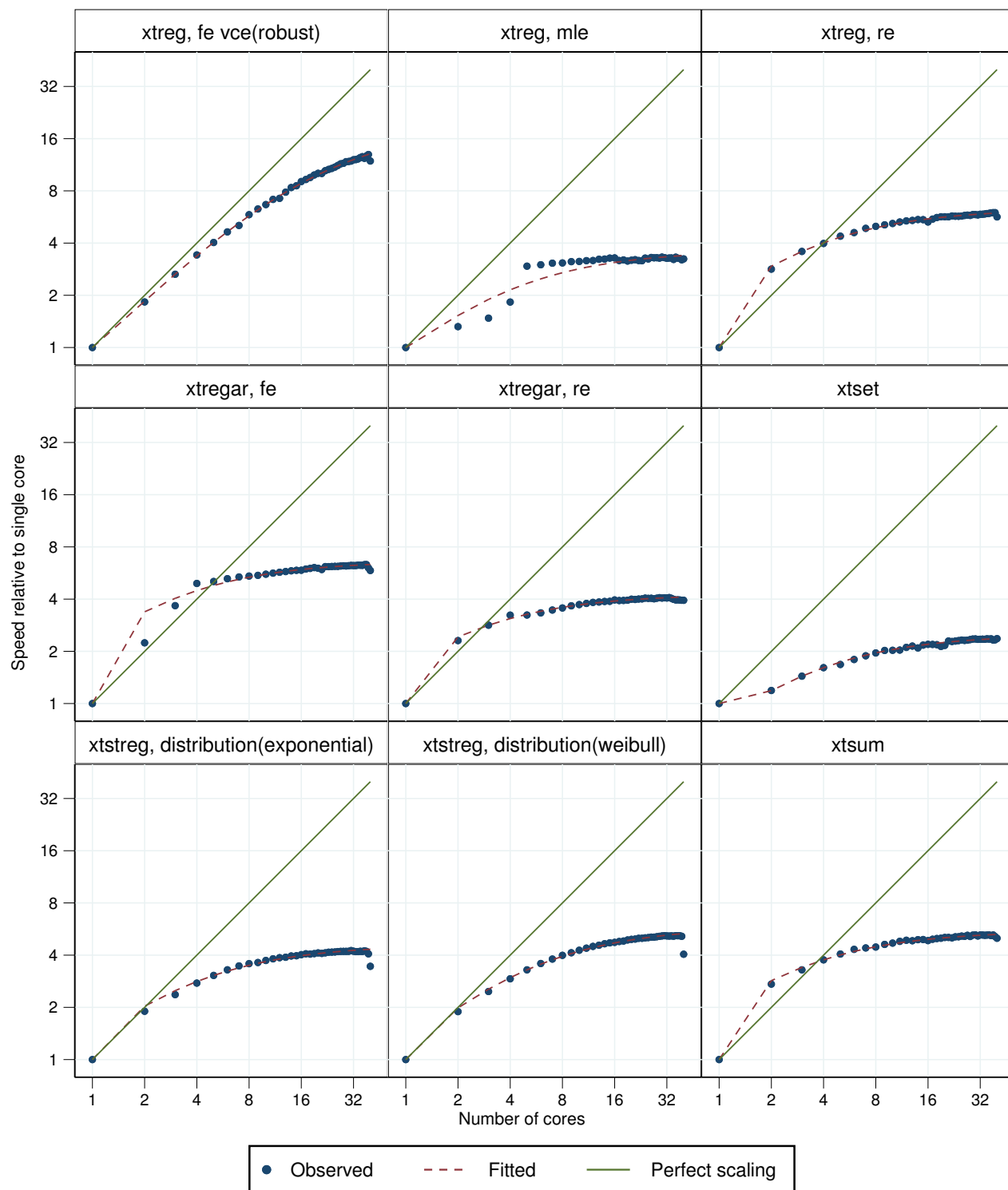


Figure 673. Parallelization performance plots.

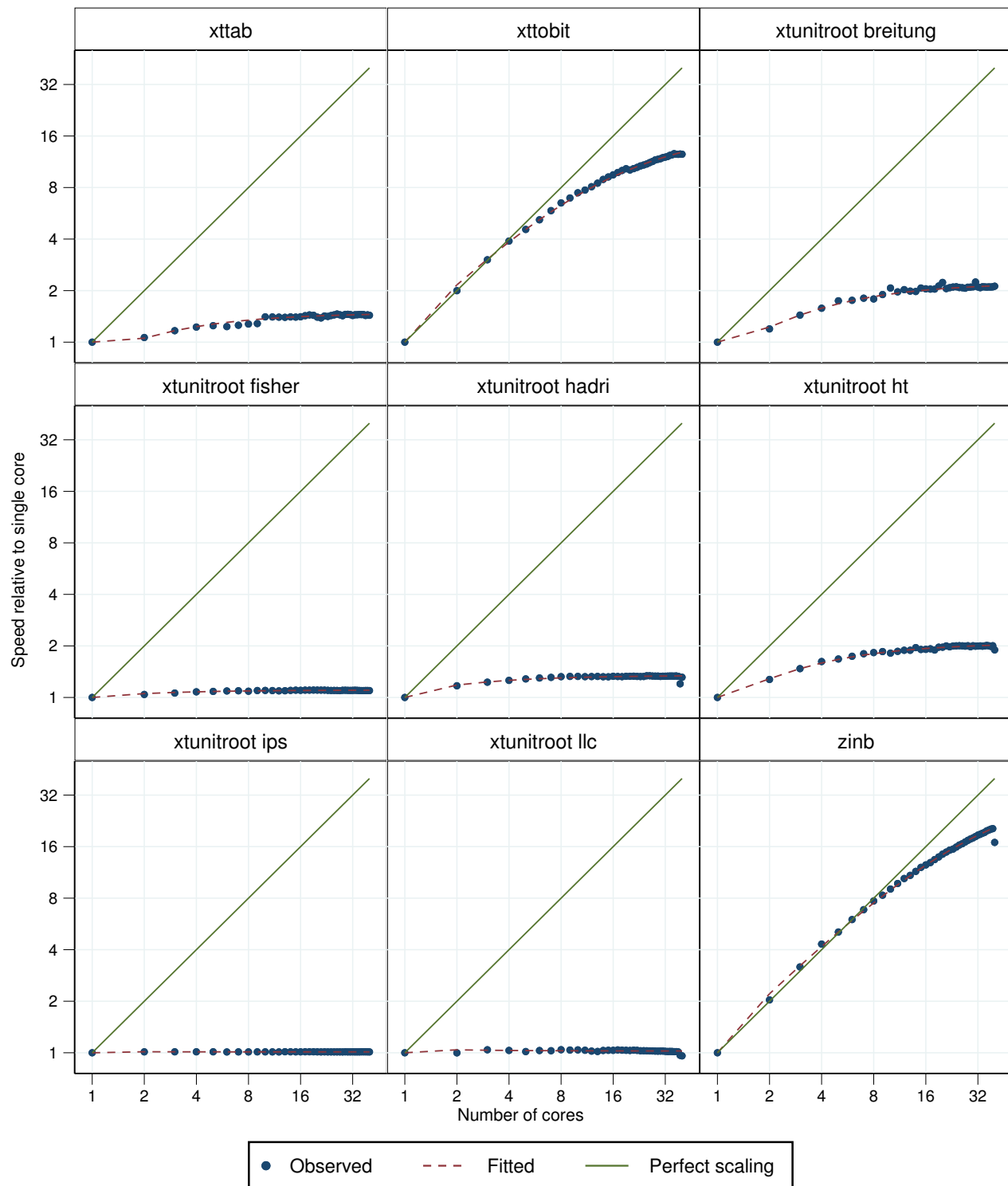


Figure 674. Parallelization performance plots.

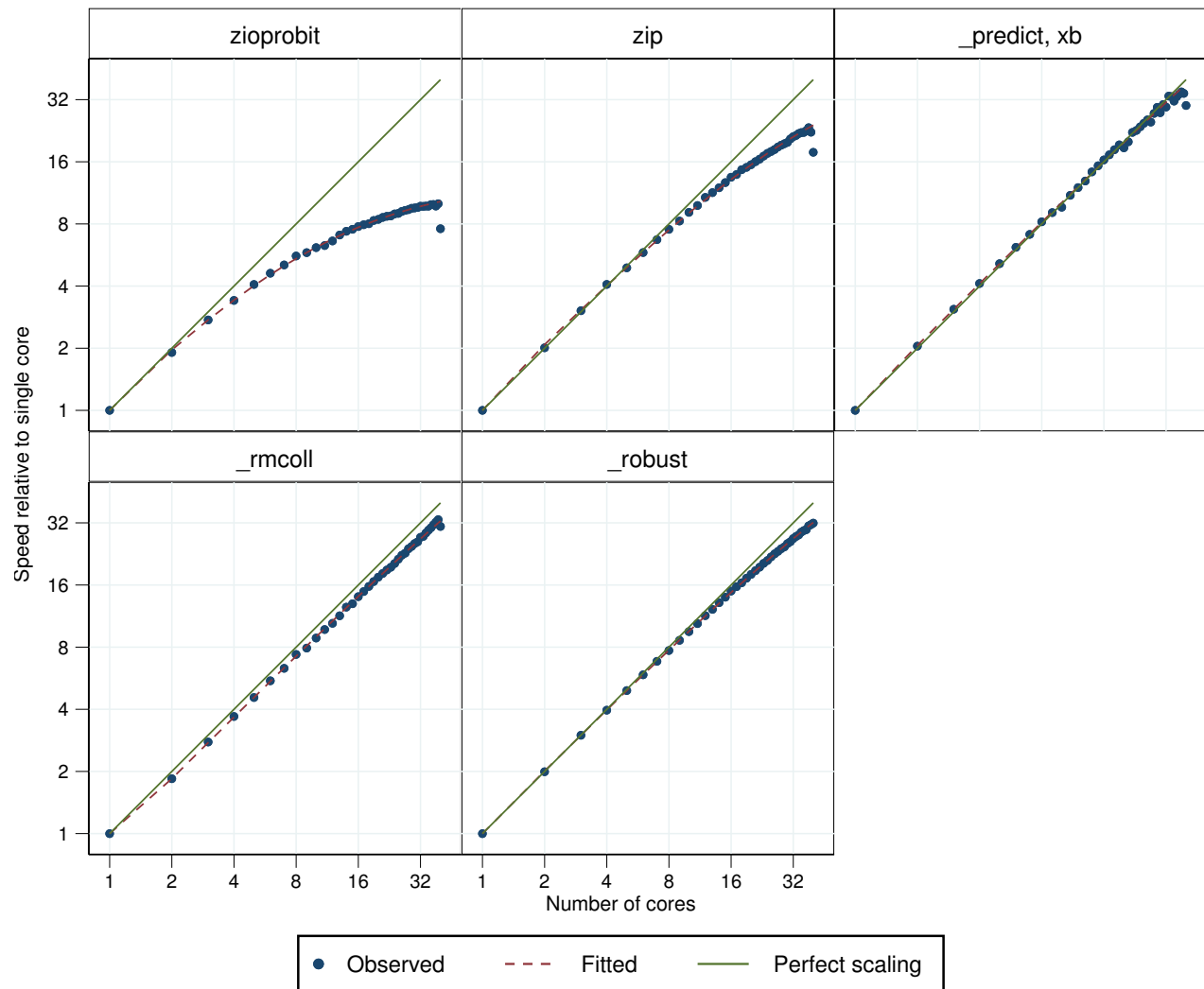


Figure 675. Parallelization performance plots.

## C Command names and descriptions

Table 2. Command descriptions

| Command                            | Description                                                                             |
|------------------------------------|-----------------------------------------------------------------------------------------|
| <code>alpha</code>                 | Cronbach's alpha                                                                        |
| <code>ameans</code>                | Arithmetic, geometric, and harmonic means                                               |
| <code>anova</code> (one-way)       | Analysis of variance and covariance—one-way                                             |
| <code>anova</code> (two-way)       | Analysis of variance and covariance—two-way                                             |
| <code>arch</code>                  | Autoregressive conditional heteroskedasticity (ARCH) family of estimators               |
| <code>areg</code>                  | Linear regression with a large dummy-variable set                                       |
| <code>areg, vce(cluster)</code>    | Linear regression with a large dummy-variable set, cluster-robust standard errors       |
| <code>areg, vce(robust)</code>     | Linear regression with a large dummy-variable set, robust (Huber/White) standard errors |
| <code>arfima</code>                | Autoregressive fractionally integrated moving-average models                            |
| <code>arima</code>                 | ARIMA, ARMAX, and other dynamic regression models                                       |
| <code>bayes: logit</code>          | Bayesian logistic regression                                                            |
| <code>bayes: poisson</code>        | Bayesian Poisson regression                                                             |
| <code>bayes: regress</code>        | Bayesian linear regression                                                              |
| <code>bayesmh logit</code>         | Bayesian logistic regression using Metropolis-Hastings algorithm                        |
| <code>bayesmh mvn</code>           | Bayesian multivariate normal regression using Metropolis-Hastings algorithm             |
| <code>bayesmh mylogit</code>       | Bayesian logistic regression using Metropolis-Hastings algorithm (custom evaluator)     |
| <code>bayesmh nl</code>            | Bayesian nonlinear regression using Metropolis-Hastings algorithm                       |
| <code>bayesmh normal</code>        | Bayesian linear regression using Metropolis-Hastings algorithm                          |
| <code>bayesmh normal gibbs</code>  | Bayesian linear regression using Gibbs sampling                                         |
| <code>bayesmh normal re</code>     | Bayesian linear regression with random effects using Metropolis-Hastings algorithm      |
| <code>betareg, link(logit)</code>  | Beta regression, logit link                                                             |
| <code>betareg, link(probit)</code> | Beta regression, probit link                                                            |
| <code>binreg</code>                | Generalized linear models: extensions to the binomial family                            |
| <code>biplot</code>                | Biplots                                                                                 |
| <code>biprobit</code>              | Bivariate probit regression                                                             |

Table 2. Command descriptions

| Command                                     | Description                                                        |
|---------------------------------------------|--------------------------------------------------------------------|
| <code>biprobit</code> (seemingly unrelated) | Seemingly unrelated probit regression                              |
| <code>bittest</code>                        | Binomial probability test                                          |
| <code>blogit</code>                         | Logistic regression for grouped data                               |
| <code>boxcox</code>                         | Box–Cox regression models                                          |
| <code>bprobit</code>                        | Probit regression for grouped data                                 |
| <code>brier</code>                          | Brier score decomposition                                          |
| <code>bsample</code>                        | Sampling with replacement                                          |
| <code>bstat</code>                          | Compute and report bootstrap statistics                            |
| <code>by: generate</code>                   | Create new variables over longitudinal/panel data                  |
| <code>by: generate</code> (small groups)    | Create new variables over longitudinal/panel data, small panels    |
| <code>by: replace</code>                    | Replace variable values over longitudinal/panel data               |
| <code>by: replace</code> (small groups)     | Replace variable values over longitudinal/panel data, small panels |
| <code>ca</code>                             | Simple correspondence analysis                                     |
| <code>candisc</code>                        | Canonical linear discriminant analysis                             |
| <code>canon</code>                          | Canonical correlations                                             |
| <code>cc</code>                             | Case–control odds ratio                                            |
| <code>by: cc</code>                         | Case–control odds ratio over groups                                |
| <code>centile</code>                        | Report centile and confidence interval                             |
| <code>churdle linear</code>                 | Cragg hurdle regression                                            |
| <code>ci means</code>                       | Confidence intervals for means, normal distribution                |
| <code>ci means, poisson</code>              | Confidence intervals for means, Poisson distribution               |
| <code>ci proportions</code>                 | Confidence intervals for proportions                               |
| <code>clogit</code> (k1 to k2 matching)     | Conditional (fixed-effects) logistic regression, k1 to k2 matching |
| <code>clogit</code> (1 to k matching)       | Conditional (fixed-effects) logistic regression, 1 to k matching   |
| <code>cloglog</code>                        | Complementary log-log regression                                   |

Table 2. Command descriptions

| Command                              | Description                                                     |
|--------------------------------------|-----------------------------------------------------------------|
| <code>cluster averagelinkage</code>  | Hierarchical cluster analysis—average linkage                   |
| <code>cluster centroidlinkage</code> | Hierarchical cluster analysis—centroid linkage                  |
| <code>cluster completelinkage</code> | Hierarchical cluster analysis—complete linkage                  |
| <code>cluster generate</code>        | Generate summary and grouping variables from a cluster analysis |
| <code>cluster kmeans</code>          | Kmeans cluster analysis                                         |
| <code>cluster kmedians</code>        | Kmedians cluster analysis                                       |
| <code>cluster medianlinkage</code>   | Hierarchical cluster analysis—median linkage                    |
| <code>cluster singlelinkage</code>   | Hierarchical cluster analysis—single linkage                    |
| <code>cluster wardslinkage</code>    | Hierarchical cluster analysis—Ward’s linkage                    |
| <code>cluster waveragelinkage</code> | Hierarchical cluster analysis—Ward’s average linkage            |
| <code>cmclogit</code>                | Conditional logit (McFadden’s) choice model                     |
| <code>cmmprobit</code>               | Multinomial probit choice model                                 |
| <code>cmroprobit</code>              | Rank-ordered probit choice model                                |
| <code>cnsreg</code>                  | Constrained linear regression                                   |
| <code>codebook</code>                | Describe data contents                                          |
| <code>collapse</code>                | Make dataset of summary datasets                                |
| <code>compare</code>                 | Compare two variables                                           |
| <code>compress</code>                | Compress data in memory                                         |
| <code>contract</code>                | Make dataset of frequencies and percentages                     |
| <code>corr2data</code>               | Create dataset with specified correlation structure             |
| <code>correlate</code>               | Correlations (covariances) of variables or estimators           |
| <code>corrgram</code>                | Tabulate and graph autocorrelations                             |
| <code>count</code>                   | Count observations satisfying specified condition               |
| <code>cpoisson</code>                | Censored Poisson regression                                     |
| <code>cs</code>                      | Cohort study risk-ratio                                         |



Table 2. Command descriptions

| Command                           | Description                                             |
|-----------------------------------|---------------------------------------------------------|
| <code>by: cs</code>               | Cohort study risk-ratio over groups                     |
| <code>ctset</code>                | Declare data to be count-time data                      |
| <code>cttost</code>               | Convert count-time data to survival-time data           |
| <code>cumul</code>                | Cumulative distribution                                 |
| <code>cusum</code>                | Cusum plots and tests for binary variables              |
| <code>datasignature</code>        | Determine whether data have changed                     |
| <code>decode</code>               | Decode labeled numeric into string                      |
| <code>destring</code>             | Convert string variables to numeric variables           |
| <code>dfactor</code>              | Dynamic-factor models                                   |
| <code>dfgls</code>                | DF-GLS unit-root test                                   |
| <code>dfuller</code>              | Augmented Dickey–Fuller unit-root test                  |
| <code>discrim knn</code>          | Discriminant analysis— <i>k</i> th-nearest-neighbor     |
| <code>discrim lda</code>          | Discriminant analysis—linear                            |
| <code>discrim logistic</code>     | Discriminant analysis—logistic                          |
| <code>discrim qda</code>          | Discriminant analysis—quadratic                         |
| <code>dotplot</code>              | Comparative scatterplots                                |
| <code>drawnorm</code>             | Draw sample from multivariate normal distribution       |
| <code>drop if <i>exp</i></code>   | Eliminate observations using <code>if</code> expression |
| <code>drop in <i>range</i></code> | Eliminate observations using <code>in range</code>      |
| <code>dsge</code>                 | Linearized dynamic stochastic general equilibrium model |
| <code>dsge1</code>                | Nonlinear dynamic stochastic general equilibrium model  |
| <code>dslogit</code>              | Double-selection lasso logistic regression              |
| <code>ds poisson</code>           | Double-selection lasso Poisson regression               |
| <code>dsregress</code>            | Double-selection lasso linear regression                |
| <code>dstdize</code>              | Direct and indirect standardization                     |

Table 2. Command descriptions

| Command                         | Description                                                                                             |
|---------------------------------|---------------------------------------------------------------------------------------------------------|
| dvech                           | Diagonal vech multivariate GARCH models                                                                 |
| egen group()                    | Extensions to generate—create grouping variable                                                         |
| by: egen mean                   | Extensions to generate—create means over groups                                                         |
| eivreg                          | Errors-in-variables regression                                                                          |
| encode                          | Encode string into numeric                                                                              |
| eregress                        | Extended linear regression with endogenous covariates, treatment assignment, and sample selection       |
| esize twosample                 | Effect size for two independent samples using groups                                                    |
| esize unpaired                  | Effect size for two independent samples using variables                                                 |
| eteffects (exponential),<br>ate | Endogenous treatment-effects estimation, exponential-mean model, average treatment effect in population |
| eteffects (linear), ate         | Endogenous treatment-effects estimation, linear model, average treatment effect in population           |
| eteffects (linear),<br>pomeans  | Endogenous treatment-effects estimation, linear model, potential-outcome means                          |
| eteffects (probit), ate         | Endogenous treatment-effects estimation, probit model, average treatment effect in population           |
| etpoisson                       | Poisson regression with endogenous treatment effects                                                    |
| etregress, poutcomes            | Linear regression with endogenous treatment effects, ML estimation with potential outcomes              |
| etregress, twostep              | Linear regression with endogenous treatment effects, two-step estimation                                |
| exlogistic                      | Exact logistic regression                                                                               |
| expand #                        | Duplicate observations                                                                                  |
| expand <i>varname</i>           | Duplicate observations using a variable                                                                 |
| expandcl #                      | Duplicate clustered observations                                                                        |
| expandcl <i>varname</i>         | Duplicate clustered observations using a variable                                                       |
| expoisson                       | Exact Poisson regression                                                                                |
| factor                          | Factor analysis                                                                                         |
| fcast compute                   | Dynamic forecasts after VAR or VEC estimation                                                           |
| fillin                          | Rectangularize dataset                                                                                  |
| fmm 2: poisson                  | Finite mixture model with two Poisson outcomes                                                          |

Table 2. Command descriptions

| Command                                   | Description                                                 |
|-------------------------------------------|-------------------------------------------------------------|
| <code>fmm 2: regress</code>               | Finite mixture model with two linear outcomes               |
| <code>fmm 3: poisson</code>               | Finite mixture model with three Poisson outcomes            |
| <code>fmm 3: regress</code>               | Finite mixture model with three linear outcomes             |
| <code>fracreg probit</code>               | Fractional probit regression                                |
| <code>frontier</code>                     | Stochastic frontier models                                  |
| <code>fvrevar (factors)</code>            | Create indicators for factor variables                      |
| <code>fvrevar (interaction)</code>        | Create indicators for factor variables—interactions         |
| <code>generate (small expressions)</code> | Create or change contents of variable—small expressions     |
| <code>generate</code>                     | Create or change contents of variable                       |
| <code>glm, family(gamma)</code>           | Generalized linear models—gamma distribution                |
| <code>glm, family(gaussian)</code>        | Generalized linear models—Gaussian distribution             |
| <code>glm, family(igaussian)</code>       | Generalized linear models—inverse Gaussian distribution     |
| <code>glm, family(nbinomial)</code>       | Generalized linear models—negative binomial distribution    |
| <code>glm, family(poisson)</code>         | Generalized linear models—Poisson distribution              |
| <code>glogit</code>                       | Weighted least-squares logistic regression for grouped data |
| <code>gmm</code>                          | Generalized method of moments estimation                    |
| <code>gmm (with derivatives)</code>       | Generalized method of moments estimation with derivatives   |
| <code>gprobit</code>                      | Weighted least-squares probit regression for grouped data   |
| <code>graph bar</code>                    | Bar charts                                                  |
| <code>graph box</code>                    | Box plots                                                   |
| <code>graph pie</code>                    | Pie charts                                                  |
| <code>grmeanby</code>                     | Graph means and medians by categorical variables            |
| <code>gsem, oprobit (CFA, 2-level)</code> | Ordered probit multilevel confirmatory factor analysis      |
| <code>gsem, oprobit (CFA)</code>          | Ordered probit confirmatory factor analysis                 |
| <code>gsem, logit group()</code>          | GSEM: Logistic regression on 5 groups                       |

Table 2. Command descriptions

| Command                            | Description                                              |
|------------------------------------|----------------------------------------------------------|
| <code>gsem, group()</code>         | GSEM: Linear regression on 5 groups                      |
| <code>gsem, ologit group()</code>  | GSEM: Ordinal logistic regression on 5 groups            |
| <code>gsem, poisson group()</code> | GSEM: Poisson regression on 5 groups                     |
| <code>gsort</code>                 | Ascending and descending sort                            |
| <code>hausman</code>               | Hausman specification test                               |
| <code>heckman</code>               | Heckman selection model—maximum likelihood estimator     |
| <code>heckman, twostep</code>      | Heckman selection model—two-step estimator               |
| <code>heckprobit</code>            | Ordered probit model with sample selection               |
| <code>heckpoisson</code>           | Poisson regression with sample selection                 |
| <code>heckprob</code>              | Probit model with selection                              |
| <code>hetoprobit</code>            | Heteroskedastic ordered probit regression                |
| <code>hetprob</code>               | Heteroskedastic probit model                             |
| <code>hetregress</code>            | Heteroskedastic linear regression, ML estimation         |
| <code>hetregress, twostep</code>   | Heteroskedastic linear regression, two-step estimation   |
| <code>histogram</code>             | Histograms for continuous and categorical variables      |
| <code>hotelling</code>             | Hotelling's $T$ -squared generalized means test          |
| <code>icc, mixed</code>            | Intraclass correlations for two-way mixed-effects model  |
| <code>icc (one-way)</code>         | Intraclass correlations for one-way random-effects model |
| <code>icc (two-way)</code>         | Intraclass correlations for two-way random-effects model |
| <code>intreg</code>                | Interval regression                                      |
| <code>ir</code>                    | Incidence-rate ratio                                     |
| <code>by: ir</code>                | Incidence-rate ratio over groups                         |
| <code>irf create</code>            | Create IRFs and FEVDs after VAR and VEC estimation       |
| <code>irt 1pl</code>               | Item response theory one-parameter logistic model        |
| <code>irt 2pl</code>               | Item response theory two-parameter logistic model        |

Table 2. Command descriptions

| Command                                    | Description                                                                         |
|--------------------------------------------|-------------------------------------------------------------------------------------|
| <code>irt 3pl</code>                       | Item response theory three-parameter logistic model                                 |
| <code>irt grm</code>                       | Item response theory graded response model                                          |
| <code>irt nrm</code>                       | Item response theory nominal response model                                         |
| <code>irt pcm</code>                       | Item response theory partial credit model                                           |
| <code>irt rsm</code>                       | Item response theory rating scale model                                             |
| <code>istdize</code>                       | Indirect standardization                                                            |
| <code>ivpoisson cfunction</code>           | Poisson regression with endogenous regressors, control-function estimator           |
| <code>ivpoisson gmm, additive</code>       | Poisson regression with endogenous regressors, GMM with additive regression errors  |
| <code>ivpoisson gmm, multiplicative</code> | Poisson regression with endogenous regressors, GMM multiplicative regression errors |
| <code>ivprobit</code>                      | Probit model with endogenous regressors                                             |
| <code>ivregress 2sls</code>                | Instrumental-variables regression—two-stage least squares                           |
| <code>ivregress gmm</code>                 | Instrumental-variables regression—GMM                                               |
| <code>ivregress liml</code>                | Instrumental-variables regression—LIML                                              |
| <code>ivtobit</code>                       | Tobit model with endogenous regressors                                              |
| <code>kap</code>                           | Interrater agreement                                                                |
| <code>kappa</code>                         | Interrater agreement                                                                |
| <code>kdensity</code>                      | Univariate kernel density estimation                                                |
| <code>keep if exp</code>                   | Retain observations using <code>if</code> expression                                |
| <code>keep in range</code>                 | Retain observations using <code>in range</code>                                     |
| <code>keep varlist</code>                  | Retain variables                                                                    |
| <code>ksmirnov</code>                      | Kolmogorov–Smirnov equality-of-distributions test                                   |
| <code>ksmirnov, by()</code>                | Kolmogorov–Smirnov equality-of-distributions test over groups                       |
| <code>ktau</code>                          | Kendall’s rank correlation coefficients                                             |
| <code>kwallis</code>                       | Kruskal–Wallis equality-of-populations rank test                                    |
| <code>ladder</code>                        | Ladder of powers                                                                    |

Table 2. Command descriptions

| Command                            | Description                                                            |
|------------------------------------|------------------------------------------------------------------------|
| <code>lasso_linear</code>          | Linear lasso for prediction and model selection                        |
| <code>lasso_logit</code>           | Logistic lasso for prediction and model selection                      |
| <code>lasso_poisson</code>         | Poisson lasso for prediction and model selection                       |
| <code>gsem, lclass(C 2)</code>     | Latent Class Analysis, logit outcomes, 2 classes                       |
| <code>gsem, lclass(C 3)</code>     | Latent Class Analysis, logit outcomes, 3 classes                       |
| <code>levelsof</code>              | Levels of variable                                                     |
| <code>loadingplot</code>           | Score and loading plots after <code>factor</code> and <code>pca</code> |
| <code>logistic</code>              | Logistic regression, reporting odds ratios                             |
| <code>logit</code>                 | Logistic regression, reporting coefficients                            |
| <code>loneaway</code>              | Large one-way ANOVA, random effects, and reliability                   |
| <code>lowess</code>                | Lowess smoothing                                                       |
| <code>lpoly</code>                 | Kernel-weighted local polynomial smoothing                             |
| <code>ltable</code>                | Life tables for survival data                                          |
| <code>manova (one-way)</code>      | Multivariate analysis of variance and covariance, one-way              |
| <code>manova (two-way)</code>      | Multivariate analysis of variance and covariance, two-way              |
| <code>margins</code>               | Marginal means and predictive margins                                  |
| <code>margins, dydx() exp()</code> | Marginal effects of an expression                                      |
| <code>margins, dydx()</code>       | Marginal effects                                                       |
| <code>margins, exp()</code>        | Predictive margins of an expression                                    |
| <code>markout</code>               | Mark observations for exclusion                                        |
| <code>marksample</code>            | Mark observations for inclusion                                        |
| <code>marksample if exp</code>     | Mark observations for inclusion, with <code>if</code> expression       |
| <code>matrix accum</code>          | Form cross-product matrices of variables over observations             |
| <code>matrix eigenvalues</code>    | Eigenvalues of a matrix                                                |
| <code>matrix score</code>          | Inner product of matrix with variables over observations               |

Table 2. Command descriptions

| Command                                                     | Description                                                                            |
|-------------------------------------------------------------|----------------------------------------------------------------------------------------|
| <code>matrix svd</code>                                     | Singular value decomposition                                                           |
| <code>matrix symeigen</code>                                | Eigenvalues of a symmetric matrix                                                      |
| <code>matrix syminv</code>                                  | Inversion of a symmetric matrix                                                        |
| <code>mca</code>                                            | Multiple and joint correspondence analysis                                             |
| <code>mcc</code>                                            | Matched case-control studies                                                           |
| <code>mds</code>                                            | Multidimensional scaling for two-way data                                              |
| <code>mdslong</code>                                        | Multidimensional scaling of proximity data in long format                              |
| <code>mean</code>                                           | Estimate means                                                                         |
| <code>mecloglog</code>                                      | Multilevel mixed-effects complimentary log-log regression                              |
| <code>median</code>                                         | Equality tests on unmatched data                                                       |
| <code>meintreg</code>                                       | Multilevel mixed-effects interval regression                                           |
| <code>melogit</code>                                        | Multilevel mixed-effects logistic regression                                           |
| <code>menbreg,</code><br><code>dispersion(constant)</code>  | Multilevel mixed-effects negative binomial regression, constant dispersion             |
| <code>menbreg, dispersion(mean)</code>                      | Multilevel mixed-effects negative binomial regression, mean dispersion                 |
| <code>menl</code>                                           | Nonlinear mixed-effects regression for a linear outcome                                |
| <code>meologit</code>                                       | Multilevel mixed-effects ordered logistic regression                                   |
| <code>meoprobit</code>                                      | Multilevel mixed-effects ordered probit regression                                     |
| <code>mepoisson</code>                                      | Multilevel mixed-effects Poisson regression                                            |
| <code>meprobit</code>                                       | Multilevel mixed-effects probit regression                                             |
| <code>mestreg,</code><br><code>distribution(exp)</code>     | Multilevel mixed-effects survival models, exponential distribution                     |
| <code>mestreg,</code><br><code>distribution(weibull)</code> | Multilevel mixed-effects survival models, Weibull distribution                         |
| <code>metobit</code>                                        | Multilevel mixed-effects Tobit regression                                              |
| <code>mgarch</code>                                         | Multivariate generalized autoregressive conditional-heteroskedasticity (MGARCH) models |
| <code>mhodds</code>                                         | Ratio of odds of failure for two categories                                            |
| <code>mhodds (adjusted)</code>                              | Ratio of odds of failure for two categories adjusting for levels                       |

Table 2. Command descriptions

| Command                                      | Description                                                                 |
|----------------------------------------------|-----------------------------------------------------------------------------|
| <code>by: mhdods</code>                      | Ratio of odds of failure for two categories over groups                     |
| <code>mhdods (trend)</code>                  | Ratio of odds of failure testing for trend                                  |
| <code>mi estimate: logit (flong)</code>      | Logistic regression with multiply imputed data— <b>flong</b> style data     |
| <code>mi estimate: logit (flongsep)</code>   | Logistic regression with multiply imputed data— <b>flongsep</b> style data  |
| <code>mi estimate: logit (mlong)</code>      | Logistic regression with multiply imputed data— <b>mlong</b> style data     |
| <code>mi estimate: logit (wide)</code>       | Logistic regression with multiply imputed data— <b>wide</b> style data      |
| <code>mi estimate: mlogit</code>             | Multinomial logistic regression with multiply imputed data                  |
| <code>mi estimate: ologit</code>             | Ordered logistic regression with multiply imputed data                      |
| <code>mi estimate: regress (flong)</code>    | Linear regression with multiply imputed data— <b>flong</b> style data       |
| <code>mi estimate: regress (flongsep)</code> | Linear regression with multiply imputed data— <b>flongsep</b> style data    |
| <code>mi estimate: regress (mlong)</code>    | Linear regression with multiply imputed data— <b>mlong</b> style data       |
| <code>mi estimate: regress (wide)</code>     | Linear regression with multiply imputed data— <b>wide</b> style data        |
| <code>mi impute chained (flong)</code>       | Impute missing values using chained equations— <b>flong</b> style data      |
| <code>mi impute chained (flongsep)</code>    | Impute missing values using chained equations— <b>flongsep</b> style data   |
| <code>mi impute chained (mlong)</code>       | Impute missing values using chained equations— <b>mlong</b> style data      |
| <code>mi impute chained (wide)</code>        | Impute missing values using chained equations— <b>wide</b> style data       |
| <code>mi impute logit (flong)</code>         | Impute missing values using logistic regression— <b>flong</b> style data    |
| <code>mi impute logit (flongsep)</code>      | Impute missing values using logistic regression— <b>flongsep</b> style data |
| <code>mi impute logit (mlong)</code>         | Impute missing values using logistic regression— <b>mlong</b> style data    |
| <code>mi impute logit (wide)</code>          | Impute missing values using logistic regression— <b>wide</b> style data     |
| <code>mi impute mlogit</code>                | Impute missing values using multinomial logistic regression                 |
| <code>mi impute mono pmm</code>              | Impute missing values using monotone predictive mean matching               |
| <code>mi impute mono regress</code>          | Impute missing values using monotone linear regression                      |
| <code>mi impute mvn</code>                   | Impute missing values using multivariate normal                             |
| <code>mi impute ologit</code>                | Impute missing values using ordinal logistic regression                     |



Table 2. Command descriptions

| Command                                      | Description                                                                                         |
|----------------------------------------------|-----------------------------------------------------------------------------------------------------|
| <code>mi impute pmm</code>                   | Impute missing values using predictive mean matching                                                |
| <code>mi impute regress</code>               |                                                                                                     |
| <code>misstable nested</code>                | Analyze missing values—list the nesting rules                                                       |
| <code>misstable patterns</code>              | Analyze missing values—report patterns                                                              |
| <code>misstable summarize</code>             | Analyze missing values—report counts                                                                |
| <code>misstable tree</code>                  | Analyze missing values—present tree view                                                            |
| <code>mixed</code>                           | Multilevel mixed-effects linear regression                                                          |
| <code>mixed (crossed effects)</code>         | Multilevel mixed-effects linear regression—crossed effects                                          |
| <code>mkspline</code>                        | Linear spline construction                                                                          |
| <code>mleval</code>                          | Helper command for user-programmed MLEs: Evaluate likelihood of coefficient vector                  |
| <code>mleval, nocons</code>                  | Helper command for user-programmed MLEs: Evaluate likelihood of coefficient vector without constant |
| <code>mlmatbysum</code>                      | Helper command for user-programmed MLEs: Compute Hessians of panel-data estimators                  |
| <code>mlmatsum</code>                        | Helper command for user-programmed MLEs: Compute Hessians of coefficient vector                     |
| <code>mlogit</code>                          | Multinomial (polytomous) logistic regression                                                        |
| <code>mlsum</code>                           | Helper command for user-programmed MLEs: Sum likelihood of coefficient vector                       |
| <code>mlvecsum</code>                        | Helper command for user-programmed MLEs: Compute gradients of coefficient vector                    |
| <code>mprobit</code>                         | Multinomial probit regression                                                                       |
| <code>mswitch ar</code>                      | Markov-switching regression models, autoregression                                                  |
| <code>mswitch dr</code>                      | Markov-switching regression models, dynamic regression                                              |
| <code>mvdecode</code>                        | Recode numeric values to missing                                                                    |
| <code>mvencode</code>                        | Recode missing values to numeric                                                                    |
| <code>mvreg</code>                           | Multivariate regression                                                                             |
| <code>mvtest correlations</code>             | Multivariate test—correlations                                                                      |
| <code>mvtest covariances</code>              | Multivariate test—covariances                                                                       |
| <code>mvtest means,<br/>heterogeneous</code> | Multivariate test—means, heterogenous covariances                                                   |

Table 2. Command descriptions

| Command                                | Description                                                       |
|----------------------------------------|-------------------------------------------------------------------|
| <code>mvtest means, homogeneous</code> | Multivariate test—means, homogeneous covariances                  |
| <code>mvtest means, lr</code>          | Multivariate test—means, likelihood-ratio test                    |
| <code>mvtest normality</code>          | Multivariate test—normality                                       |
| <code>nbreg</code>                     | Negative binomial regression                                      |
| <code>newey</code>                     | Regression with Newey–West standard errors                        |
| <code>nl</code>                        | Nonlinear least-squares estimation                                |
| <code>nlogit</code>                    | Nested logit regression                                           |
| <code>nlshr</code>                     | Estimation of nonlinear systems of equations                      |
| <code>npregress kernel</code>          | Nonparametric kernel regression                                   |
| <code>nptrend</code>                   | Test for trend across ordered groups                              |
| <code>ologit</code>                    | Ordered logistic regression                                       |
| <code>ologit, vce(cluster)</code>      | Ordered logistic regression, cluster–robust standard errors       |
| <code>ologit, vce(robust)</code>       | Ordered logistic regression, robust (Huber/White) standard errors |
| <code>oneway</code>                    | One-way analysis of variance                                      |
| <code>oprobit</code>                   | Ordered probit regression                                         |
| <code>oprobit, vce(cluster)</code>     | Ordered probit regression, cluster–robust standard errors         |
| <code>oprobit, vce(robust)</code>      | Ordered probit regression, robust (Huber/White) standard errors   |
| <code>orthog</code>                    | Orthogonalize variables and compute orthogonal polynomials        |
| <code>pca</code>                       | Principal component analysis                                      |
| <code>pcorr</code>                     | Partial correlation coefficients                                  |
| <code>pctile</code>                    | Create variable containing percentiles                            |
| <code>pergram</code>                   | Periodogram                                                       |
| <code>pkcollapse</code>                | Generate pharmacokinetic measurement dataset                      |
| <code>pkexamine</code>                 | Calculate pharmacokinetic measures                                |
| <code>pksumm</code>                    | Summarize pharmacokinetic data                                    |

Table 2. Command descriptions

| Command                            | Description                                                           |
|------------------------------------|-----------------------------------------------------------------------|
| <code>poisson</code>               | Poisson regression                                                    |
| <code>poisson, vce(cluster)</code> | Poisson regression, cluster-robust standard errors                    |
| <code>poisson, exposure()</code>   | Poisson regression, with exposure                                     |
| <code>poisson, vce(robust)</code>  | Poisson regression, robust (Huber/White) standard errors              |
| <code>pologit</code>               | Partialling-out lasso logistic regression                             |
| <code>popoisson</code>             | Partialling-out lasso Poisson regression                              |
| <code>poregress</code>             | Partialling-out lasso linear regression                               |
| <code>pperron</code>               | Phillips–Perron unit-root test                                        |
| <code>prais</code>                 | Prais–Winsten and Cochrane–Orcutt regression                          |
| <code>predict, cooksd</code>       | Obtain Cook’s distance predictions after estimation                   |
| <code>predict, covratio</code>     | Obtain COVRATIO predictions after estimation                          |
| <code>predict, dfbeta</code>       | Obtain DFBETAs for a variable after estimation                        |
| <code>predict, dfits</code>        | Obtain DFITS predictions after estimation                             |
| <code>predict, e</code>            | Obtain predictions given upper and lower truncation after estimation  |
| <code>predict, leverage</code>     | Obtain leverage of observations after estimation                      |
| <code>predict, pr</code>           | Obtain probability-in-range predictions after estimation              |
| <code>predict, residuals</code>    | Obtain residuals after estimation                                     |
| <code>predict, rstandard</code>    | Obtain standardized residuals after estimation                        |
| <code>predict, rstudent</code>     | Obtain Studentized residuals after estimation                         |
| <code>predict, stdf</code>         | Obtain standard errors of predictions after estimation                |
| <code>predict, stdp</code>         | Obtain standard errors of forecasts after estimation                  |
| <code>predict, stdr</code>         | Obtain standard errors of residuals after estimation                  |
| <code>predict, welsch</code>       | Obtain Welsch distances after estimation                              |
| <code>predict, ystar</code>        | Obtain truncated predictions in a range after estimation              |
| <code>predictnl</code>             | Obtain nonlinear predictions, standard errors, etc., after estimation |

Table 2. Command descriptions

| Command                                  | Description                                                                     |
|------------------------------------------|---------------------------------------------------------------------------------|
| <code>probit</code>                      | Probit regression                                                               |
| <code>procrustes</code>                  | Procrustes transformation                                                       |
| <code>proportion</code>                  | Estimate proportions                                                            |
| <code>prtest1</code>                     | One-sample tests of proportions                                                 |
| <code>prtest2</code>                     | Two-sample tests of proportions                                                 |
| <code>prtest, by()</code>                | Tests of proportions computed over groups                                       |
| <code>pwcorr</code>                      | Pairwise correlation coefficients                                               |
| <code>qreg</code>                        | Quantile (including median) regression                                          |
| <code>ranksum</code>                     | Equality tests on unmatched data                                                |
| <code>ratio</code>                       | Estimate ratio with SE and CI                                                   |
| <code>ratio (exp1) (exp2)</code>         | Estimate two ratios with SE and CI                                              |
| <code>recode</code>                      | Recode categorical variables                                                    |
| <code>reg3</code>                        | Three-stage estimation for systems of simultaneous equations                    |
| <code>regress</code>                     | Linear regression                                                               |
| <code>regress, vce(cluster)</code>       | Linear regression, cluster-robust standard errors                               |
| <code>regress, vce(robust)</code>        | Linear regression, robust (Huber/White) standard errors                         |
| <code>replace</code>                     | Create or change contents of variable                                           |
| <code>replace (small expressions)</code> | Create or change contents of variable, simple expression                        |
| <code>reshape long</code>                | Convert data from wide to long                                                  |
| <code>reshape wide</code>                | Convert data from long to wide                                                  |
| <code>robvar</code>                      | Robust tests for equality of variance                                           |
| <code>rocfit</code>                      | Fit ROC models                                                                  |
| <code>roctab</code>                      | Receiver operating characteristic (ROC) analysis                                |
| <code>rotate</code>                      | Orthogonal and oblique rotations after <code>factor</code> and <code>pca</code> |
| <code>rotatemat</code>                   | Orthogonal and oblique rotations of a Stata matrix                              |

Table 2. Command descriptions

| Command                              | Description                                                            |
|--------------------------------------|------------------------------------------------------------------------|
| <code>rreg</code>                    | Robust regression                                                      |
| <code>runtest</code>                 | Test for random order                                                  |
| <code>scobit</code>                  | Skewed logistic regression                                             |
| <code>scoreplot</code>               | Score and loading plots after <code>factor</code> and <code>pca</code> |
| <code>screeplot</code>               | Scree plot of eigenvalues                                              |
| <code>sdtest1</code>                 | Variance-comparison test against constant                              |
| <code>sdtest2</code>                 | Variance-comparison test between variables                             |
| <code>sdtest, by()</code>            | Variance-comparison test over groups                                   |
| <code>sem, method(adf) (CFA)</code>  | Confirmatory factor analysis, ADF estimation                           |
| <code>sem, method(ml) (CFA)</code>   | Confirmatory factor analysis, ML estimation                            |
| <code>sem, method(mlmv) (CFA)</code> | Confirmatory factor analysis, ML estimation with missing values        |
| <code>sem (SEM latent)</code>        | Structural equations model with latent variables, ML estimation        |
| <code>sem (SEM observed)</code>      | Structural equations model on observed variables, ML estimation        |
| <code>separate</code>                | Create separate variables                                              |
| <code>sfrancia</code>                | Shapiro–Francia test for normality                                     |
| <code>signrank</code>                | Equality tests on matched data                                         |
| <code>signtest</code>                | Equality tests on matched data                                         |
| <code>sktest</code>                  | Skewness and kurtosis test for normality                               |
| <code>slogit</code>                  | Stereotype logistic regression                                         |
| <code>sort</code>                    | Sort data                                                              |
| <code>spearman</code>                | Spearman’s rank correlation coefficients                               |
| <code>sspace</code>                  | State-space models                                                     |
| <code>stack</code>                   | Stack data                                                             |
| <code>stci</code>                    | Confidence intervals for means and percentiles of survival time        |
| <code>stcox</code>                   | Fit Cox proportional hazards model                                     |



Table 2. Command descriptions

| Command                                                           | Description                                                                                        |
|-------------------------------------------------------------------|----------------------------------------------------------------------------------------------------|
| <code>stcrreg</code>                                              | Competing-risks regression                                                                         |
| <code>stgen</code>                                                | Generate variables reflecting entire histories                                                     |
| <code>stintreg, d(exponential)</code>                             | Fit parametric models for interval-censored survival-time data, exponential distribution           |
| <code>stintreg, d(weibull)</code>                                 | Fit parametric models for interval-censored survival-time data, Weibull distribution               |
| <code>stir</code>                                                 | Report incidence-rate comparison                                                                   |
| <code>stmcc</code>                                                | Calculate rate ratios with the Mantel–Cox method                                                   |
| <code>by: stmcc</code>                                            | Calculate rate ratios with the Mantel–Cox method over groups                                       |
| <code>stmh</code>                                                 | Calculate rate ratios with the Mantel–Haenszel method                                              |
| <code>by: stmh</code>                                             | Calculate rate ratios with the Mantel–Haenszel method over groups                                  |
| <code>stptime</code>                                              | Calculate person-time, incidence rates, and SMR                                                    |
| <code>strate</code>                                               | Tabulate failure rates and rate ratios                                                             |
| <code>streg,</code><br><code>distribution(exponential)</code>     | Fit parametric survival models, exponential distribution                                           |
| <code>streg, dist(exp)</code><br><code>vce(cluster)</code>        | Fit parametric survival models, exponential distribution with cluster-robust standard errors       |
| <code>streg, dist(exp)</code><br><code>frailty()</code>           | Fit parametric survival models, exponential distribution with individual frailty                   |
| <code>streg, dist(exp)</code><br><code>frailty() shared()</code>  | Fit parametric survival models, exponential distribution with shared frailty                       |
| <code>streg, dist(exp)</code><br><code>vce(robust)</code>         | Fit parametric survival models, exponential distribution with robust standard errors               |
| <code>streg,</code><br><code>distribution(ggamma)</code>          | Fit parametric survival models, generalized-gamma distribution                                     |
| <code>streg, dist(ggamma)</code><br><code>vce(cluster)</code>     | Fit parametric survival models, generalized-gamma distribution with cluster-robust standard errors |
| <code>streg, dist(ggamma)</code><br><code>vce(robust)</code>      | Fit parametric survival models, generalized-gamma distribution with robust standard errors         |
| <code>streg,</code><br><code>distribution(gompertz)</code>        | Fit parametric survival models, Gompertz distribution                                              |
| <code>streg, dist(gompertz)</code><br><code>vce(cluster)</code>   | Fit parametric survival models, Gompertz distribution with cluster-robust standard errors          |
| <code>streg, dist(gompertz)</code><br><code>frailty()</code>      | Fit parametric survival models, Gompertz distribution with individual frailty                      |
| <code>streg, dist(gomp)</code><br><code>frailty() shared()</code> | Fit parametric survival models, Gompertz distribution with shared frailty                          |
| <code>streg, dist(gompertz)</code><br><code>vce(robust)</code>    | Fit parametric survival models, Gompertz distribution with robust standard errors                  |
| <code>streg,</code><br><code>distribution(llogistic)</code>       | Fit parametric survival models, log-logistic distribution                                          |







Table 2. Command descriptions

| Command                                            | Description                                                                                                              |
|----------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|
| <code>streg, dist(llogistic) vce(cluster)</code>   | Fit parametric survival models, log-logistic distribution with cluster-robust standard errors                            |
| <code>streg, dist(llogistic) frailty()</code>      | Fit parametric survival models, log-logistic distribution with individual frailty                                        |
| <code>streg, dist(llog) frailty() shared()</code>  | Fit parametric survival models, log-logistic distribution with shared frailty                                            |
| <code>streg, dist(llogistic) vce(robust)</code>    | Fit parametric survival models, log-logistic distribution with robust standard errors                                    |
| <code>streg, distribution(lnormal)</code>          | Fit parametric survival models, log-normal distribution                                                                  |
| <code>streg, dist(lnormal) vce(cluster)</code>     | Fit parametric survival models, log-normal distribution with cluster-robust standard errors                              |
| <code>streg, dist(lnormal) frailty()</code>        | Fit parametric survival models, log-normal distribution with individual frailty                                          |
| <code>streg, dist(lnorm) frailty() shared()</code> | Fit parametric survival models, log-normal distribution with shared frailty                                              |
| <code>streg, dist(lnormal) vce(robust)</code>      | Fit parametric survival models, log-normal distribution with robust standard errors                                      |
| <code>streg, distribution(weibull)</code>          | Fit parametric survival models, Weibull distribution                                                                     |
| <code>streg, dist(weibull) vce(cluster)</code>     | Fit parametric survival models, Weibull distribution with cluster-robust standard errors                                 |
| <code>streg, dist(weibull) frailty()</code>        | Fit parametric survival models, Weibull distribution with individual frailty                                             |
| <code>streg, dist(weib) frailty() shared()</code>  | Fit parametric survival models, Weibull distribution with shared frailty                                                 |
| <code>streg, dist(weibull) vce(robust)</code>      | Fit parametric survival models, Weibull distribution with robust standard errors                                         |
| <code>sts generate</code>                          | Create new variables containing survival, hazard, and related functions                                                  |
| <code>sts graph</code>                             | Compute and graph survival, hazard, and related functions                                                                |
| <code>sts list</code>                              | Compute and list survival and related functions                                                                          |
| <code>sts test</code>                              | Test the equality of the survival function across groups                                                                 |
| <code>stset</code>                                 | Declare data to be survival-time data                                                                                    |
| <code>stsplrit</code>                              | Split time-span records                                                                                                  |
| <code>stsum</code>                                 | Summarize survival-time data                                                                                             |
| <code>stteffects ipw (weibull)</code>              | Treatment-effects estimation for survival data, inverse-probability weighting, Weibull distribution                      |
| <code>stteffects ipwra (weibull)</code>            | Treatment-effects estimation for survival data, inverse-probability weighted regression adjustment, Weibull distribution |
| <code>stteffects ra (weibull)</code>               | Treatment-effects estimation for survival data, regression adjustment, Weibull distribution                              |
| <code>stteffects wra (weibull)</code>              | Treatment-effects estimation for survival data, weighted regression adjustment, Weibull distribution                     |



Table 2. Command descriptions

| Command                              | Description                                                          |
|--------------------------------------|----------------------------------------------------------------------|
| <code>stvary</code>                  | Report variables that vary over time                                 |
| <code>suest</code>                   | Seemingly unrelated estimation                                       |
| <code>summarize</code>               | Summary statistics                                                   |
| <code>sunflower</code>               | Density-distribution sunflower plots                                 |
| <code>sureg</code>                   | Zellner's seemingly unrelated regression                             |
| <code>svar</code>                    | Structural vector autoregression models                              |
| <code>svmat</code>                   | Convert variables to matrix and vice versa                           |
| <code>svy brr: logit</code>          | Logistic regression using survey data—balanced repeated replications |
| <code>svy brr: poisson</code>        | Poisson regression using survey data—balanced repeated replications  |
| <code>svy brr: regress</code>        | Linear regression using survey data—balanced repeated replications   |
| <code>svy jackknife: logit</code>    | Logistic regression using survey data—jackknife                      |
| <code>svy jackknife: poisson</code>  | Poisson regression using survey data—jackknife                       |
| <code>svy jackknife: regress</code>  | Linear regression using survey data—jackknife                        |
| <code>svy linearized: logit</code>   | Logistic/logit regression using survey data—linearization            |
| <code>svy linearized: poisson</code> | Poisson regression using count survey data—linearization             |
| <code>svy linearized: regress</code> | Linear regression using survey data—linearization                    |
| <code>swilk</code>                   | Shapiro–Wilk test for normality                                      |
| <code>symmetry</code>                | Symmetry and marginal homogeneity tests                              |
| <code>table (one-way)</code>         | Table of summary statistics, one-way                                 |
| <code>table (two-way)</code>         | Table of summary statistics, two-way                                 |
| <code>tabodds</code>                 | Tabulate odds of failure by category                                 |
| <code>tabodds (adjusted)</code>      | Tabulate odds of failure by category adjusting for levels            |
| <code>tabstat</code>                 | Display table of summary statistics                                  |
| <code>tabstat, by()</code>           | Display table of summary statistics over groups                      |
| <code>tabulate (one-way)</code>      | Tables of frequencies, one-way                                       |

Table 2. Command descriptions

| Command                                                  | Description                                                                                                      |
|----------------------------------------------------------|------------------------------------------------------------------------------------------------------------------|
| <code>tabulate (two-way)</code>                          | Tables of frequencies, two-way                                                                                   |
| <code>teffects aipw (linear)</code>                      | Treatment-effects estimation for linear regression, augmented inverse-probability weighting                      |
| <code>teffects aipw (probit)</code>                      | Treatment-effects estimation for probit regression, augmented inverse-probability weighting                      |
| <code>teffects ipw (logit)</code>                        | Treatment-effects estimation for linear regression, inverse-probability weighting                                |
| <code>teffects ipwra (linear)</code>                     | Treatment-effects estimation for linear regression, inverse-probability weight regression adjustment             |
| <code>teffects ipwra (probit)</code>                     | Treatment-effects estimation for probit regression, augmented inverse-probability weighted regression adjustment |
| <code>teffects nnmatch</code>                            | Treatment-effects estimation, nearest-neighbor matching                                                          |
| <code>teffects psmatch, logit</code>                     | Treatment-effects estimation, propensity-score matching                                                          |
| <code>teffects ra (linear)</code>                        | Treatment-effects estimation for linear regression, regression adjustment                                        |
| <code>teffects ra (probit)</code>                        | Treatment-effects estimation for probit regression, regression adjustment                                        |
| <code>tetrachoric</code>                                 | Tetrachoric correlations for binary variables                                                                    |
| <code>threshold, threshvar()</code>                      | Threshold regression, single threshold for the intercept                                                         |
| <code>threshold, threshvar()<br/>    regionvars()</code> | Threshold regression, single threshold for the intercept and some coefficients                                   |
| <code>tnbreg</code>                                      | Truncated negative binomial regression                                                                           |
| <code>tobit</code>                                       | Tobit regression                                                                                                 |
| <code>tostring</code>                                    | Convert numeric variables to string variables                                                                    |
| <code>total</code>                                       | Estimate totals                                                                                                  |
| <code>tpoisson</code>                                    | Truncated Poisson regression                                                                                     |
| <code>truncreg</code>                                    | Truncated regression                                                                                             |
| <code>tsfilter bk</code>                                 | Time-series filter, Baxter-King                                                                                  |
| <code>tsfilter bw</code>                                 | Time-series filter, Butterworth                                                                                  |
| <code>tsfilter cf</code>                                 | Time-series filter, Christiano-Fitzgerald                                                                        |
| <code>tsfilter hp</code>                                 | Time-series filter, Hodrick-Prescott                                                                             |
| <code>tsrevar</code>                                     | Create time-series operated temporary variables                                                                  |
| <code>tsset</code>                                       | Declare a dataset to be time-series data                                                                         |

Table 2. Command descriptions

| Command                          | Description                                                                                   |
|----------------------------------|-----------------------------------------------------------------------------------------------|
| <code>tssmooth exp</code>        | Exponential smoothing of univariate time-series data                                          |
| <code>tssmooth ma</code>         | Moving average smoothing of univariate time-series data                                       |
| <code>ttest1</code>              | Mean comparison test against constant null hypothesis                                         |
| <code>ttest2</code>              | Mean comparison test against between variables                                                |
| <code>ttest, by()</code>         | Mean comparison test against over groups                                                      |
| <code>twoway fpfit</code>        | Compute and graph fractional-polynomial fit                                                   |
| <code>twoway lfitci</code>       | Compute and graph linear fit with confidence intervals                                        |
| <code>twoway mband</code>        | Compute and graph median bands                                                                |
| <code>twoway mspline</code>      | Compute and graph spline smooth                                                               |
| <code>ucm, model(rwdrift)</code> | Unobserved-components model, random walk with drift                                           |
| <code>var</code>                 | Vector autoregression models                                                                  |
| <code>vargranger</code>          | Perform pairwise Granger causality tests after <code>var</code> or <code>svar</code>          |
| <code>varlmar</code>             | Obtain LM statistics for residual autocorrelation after <code>var</code> or <code>svar</code> |
| <code>varnorm</code>             | Test for normally distributed disturbances after <code>var</code> or <code>svar</code>        |
| <code>varsoc</code>              | Obtain lag-order selection statistics for VARs and VECMs                                      |
| <code>varstable</code>           | Check the stability condition of VAR or SVAR estimates                                        |
| <code>vec</code>                 | Vector error-correction models                                                                |
| <code>veclmar</code>             | Obtain LM statistics for residual autocorrelation after <code>vec</code>                      |
| <code>vecnorm</code>             | Test for normally distributed disturbances after <code>vec</code>                             |
| <code>vecrank</code>             | Estimate the cointegrating rank using Johansen's framework                                    |
| <code>vecstable</code>           | Check the stability condition of VECM estimates                                               |
| <code>vwls</code>                | Variance-weighted least squares                                                               |
| <code>wntestb</code>             | Bartlett's periodogram-based test for white noise                                             |
| <code>wntestq</code>             | Portmanteau ( $Q$ ) test for white noise                                                      |
| <code>xcorr</code>               | Cross-correlogram for bivariate time series                                                   |

Table 2. Command descriptions

| Command                                            | Description                                                               |
|----------------------------------------------------|---------------------------------------------------------------------------|
| <code>xpologit</code>                              | Cross-fit partialling-out lasso logistic regression                       |
| <code>xpopoisson</code>                            | Cross-fit partialling-out lasso Poisson regression                        |
| <code>xporegress</code>                            | Cross-fit partialling-out lasso linear regression                         |
| <code>xtabond</code>                               | Arellano–Bond linear, dynamic panel-data estimation                       |
| <code>xtabond, twostep</code>                      | Arellano–Bond linear, dynamic panel-data estimation, two-step estimation  |
| <code>xtcloglog, re</code>                         | Random-effects cloglog models                                             |
| <code>xtdata, be</code>                            | Compute between transform of panel data                                   |
| <code>xtdata, fe</code>                            | Compute within (fixed-effects) transform of panel data                    |
| <code>xtdata, re</code>                            | Compute random-effects transform of panel data                            |
| <code>xtdpd</code>                                 | Linear dynamic panel-data estimation                                      |
| <code>xtdpdsys</code>                              | Arellano–Bover/Blundell–Bond linear dynamic panel-data estimation         |
| <code>xteregress</code>                            | Extended linear regression with random effects                            |
| <code>xtfrontier</code>                            | Stochastic frontier models for panel data                                 |
| <code>xtgee, family(gaussian)<br/>corr(ar2)</code> | GEE estimation of Gaussian panel-data model with 2-period autocorrelation |
| <code>xtgee, fam(gauss)<br/>corr(unstruct)</code>  | GEE estimation of Gaussian panel-data model with unstructured correlation |
| <code>xtcloglog, pa</code>                         | Population-averaged cloglog models                                        |
| <code>xtlogit, pa</code>                           | Population-averaged logit models                                          |
| <code>xtnbreg, pa</code>                           | Population-averaged negative binomial models                              |
| <code>xtpoisson, pa</code>                         | Population-averaged Poisson models                                        |
| <code>xtprobit, pa</code>                          | Population-averaged probit models                                         |
| <code>xtreg, pa</code>                             | Population-averaged linear models                                         |
| <code>xtgls</code>                                 | Fit panel-data models using GLS                                           |
| <code>xthtaylor</code>                             | Hausman–Taylor estimator for error-components models                      |
| <code>xtile</code>                                 | Panel-data line plots                                                     |
| <code>xtintreg</code>                              | Random-effects interval data regression models                            |

Table 2. Command descriptions

| Command                                             | Description                                                                                |
|-----------------------------------------------------|--------------------------------------------------------------------------------------------|
| <code>xtivreg, be</code>                            | Instrumental variables and two-stage least squares for panel-data models—between effects   |
| <code>xtivreg, fd</code>                            | Instrumental variables and two-stage least squares for panel-data models—first differences |
| <code>xtivreg, fe</code>                            | Instrumental variables and two-stage least squares for panel-data models—fixed effects     |
| <code>xtivreg, re</code>                            | Instrumental variables and two-stage least squares for panel-data models—random effects    |
| <code>xtlogit, fe</code>                            | Fixed-effects logit models                                                                 |
| <code>xtlogit, re</code>                            | Random-effects logit models                                                                |
| <code>xtnbreg, fe</code>                            | Fixed-effects negative binomial models                                                     |
| <code>xtnbreg, re</code>                            | Random-effects negative binomial models                                                    |
| <code>xtologit</code>                               | Random-effects ordered logistic models                                                     |
| <code>xtoprobit</code>                              | Random-effects ordered probit models                                                       |
| <code>xtpcse</code>                                 | OLS or Prais–Winsten models with panel-corrected standard errors                           |
| <code>xtpoisson, fe</code>                          | Fixed-effects Poisson models                                                               |
| <code>xtpoisson, re</code>                          | Random-effects Poisson models                                                              |
| <code>xtprobit, re</code>                           | Random-effects probit models                                                               |
| <code>xtrc</code>                                   | Random-coefficients regression                                                             |
| <code>xtreg, be</code>                              | Between-effects linear models                                                              |
| <code>xtreg, fe</code>                              | Fixed-effects linear models                                                                |
| <code>xtreg, fe vce(robust)</code>                  | Fixed-effects linear models, cluster–robust standard errors                                |
| <code>xtreg, mle</code>                             | Random-effects linear models, ML estimation                                                |
| <code>xtreg, re</code>                              | Random-effects linear models                                                               |
| <code>xtregar, fe</code>                            | Fixed-effects linear models with an AR(1) disturbance                                      |
| <code>xtregar, re</code>                            | Random-effects linear models with an AR(1) disturbance                                     |
| <code>xtset</code>                                  | Declare data to be panel data                                                              |
| <code>xtstreg,<br/>distribution(exponential)</code> | Random-effects survival models, exponential distribution                                   |
| <code>xtstreg,<br/>distribution(weibull)</code>     | Random-effects survival models, Weibull distribution                                       |



Table 2. Command descriptions

| Command                          | Description                                                                                      |
|----------------------------------|--------------------------------------------------------------------------------------------------|
| <code>xtsum</code>               | Summarize panel data                                                                             |
| <code>xttab</code>               | Tabulate panel data                                                                              |
| <code>xttobit</code>             | Random-effects tobit models                                                                      |
| <code>xtunitroot breitung</code> | Panel-data unit-root test—Breitung                                                               |
| <code>xtunitroot fisher</code>   | Panel-data unit-root test—Fisher                                                                 |
| <code>xtunitroot hadri</code>    | Panel-data unit-root test—Hadri Lagrange multiplier                                              |
| <code>xtunitroot ht</code>       | Panel-data unit-root test—Harris–Tzavalis                                                        |
| <code>xtunitroot ips</code>      | Panel-data unit-root test—Im–Pesaran–Shin                                                        |
| <code>xtunitroot llc</code>      | Panel-data unit-root test—Levin–Lin–Chu                                                          |
| <code>zinb</code>                | Zero-inflated negative binomial regression                                                       |
| <code>zioprobit</code>           | Zero-inflated ordered probit regression                                                          |
| <code>zip</code>                 | Zero-inflated Poisson regression                                                                 |
| <code>_predict, xb</code>        | Obtain predictions, residuals, etc., after estimation programming command—option <code>xb</code> |
| <code>_rmcoll</code>             | Remove collinear variables                                                                       |
| <code>_robust</code>             | Robust variance estimates                                                                        |

## D Problem sizes

The following table (table 3) shows the sizes of the problems used to measure the performance gains reported in table 1. As discussed in section 9, these are intentionally large problems requiring considerable time to run. If a command was so fast that a sufficiently large problem would have required too much memory to be run on a variety of computers, then a smaller problem was run several times (several iterations) for an accurate read of the timing required to run the command.

The second through fourth columns of table 3 record the number of observations for the problem, either as a simple number of observations  $N$  or as a number of panels  $m$  and a number of time periods  $t$  within a panel. Columns 3 and 4 provide more information on problem size for longitudinal panel-data problems, and the number of observations,  $N$ , is just the product of  $m$  and  $t$ . Some such problems are not really panel data but merely grouped data; in these cases, the time periods should just be considered the number of observations within group. Almost all the panel-data problems were created with balanced panels (an equal number of observations within panel). Rarely would unbalanced panels affect the performance gains of Stata/MP.

The column labeled  $k$  records the number of covariates in the problem or, for matrix commands, the row and column dimensions of the matrix.

The column labeled  $d_1$  records a miscellaneous dimension, such as number of equations for problems that involve multiple equations. Some commands have more miscellaneous dimensions; only  $d_1$  is shown here.

The column labeled  $n_{\text{iter}}$  records the number of times the command was run on the problem to generate a single timing.

Table 3. Problem sizes

| Command               | Observations |     |     | $k$ | $d_1$ | $n_{\text{iter}}$ |
|-----------------------|--------------|-----|-----|-----|-------|-------------------|
|                       | $N$          | $m$ | $t$ |     |       |                   |
| alpha                 | 2250000      |     |     | 20  |       | 1                 |
| ameans                | 3000000      |     |     | 5   |       | 1                 |
| anova (one-way)       | 80000000     |     |     | 200 |       | 1                 |
| anova (two-way)       | 10000000     |     |     | 10  |       | 1                 |
| arch                  | 80000        |     |     | 1   |       | 1                 |
| areg                  | 6000000      |     |     | 20  | 30000 | 1                 |
| areg, vce(cluster)    | 2000000      |     |     | 20  | 20000 | 1                 |
| areg, vce(robust)     | 2000000      |     |     | 20  | 20000 | 1                 |
| arfima                | 1000         |     |     | 1   |       | 1                 |
| arima                 | 80000        |     |     | 1   |       | 1                 |
| bayes: logit          | 300000       |     |     | 20  |       | 1                 |
| bayes: poisson        | 200000       |     |     | 20  |       | 1                 |
| bayes: regress        | 300000       |     |     | 20  |       | 1                 |
| bayesmh logit         | 10000        |     |     | 50  |       | 1                 |
| bayesmh mvn           | 20000        |     |     | 30  | 3     | 1                 |
| bayesmh mylogit       | 10000        |     |     | 10  |       | 1                 |
| bayesmh nl            | 10000        |     |     | 10  |       | 1                 |
| bayesmh normal        | 10000        |     |     | 100 |       | 1                 |
| bayesmh normal gibbs  | 10000        |     |     | 10  |       | 1                 |
| bayesmh normal re     |              | 10  | 100 | 100 |       | 1                 |
| betareg, link(logit)  | 100000       |     |     | 200 |       | 1                 |
| betareg, link(probit) | 100000       |     |     | 200 |       | 1                 |
| binreg                | 200000       |     |     | 200 |       | 1                 |
| biplot                | 4000         |     |     | 2   |       | 1                 |
| biprobit              | 160000       |     |     | 40  | 40    | 1                 |

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.

Table 3. Problem sizes

| Command                        | Observations |         |       | $k$ | $d_1$ | $n_{\text{iter}}$ |
|--------------------------------|--------------|---------|-------|-----|-------|-------------------|
|                                | $N$          | $m$     | $t$   |     |       |                   |
| biprobit (seemingly unrelated) | 160000       |         |       | 40  | 40    | 1                 |
| bitest                         | 10000000     |         |       | 1   | 2     | 10                |
| blogit                         | 200000       |         |       | 200 | 50    | 1                 |
| boxcox                         | 100000       |         |       | 200 |       | 1                 |
| bprobit                        | 200000       |         |       | 200 | 50    | 1                 |
| brier                          | 150000       |         |       |     |       | 1                 |
| bsample                        | 100000       |         |       | 100 |       | 20                |
| bstat                          | 1000000      |         |       | 10  |       | 1                 |
| by: generate                   |              | 1000000 | 100   |     |       | 6                 |
| by: generate (small groups)    |              | 9000000 | 10    |     |       | 2                 |
| by: replace                    |              | 1000000 | 100   |     |       | 6                 |
| by: replace (small groups)     |              | 9000000 | 10    |     |       | 2                 |
| ca                             | 10000000     |         |       |     | 5     | 1                 |
| candisc                        |              | 5       | 40000 | 150 |       | 1                 |
| canon                          | 4000000      |         |       |     | 30    | 1                 |
| cc                             | 500000       |         |       |     |       | 1                 |
| by: cc                         | 100000       |         |       |     | 20    | 1                 |
| centile                        | 1000000      |         |       | 2   |       | 1                 |
| churdle linear                 | 200000       |         |       | 50  | 50    | 1                 |
| ci means                       | 1000000      |         |       | 50  |       | 1                 |
| ci means, poisson              | 100000       |         |       | 50  |       | 8                 |
| ci proportions                 | 1000000      |         |       | 50  |       | 1                 |
| clogit (k1 to k2 matching)     |              | 20000   | 10    | 30  |       | 1                 |
| clogit (1 to k matching)       |              | 50000   | 10    | 50  |       | 1                 |
| cloglog                        | 200000       |         |       | 100 |       | 1                 |

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.

Table 3. Problem sizes

| Command                 | Observations |     |     | $k$ | $d_1$ | $n_{\text{iter}}$ |
|-------------------------|--------------|-----|-----|-----|-------|-------------------|
|                         | $N$          | $m$ | $t$ |     |       |                   |
| cluster averagelinkage  | 4000         |     |     | 200 |       | 1                 |
| cluster centroidlinkage | 4000         |     |     | 200 |       | 1                 |
| cluster completelinkage | 4000         |     |     | 200 |       | 1                 |
| cluster generate        | 2000         |     |     | 200 |       | 1                 |
| cluster kmeans          | 50000        |     |     | 30  |       | 1                 |
| cluster kmedians        | 50000        |     |     | 30  |       | 1                 |
| cluster medianlinkage   | 5000         |     |     | 200 |       | 1                 |
| cluster singlelinkage   | 5000         |     |     | 5   |       | 1                 |
| cluster wardslinkage    | 3000         |     |     | 200 |       | 1                 |
| cluster waveragelinkage | 3000         |     |     | 200 |       | 1                 |
| cmclogit                | 3300         |     |     | 100 | 10    | 1                 |
| cmmprobit               |              | 200 | 3   | 2   | 2     | 1                 |
| cmroprobit              | 300          |     |     | 2   | 3     | 1                 |
| cnsreg                  | 1400000      |     |     | 200 |       | 1                 |
| codebook                | 150000       |     |     | 25  |       | 1                 |
| collapse                | 300000       |     |     | 50  | 100   | 1                 |
| compare                 | 6000000      |     |     | 2   |       | 2                 |
| compress                | 500000       |     |     | 50  | 50    | 1                 |
| contract                | 1000000      |     |     | 20  | 100   | 1                 |
| corr2data               | 200000       |     |     | 50  |       | 1                 |
| correlate               | 3000000      |     |     | 200 |       | 1                 |
| corrgram                | 80000        |     |     | 1   |       | 1                 |
| count                   | 20000000     |     |     |     |       | 20                |
| cpoisson                | 100000       |     |     | 100 |       | 1                 |
| cs                      | 10000000     |     |     |     |       | 1                 |

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.

Table 3. Problem sizes

| Command              | Observations |       |      | $k$ | $d_1$ | $n_{\text{iter}}$ |
|----------------------|--------------|-------|------|-----|-------|-------------------|
|                      | $N$          | $m$   | $t$  |     |       |                   |
| by: cs               | 60000        |       |      |     | 100   | 1                 |
| ctset                | 40000000     |       |      |     |       | 15                |
| cttost               | 50000        |       |      |     |       | 1                 |
| cumul                | 1000000      |       |      | 2   |       | 1                 |
| cusum                | 1500000      |       |      | 1   |       | 1                 |
| datasignature        | 500000       |       |      | 300 |       | 1                 |
| decode               |              | 10000 | 1000 |     |       | 1                 |
| destring             |              | 4000  | 2000 |     |       | 1                 |
| dfactor              | 2000         |       |      | 3   |       | 1                 |
| dfgls                | 20000        |       |      | 1   |       | 1                 |
| dfuller              | 5000000      |       |      | 1   |       | 3                 |
| discrim knn          |              | 5     | 1000 | 20  |       | 1                 |
| discrim lda          |              | 50    | 2000 | 10  |       | 1                 |
| discrim logistic     |              | 50    | 400  | 10  |       | 1                 |
| discrim qda          |              | 50    | 2000 | 10  |       | 1                 |
| dotplot              | 100000       |       |      | 10  |       | 1                 |
| drawnorm             | 100000       |       |      | 150 |       | 1                 |
| drop if <i>exp</i>   | 10000000     |       |      | 4   |       | 1                 |
| drop in <i>range</i> | 10000000     |       |      | 4   |       | 1                 |
| dsge                 | 10000        |       |      | 4   |       | 1                 |
| dsgenl               | 10000        |       |      | 4   |       | 1                 |
| dslogit              | 100000       |       |      | 40  |       | 1                 |
| ds poisson           | 100000       |       |      | 40  |       | 1                 |
| dsregress            | 100000       |       |      | 40  |       | 1                 |
| dstdize              |              | 10    | 150  | 200 |       | 1                 |

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.

Table 3. Problem sizes

| Command                         | Observations |       |        | $k$ | $d_1$ | $n_{\text{iter}}$ |
|---------------------------------|--------------|-------|--------|-----|-------|-------------------|
|                                 | $N$          | $m$   | $t$    |     |       |                   |
| dvech                           | 500          |       |        | 2   |       | 1                 |
| egen group()                    |              | 1     | 800000 | 500 |       | 1                 |
| by: egen mean                   |              | 400   | 10000  | 2   |       | 1                 |
| eivreg                          | 1400000      |       |        | 200 |       | 1                 |
| encode                          |              | 50    | 220000 |     |       | 1                 |
| eregress                        | 20000        |       |        | 10  | 10    | 1                 |
| esize twosample                 | 10000000     |       |        |     |       | 1                 |
| esize unpaired                  | 30000000     |       |        |     |       | 1                 |
| eteffects (exponential),<br>ate | 20000        |       |        | 20  |       | 1                 |
| eteffects (linear), ate         | 10000        |       |        | 100 |       | 1                 |
| eteffects (linear),<br>pomeans  | 10000        |       |        | 100 |       | 1                 |
| eteffects (probit), ate         | 10000        |       |        | 100 |       | 1                 |
| etpoisson                       | 10000        |       |        | 10  | 10    | 1                 |
| etregress, poutcomes            | 10000        |       |        | 30  | 30    | 1                 |
| etregress, twostep              | 800000       |       |        | 50  | 50    | 1                 |
| exlogistic                      | 100          |       |        | 3   |       | 1                 |
| expand #                        | 10000        |       |        | 800 |       | 1                 |
| expand varname                  | 100000       |       |        | 100 | 5     | 1                 |
| expandcl #                      |              | 12000 | 10     | 100 |       | 1                 |
| expandcl varname                |              | 30000 | 10     | 80  | 5     | 1                 |
| expoisson                       | 50           |       |        | 20  |       | 1                 |
| factor                          | 10000000     |       |        | 50  |       | 1                 |
| fcast compute                   | 10000        |       |        | 2   | 5     | 1                 |
| fillin                          |              | 80    | 1      |     |       | 1                 |
| fmm 2: poisson                  | 50000        |       |        | 30  | 2     | 1                 |

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.

Table 3. Problem sizes

| Command                         | Observations |      |       | $k$  | $d_1$ | $n_{\text{iter}}$ |
|---------------------------------|--------------|------|-------|------|-------|-------------------|
|                                 | $N$          | $m$  | $t$   |      |       |                   |
| fmm 2: regress                  | 50000        |      |       | 30   | 2     | 1                 |
| fmm 3: poisson                  | 50000        |      |       | 20   | 3     | 1                 |
| fmm 3: regress                  | 50000        |      |       | 20   | 3     | 1                 |
| fracreg probit                  | 200000       |      |       | 200  |       | 1                 |
| frontier                        | 400000       |      |       | 200  |       | 1                 |
| fvrevar (factors)               | 1000000      |      |       | 4    | 80    | 1                 |
| fvrevar (interaction)           | 5000000      |      |       | 2    | 8     | 1                 |
| generate (small expressions)    | 60000        |      |       | 4000 |       | 1                 |
| generate                        | 5000000      |      |       |      |       | 1                 |
| glm, family(gamma)              | 700000       |      |       | 100  |       | 1                 |
| glm, family(gaussian)           | 700000       |      |       | 200  |       | 1                 |
| glm, family(igaussian)          | 500000       |      |       | 200  |       | 1                 |
| glm, family(nbinomial)          | 300000       |      |       | 200  |       | 1                 |
| glm, family(poisson)            | 300000       |      |       | 200  |       | 1                 |
| glogit                          | 2000000      |      |       | 100  | 50    | 1                 |
| gmm                             | 1000         |      |       | 10   |       | 1                 |
| gmm (with derivatives)          | 100000       |      |       | 10   |       | 1                 |
| gprobit                         | 3000000      |      |       | 100  | 50    | 1                 |
| graph bar                       | 500000       |      |       | 10   | 3     | 1                 |
| graph box                       | 200000       |      |       | 2    | 10    | 1                 |
| graph pie                       | 2500000      |      |       | 10   | 10    | 1                 |
| grmeanby                        | 300000       |      |       | 4    | 10    | 1                 |
| gsem, oprobit (CFA,<br>2-level) |              | 1000 | 10    | 4    | 1     | 1                 |
| gsem, oprobit (CFA)             | 5000         |      |       | 4    | 1     | 1                 |
| gsem, logit group()             |              | 5    | 50000 | 40   |       | 1                 |

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.



Table 3. Problem sizes

| Command               | Observations |     |       | $k$ | $d_1$ | $n_{\text{iter}}$ |
|-----------------------|--------------|-----|-------|-----|-------|-------------------|
|                       | $N$          | $m$ | $t$   |     |       |                   |
| gsem, group()         |              | 5   | 50000 | 40  |       | 1                 |
| gsem, ologit group()  |              | 5   | 50000 | 40  |       | 1                 |
| gsem, poisson group() |              | 5   | 50000 | 40  |       | 1                 |
| gsort                 | 1000000      |     |       | 5   |       | 1                 |
| hausman               | 200          |     |       |     |       | 1                 |
| heckman               | 500000       |     |       | 100 | 50    | 1                 |
| heckman, twostep      | 1000000      |     |       | 100 | 50    | 1                 |
| heckoprobit           | 100000       |     |       | 10  | 50    | 1                 |
| heckpoisson           | 10000        |     |       | 40  | 20    | 1                 |
| heckprob              | 200000       |     |       | 50  | 50    | 1                 |
| hetoprobit            | 300000       |     |       | 10  | 10    | 1                 |
| hetprob               | 300000       |     |       | 10  | 10    | 1                 |
| hetregress            | 500000       |     |       | 100 | 50    | 1                 |
| hetregress, twostep   | 1000000      |     |       | 50  | 5     | 1                 |
| histogram             | 4000000      |     |       | 1   |       | 1                 |
| hotelling             | 4000000      |     |       | 100 |       | 1                 |
| icc, mixed            | 1000000      |     |       | 100 |       | 1                 |
| icc (one-way)         | 3000000      |     |       | 300 |       | 1                 |
| icc (two-way)         | 1000000      |     |       | 100 |       | 1                 |
| intreg                | 200000       |     |       | 200 |       | 1                 |
| ir                    | 10000000     |     |       |     |       | 1                 |
| by: ir                | 10000        |     |       |     | 200   | 1                 |
| irf create            | 1000000      |     |       | 2   | 3     | 1                 |
| irt 1pl               | 40000        |     |       | 20  |       | 1                 |
| irt 2pl               | 40000        |     |       | 20  |       | 1                 |

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.

Table 3. Problem sizes

| Command                          | Observations |     |     | $k$   | $d_1$ | $n_{\text{iter}}$ |
|----------------------------------|--------------|-----|-----|-------|-------|-------------------|
|                                  | $N$          | $m$ | $t$ |       |       |                   |
| irt 3pl                          | 40000        |     |     | 10    |       | 1                 |
| irt grm                          | 20000        |     |     | 10    |       | 1                 |
| irt nrm                          | 20000        |     |     | 10    |       | 1                 |
| irt pcm                          | 20000        |     |     | 10    |       | 1                 |
| irt rsm                          | 20000        |     |     | 10    |       | 1                 |
| istdize                          |              | 50  | 100 | 10000 |       | 1                 |
| ivpoisson cfunction              | 60000        |     |     | 5     | 5     | 1                 |
| ivpoisson gmm, additive          | 80000        |     |     | 5     | 5     | 1                 |
| ivpoisson gmm,<br>multiplicative | 160000       |     |     | 5     | 5     | 1                 |
| ivprobit                         | 150000       |     |     | 30    | 20    | 1                 |
| ivregress 2sls                   | 800000       |     |     | 50    | 20    | 1                 |
| ivregress gmm                    | 1500000      |     |     | 20    | 20    | 1                 |
| ivregress liml                   | 2000000      |     |     | 20    | 20    | 1                 |
| ivtobit                          | 150000       |     |     | 50    | 20    | 1                 |
| kap                              | 500000       |     |     | 2     | 10    | 4                 |
| kappa                            | 2000000      |     |     | 10    | 20    | 1                 |
| kdensity                         | 10000000     |     |     |       |       | 1                 |
| keep if <i>exp</i>               | 10000        |     |     | 4000  |       | 1                 |
| keep in <i>range</i>             | 20000        |     |     | 4000  |       | 1                 |
| keep <i>varlist</i>              | 50000        |     |     | 4000  |       | 1                 |
| ksmirnov                         | 2000000      |     |     |       |       | 1                 |
| ksmirnov, by()                   | 1000000      |     |     |       |       | 1                 |
| ktau                             | 5000         |     |     | 5     |       | 1                 |
| kwallis                          | 1500000      |     |     | 10    |       | 1                 |
| ladder                           | 2000000      |     |     |       |       | 1                 |

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.

Table 3. Problem sizes

| Command                            | Observations |     |     | $k$  | $d_1$ | $n_{\text{iter}}$ |
|------------------------------------|--------------|-----|-----|------|-------|-------------------|
|                                    | $N$          | $m$ | $t$ |      |       |                   |
| <code>lasso_linear</code>          | 100000       |     |     | 20   |       | 1                 |
| <code>lasso_logit</code>           | 20000        |     |     | 20   |       | 1                 |
| <code>lasso_poisson</code>         | 20000        |     |     | 20   |       | 1                 |
| <code>gsem, lclass(C 2)</code>     | 500000       |     |     | 5    | 2     | 1                 |
| <code>gsem, lclass(C 3)</code>     | 50000        |     |     | 10   | 3     | 1                 |
| <code>levelsof</code>              | 20000000     |     |     |      | 20    | 1                 |
| <code>loadingplot</code>           | 2000000      |     |     | 60   |       | 1                 |
| <code>logistic</code>              | 300000       |     |     | 200  |       | 1                 |
| <code>logit</code>                 | 300000       |     |     | 200  |       | 1                 |
| <code>loneaway</code>              | 2000000      |     |     | 500  |       | 1                 |
| <code>lowess</code>                | 90000        |     |     | 1    |       | 1                 |
| <code>lpoly</code>                 | 1000000      |     |     |      |       | 1                 |
| <code>ltable</code>                | 50000        |     |     | 1    |       | 40                |
| <code>manova (one-way)</code>      | 20000000     |     |     | 50   | 3     | 1                 |
| <code>manova (two-way)</code>      | 2000000      |     |     | 20   | 3     | 1                 |
| <code>margins</code>               | 250000       |     |     | 40   | 10    | 1                 |
| <code>margins, dydx() exp()</code> | 30000        |     |     | 40   | 10    | 1                 |
| <code>margins, dydx()</code>       | 20000        |     |     | 40   | 10    | 1                 |
| <code>margins, exp()</code>        | 40000        |     |     | 40   | 10    | 1                 |
| <code>markout</code>               | 500000       |     |     | 500  |       | 1                 |
| <code>marksample</code>            | 1200000      |     |     | 200  |       | 1                 |
| <code>marksample if exp</code>     | 2300000      |     |     | 100  |       | 1                 |
| <code>matrix accum</code>          | 3000000      |     |     | 200  |       | 1                 |
| <code>matrix eigenvalues</code>    | 500          |     |     | 500  |       | 1                 |
| <code>matrix score</code>          | 6000000      |     |     | 1000 |       | 1                 |

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.

Table 3. Problem sizes

| Command                           | Observations |      |     | $k$  | $d_1$ | $n_{\text{iter}}$ |
|-----------------------------------|--------------|------|-----|------|-------|-------------------|
|                                   | $N$          | $m$  | $t$ |      |       |                   |
| matrix svd                        | 300          |      |     | 300  |       | 1                 |
| matrix symeigen                   | 600          |      |     | 600  |       | 1                 |
| matrix syminv                     | 2000         |      |     | 2000 |       | 1                 |
| mca                               | 1000000      |      |     | 3    | 5     | 1                 |
| mcc                               | 10000000     |      |     |      |       | 1                 |
| mds                               | 800          |      |     | 400  |       | 1                 |
| mdslong                           |              | 600  | 1   |      |       | 1                 |
| mean                              | 1000000      |      |     | 200  |       | 1                 |
| mecloglog                         |              | 2000 | 10  | 2    | 1     | 1                 |
| median                            | 8000000      |      |     | 5    |       | 1                 |
| meintreg                          |              | 1000 | 50  | 5    | 1     | 1                 |
| melogit                           |              | 4000 | 10  | 10   | 1     | 1                 |
| menbreg,<br>dispersion(constant)  |              | 2000 | 5   | 2    | 1     | 1                 |
| menbreg, dispersion(mean)         |              | 4000 | 10  | 2    | 1     | 1                 |
| menl                              |              | 1000 | 5   | 3    | 1     | 1                 |
| meologit                          |              | 4000 | 10  | 5    | 1     | 1                 |
| meoprobit                         |              | 4000 | 10  | 2    | 1     | 1                 |
| mepoisson                         |              | 4000 | 10  | 2    | 1     | 1                 |
| meprobit                          |              | 4000 | 10  | 10   | 1     | 1                 |
| mestreg,<br>distribution(exp)     |              | 4000 | 10  | 10   | 1     | 1                 |
| mestreg,<br>distribution(weibull) |              | 4000 | 10  | 10   | 1     | 1                 |
| metobit                           |              | 1000 | 50  | 5    | 1     | 1                 |
| mgarch                            | 1000         |      |     | 3    | 2     | 1                 |
| mhodds                            | 3000000      |      |     |      |       | 1                 |
| mhodds (adjusted)                 | 400000       |      |     | 400  |       | 1                 |

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.

Table 3. Problem sizes

| Command                            | Observations |     |     | $k$ | $d_1$ | $n_{\text{iter}}$ |
|------------------------------------|--------------|-----|-----|-----|-------|-------------------|
|                                    | $N$          | $m$ | $t$ |     |       |                   |
| by: mhodds                         | 50000        |     |     |     | 100   | 1                 |
| mhodds (trend)                     | 1000000      |     |     |     | 100   | 1                 |
| mi estimate: logit<br>(flong)      | 100000       |     |     | 180 | 20    | 1                 |
| mi estimate: logit<br>(flongsep)   | 100000       |     |     | 180 | 20    | 1                 |
| mi estimate: logit<br>(mlong)      | 100000       |     |     | 180 | 20    | 1                 |
| mi estimate: logit (wide)          | 70000        |     |     | 180 | 20    | 1                 |
| mi estimate: mlogit                | 100000       |     |     | 100 | 10    | 1                 |
| mi estimate: ologit                | 120000       |     |     | 190 | 10    | 1                 |
| mi estimate: regress<br>(flong)    | 100000       |     |     | 300 | 20    | 1                 |
| mi estimate: regress<br>(flongsep) | 100000       |     |     | 300 | 20    | 1                 |
| mi estimate: regress<br>(mlong)    | 100000       |     |     | 300 | 20    | 1                 |
| mi estimate: regress<br>(wide)     | 60000        |     |     | 300 | 20    | 1                 |
| mi impute chained (flong)          | 20000        |     |     | 20  | 20    | 1                 |
| mi impute chained<br>(flongsep)    | 20000        |     |     | 20  | 20    | 1                 |
| mi impute chained (mlong)          | 20000        |     |     | 20  | 20    | 1                 |
| mi impute chained (wide)           | 20000        |     |     | 20  | 20    | 1                 |
| mi impute logit (flong)            | 100000       |     |     | 100 | 1     | 1                 |
| mi impute logit<br>(flongsep)      | 100000       |     |     | 100 | 1     | 1                 |
| mi impute logit (mlong)            | 100000       |     |     | 100 | 1     | 1                 |
| mi impute logit (wide)             | 200000       |     |     | 100 | 1     | 1                 |
| mi impute mlogit                   | 100000       |     |     | 100 | 1     | 1                 |
| mi impute mono pmm                 | 10000        |     |     | 50  | 3     | 1                 |
| mi impute mono regress             | 40000        |     |     | 200 | 10    | 1                 |
| mi impute mvn                      | 1000         |     |     | 10  | 10    | 1                 |
| mi impute ologit                   | 40000        |     |     | 100 | 1     | 1                 |

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.

Table 3. Problem sizes

| Command                                      | Observations |     |        | $k$ | $d_1$  | $n_{\text{iter}}$ |
|----------------------------------------------|--------------|-----|--------|-----|--------|-------------------|
|                                              | $N$          | $m$ | $t$    |     |        |                   |
| <code>mi impute pmm</code>                   | 20000        |     |        | 200 | 1      | 1                 |
| <code>mi impute regress</code>               | 40000        |     |        | 100 | 1      | 1                 |
| <code>misstable nested</code>                | 2000000      |     |        | 20  |        | 1                 |
| <code>misstable patterns</code>              | 2000000      |     |        | 20  |        | 1                 |
| <code>misstable summarize</code>             | 5000         |     |        | 10  |        | 1                 |
| <code>misstable tree</code>                  | 1000000      |     |        | 20  |        | 1                 |
| <code>mixed</code>                           |              | 500 | 10     | 5   | 5      | 1                 |
| <code>mixed (crossed effects)</code>         |              | 10  | 1000   |     |        | 1                 |
| <code>mkspline</code>                        | 12000000     |     |        | 1   |        | 1                 |
| <code>mleval</code>                          | 30000000     |     |        | 200 |        | 1                 |
| <code>mleval, nocons</code>                  | 30000000     |     |        | 200 |        | 1                 |
| <code>mlmatbysum</code>                      | 20000000     |     |        | 200 | 160000 | 1                 |
| <code>mlmatsum</code>                        | 20000000     |     |        | 200 |        | 1                 |
| <code>mlogit</code>                          | 500000       |     |        | 100 | 3      | 1                 |
| <code>mlsum</code>                           | 4.0e+08      |     |        | 1   |        | 1                 |
| <code>mlvecsum</code>                        | 20000000     |     |        | 400 |        | 1                 |
| <code>mprobit</code>                         | 800          |     |        | 10  | 3      | 1                 |
| <code>mswitch ar</code>                      |              | 100 | 100    | 20  | 5      | 1                 |
| <code>mswitch dr</code>                      |              | 100 | 100    | 20  | 5      | 1                 |
| <code>mvdecode</code>                        | 500000       |     |        | 20  | 1000   | 1                 |
| <code>mvencode</code>                        | 6000000      |     |        | 20  | 1000   | 1                 |
| <code>mvreg</code>                           | 2000000      |     |        | 100 | 3      | 1                 |
| <code>mvtest correlations</code>             |              | 2   | 600000 | 100 |        | 1                 |
| <code>mvtest covariances</code>              |              | 2   | 600000 | 100 |        | 1                 |
| <code>mvtest means,<br/>heterogeneous</code> |              | 2   | 400000 | 100 |        | 1                 |

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.

Table 3. Problem sizes

| Command                   | Observations |      |         | $k$ | $d_1$ | $n_{\text{iter}}$ |
|---------------------------|--------------|------|---------|-----|-------|-------------------|
|                           | $N$          | $m$  | $t$     |     |       |                   |
| mvtest means, homogeneous |              | 2    | 150000  | 100 |       | 1                 |
| mvtest means, lr          |              | 2    | 500000  | 100 |       | 1                 |
| mvtest normality          | 1000         |      |         | 20  |       | 1                 |
| nbreg                     | 60000        |      |         | 200 |       | 1                 |
| newey                     | 500000       |      |         | 5   |       | 1                 |
| nl                        | 1500000      |      |         |     |       | 1                 |
| nlogit                    |              | 1200 | 2       | 2   | 3     | 1                 |
| nlsur                     | 100000       |      |         | 2   |       | 1                 |
| npregress kernel          | 1000         |      |         | 2   |       | 1                 |
| nptrend                   | 300000       |      |         | 10  |       | 1                 |
| ologit                    | 700000       |      |         | 100 | 3     | 1                 |
| ologit, vce(cluster)      | 300000       |      |         | 100 | 3     | 1                 |
| ologit, vce(robust)       | 700000       |      |         | 100 | 3     | 1                 |
| oneway                    | 3000000      |      |         | 200 |       | 1                 |
| oprobit                   | 200000       |      |         | 200 | 3     | 1                 |
| oprobit, vce(cluster)     | 100000       |      |         | 200 | 3     | 1                 |
| oprobit, vce(robust)      | 200000       |      |         | 200 | 3     | 1                 |
| orthog                    | 1000000      |      |         | 10  |       | 1                 |
| pca                       | 600000       |      |         | 100 |       | 1                 |
| pcorr                     | 1300000      |      |         | 200 |       | 1                 |
| pctile                    | 16000000     |      |         | 1   |       | 1                 |
| pergram                   | 10000        |      |         | 1   |       | 1                 |
| pkcollapse                |              | 100  | 50      |     |       | 1                 |
| pkexamine                 |              | 1    | 1000000 |     |       | 1                 |
| pksumm                    |              | 200  | 10      |     |       | 1                 |

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.

Table 3. Problem sizes

| Command               | Observations |     |     | $k$  | $d_1$ | $n_{\text{iter}}$ |
|-----------------------|--------------|-----|-----|------|-------|-------------------|
|                       | $N$          | $m$ | $t$ |      |       |                   |
| poisson               | 200000       |     |     | 200  |       | 1                 |
| poisson, vce(cluster) | 100000       |     |     | 200  |       | 1                 |
| poisson, exposure()   | 200000       |     |     | 200  |       | 1                 |
| poisson, vce(robust)  | 200000       |     |     | 200  |       | 1                 |
| pologit               | 100000       |     |     | 40   |       | 1                 |
| popoisson             | 100000       |     |     | 40   |       | 1                 |
| poregress             | 100000       |     |     | 40   |       | 1                 |
| pperron               | 300000       |     |     | 1    |       | 1                 |
| prais                 | 1000000      |     |     | 5    |       | 1                 |
| predict, cooksd       | 600000       |     |     | 300  |       | 1                 |
| predict, covratio     | 600000       |     |     | 300  |       | 1                 |
| predict, dfbeta       | 400000       |     |     | 200  |       | 1                 |
| predict, dfits        | 600000       |     |     | 200  |       | 1                 |
| predict, e            | 3000000      |     |     | 1000 |       | 1                 |
| predict, leverage     | 1200000      |     |     | 200  |       | 1                 |
| predict, pr           | 2500000      |     |     | 1000 |       | 1                 |
| predict, residuals    | 6000000      |     |     | 1000 |       | 1                 |
| predict, rstandard    | 400000       |     |     | 400  |       | 1                 |
| predict, rstudent     | 400000       |     |     | 400  |       | 1                 |
| predict, stdf         | 1600000      |     |     | 200  |       | 1                 |
| predict, stdp         | 400000       |     |     | 400  |       | 1                 |
| predict, stdr         | 400000       |     |     | 400  |       | 1                 |
| predict, welsch       | 300000       |     |     | 300  |       | 1                 |
| predict, ystar        | 3000000      |     |     | 1000 |       | 1                 |
| predictnl             | 60000        |     |     | 200  |       | 1                 |

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.



Table 3. Problem sizes

| Command                     | Observations |       |     | $k$  | $d_1$ | $n_{\text{iter}}$ |
|-----------------------------|--------------|-------|-----|------|-------|-------------------|
|                             | $N$          | $m$   | $t$ |      |       |                   |
| probit                      | 500000       |       |     | 200  |       | 1                 |
| procrustes                  | 200000       |       |     | 50   | 50    | 1                 |
| proportion                  | 300000       |       |     | 10   | 5     | 1                 |
| prtest1                     | 20000000     |       |     | 1    | 2     | 3                 |
| prtest2                     | 20000000     |       |     | 2    | 2     | 2                 |
| prtest, by()                | 10000000     |       |     | 2    | 2     | 1                 |
| pwcorr                      | 30000000     |       |     | 3    |       | 1                 |
| qreg                        | 100000       |       |     | 20   |       | 1                 |
| ranksum                     | 4000000      |       |     | 2    |       | 1                 |
| ratio                       | 8000000      |       |     |      |       | 1                 |
| ratio (exp1) (exp2)         | 9000000      |       |     |      |       | 1                 |
| recode                      | 1500000      |       |     | 5    | 5     | 1                 |
| reg3                        | 90000        |       |     | 100  | 3     | 1                 |
| regress                     | 3000000      |       |     | 180  |       | 1                 |
| regress, vce(cluster)       | 1500000      |       |     | 180  |       | 1                 |
| regress, vce(robust)        | 300000       |       |     | 180  |       | 1                 |
| replace                     | 15000000     |       |     |      |       | 1                 |
| replace (small expressions) | 150000       |       |     | 4000 |       | 1                 |
| reshape long                |              | 50000 | 20  |      |       | 1                 |
| reshape wide                |              | 50000 | 15  | 5    |       | 1                 |
| robvar                      | 200000       |       |     | 2    |       | 1                 |
| rocfit                      | 100000       |       |     | 1    | 5     | 1                 |
| roctab                      | 600000       |       |     | 1    | 20    | 1                 |
| rotate                      | 10000        |       |     | 80   |       | 1                 |
| rotatemat                   | 80           |       |     | 80   |       | 1                 |

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.

Table 3. Problem sizes

| Command                 | Observations |     |     | $k$ | $d_1$ | $n_{\text{iter}}$ |
|-------------------------|--------------|-----|-----|-----|-------|-------------------|
|                         | $N$          | $m$ | $t$ |     |       |                   |
| rreg                    | 100000       |     |     | 200 |       | 1                 |
| runtest                 | 6000000      |     |     | 1   |       | 1                 |
| scobit                  | 120000       |     |     | 200 |       | 1                 |
| scoreplot               | 400000       |     |     | 20  |       | 1                 |
| screeplot               | 10000000     |     |     | 20  |       | 1                 |
| sdtest1                 | 24000000     |     |     |     |       | 3                 |
| sdtest2                 | 12000000     |     |     | 2   |       | 3                 |
| sdtest, by()            | 9000000      |     |     |     |       | 2                 |
| sem, method(adf) (CFA)  | 150000       |     |     | 5   | 3     | 1                 |
| sem, method(ml) (CFA)   | 2500000      |     |     | 10  | 3     | 1                 |
| sem, method(mlmv) (CFA) | 100000       |     |     | 4   | 3     | 1                 |
| sem (SEM latent)        | 10000000     |     |     | 4   | 3     | 1                 |
| sem (SEM observed)      | 5000000      |     |     | 20  | 3     | 1                 |
| separate                | 1000000      |     |     | 100 | 4     | 1                 |
| sfrancia                | 1000000      |     |     | 2   |       | 1                 |
| signrank                | 2500000      |     |     | 2   |       | 1                 |
| signtest                | 1.0e+08      |     |     | 2   |       | 1                 |
| sktest                  | 6000000      |     |     | 2   |       | 1                 |
| slogit                  | 20000        |     |     | 10  | 5     | 1                 |
| sort                    | 9000000      |     |     | 10  |       | 1                 |
| spearman                | 400000       |     |     | 3   |       | 1                 |
| sspace                  | 5000         |     |     | 20  |       | 1                 |
| stack                   | 500000       |     |     | 100 |       | 1                 |
| stci                    | 200000       |     |     | 1   |       | 1                 |
| stcox                   | 250000       |     |     | 10  |       | 1                 |

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.



Table 3. Problem sizes

| Command                   | Observations |     |     | $k$ | $d_1$ | $n_{\text{iter}}$ |
|---------------------------|--------------|-----|-----|-----|-------|-------------------|
|                           | $N$          | $m$ | $t$ |     |       |                   |
| stcrreg                   | 2000         |     |     | 5   |       | 1                 |
| stgen                     | 30000000     |     |     | 2   |       | 1                 |
| stintreg, d(exponential)  | 500000       |     |     | 30  |       | 1                 |
| stintreg, d(weibull)      | 200000       |     |     | 20  |       | 1                 |
| stir                      | 4500000      |     |     | 1   | 2     | 1                 |
| stmcc                     | 900000       |     |     |     |       | 1                 |
| by: stmcc                 | 600000       |     |     |     | 50    | 1                 |
| stmh                      | 1500000      |     |     |     |       | 1                 |
| by: stmh                  | 1500000      |     |     |     | 10    | 1                 |
| stptime                   | 9000000      |     |     | 1   | 60000 | 1                 |
| strate                    | 1000000      |     |     | 1   | 5     | 1                 |
| streg,                    | 600000       |     |     | 100 |       | 1                 |
| distribution(exponential) |              |     |     |     |       |                   |
| streg, dist(exp)          | 200000       |     |     | 200 | 1000  | 1                 |
| vce(cluster)              |              |     |     |     |       |                   |
| streg, dist(exp)          | 60000        |     |     | 200 |       | 1                 |
| frailty()                 |              |     |     |     |       |                   |
| streg, dist(exp)          | 200000       |     |     | 100 | 1000  | 1                 |
| frailty() shared()        |              |     |     |     |       |                   |
| streg, dist(exp)          | 200000       |     |     | 200 |       | 1                 |
| vce(robust)               |              |     |     |     |       |                   |
| streg,                    | 100000       |     |     | 200 |       | 1                 |
| distribution(ggamma)      |              |     |     |     |       |                   |
| streg, dist(ggamma)       | 200000       |     |     | 200 | 1000  | 1                 |
| vce(cluster)              |              |     |     |     |       |                   |
| streg, dist(ggamma)       | 200000       |     |     | 200 |       | 1                 |
| vce(robust)               |              |     |     |     |       |                   |
| streg,                    | 200000       |     |     | 50  |       | 1                 |
| distribution(gompertz)    |              |     |     |     |       |                   |
| streg, dist(gompertz)     | 200000       |     |     | 50  | 1000  | 1                 |
| vce(cluster)              |              |     |     |     |       |                   |
| streg, dist(gompertz)     | 200000       |     |     | 50  |       | 1                 |
| frailty()                 |              |     |     |     |       |                   |
| streg, dist(gomp)         | 200000       |     |     | 10  | 1000  | 1                 |
| frailty() shared()        |              |     |     |     |       |                   |
| streg, dist(gompertz)     | 200000       |     |     | 50  |       | 1                 |
| vce(robust)               |              |     |     |     |       |                   |
| streg,                    | 600000       |     |     | 100 |       | 1                 |
| distribution(llogistic)   |              |     |     |     |       |                   |

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.



Table 3. Problem sizes

| Command                                  | Observations |     |     | $k$ | $d_1$ | $n_{\text{iter}}$ |
|------------------------------------------|--------------|-----|-----|-----|-------|-------------------|
|                                          | $N$          | $m$ | $t$ |     |       |                   |
| streg, dist(llogistic)<br>vce(cluster)   | 200000       |     |     | 200 | 1000  | 1                 |
| streg, dist(llogistic)<br>frailty()      | 60000        |     |     | 200 |       | 1                 |
| streg, dist(llog)<br>frailty() shared()  | 200000       |     |     | 100 | 1000  | 1                 |
| streg, dist(llogistic)<br>vce(robust)    | 200000       |     |     | 200 |       | 1                 |
| streg,<br>distribution(lnormal)          | 200000       |     |     | 100 |       | 1                 |
| streg, dist(lnormal)<br>vce(cluster)     | 200000       |     |     | 200 | 1000  | 1                 |
| streg, dist(lnormal)<br>frailty()        | 60000        |     |     | 200 |       | 1                 |
| streg, dist(lnorm)<br>frailty() shared() | 200000       |     |     | 10  | 1000  | 1                 |
| streg, dist(lnormal)<br>vce(robust)      | 200000       |     |     | 200 |       | 1                 |
| streg,<br>distribution(weibull)          | 200000       |     |     | 200 |       | 1                 |
| streg, dist(weibull)<br>vce(cluster)     | 200000       |     |     | 200 | 1000  | 1                 |
| streg, dist(weibull)<br>frailty()        | 200000       |     |     | 50  |       | 1                 |
| streg, dist(weib)<br>frailty() shared()  | 100000       |     |     | 100 | 1000  | 1                 |
| streg, dist(weibull)<br>vce(robust)      | 200000       |     |     | 200 |       | 1                 |
| sts generate                             | 1000000      |     |     | 1   |       | 1                 |
| sts graph                                | 1000000      |     |     | 1   |       | 1                 |
| sts list                                 | 3000000      |     |     | 1   |       | 1                 |
| sts test                                 | 1000000      |     |     | 1   | 2     | 1                 |
| stset                                    | 3000000      |     |     |     |       | 1                 |
| stsplrit                                 | 2000000      |     |     |     | 50    | 1                 |
| stsum                                    | 200000       |     |     | 1   |       | 1                 |
| stteffects ipw (weibull)                 | 50000        |     |     | 50  |       | 1                 |
| stteffects ipwra<br>(weibull)            | 20000        |     |     | 20  |       | 1                 |
| stteffects ra (weibull)                  | 10000        |     |     | 50  |       | 1                 |
| stteffects wra (weibull)                 | 10000        |     |     | 50  |       | 1                 |

See appendix C for command descriptions.

Revision 3.2.0—15aug2019

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.



Table 3. Problem sizes

| Command                 | Observations |     |      | $k$  | $d_1$ | $n_{\text{iter}}$ |
|-------------------------|--------------|-----|------|------|-------|-------------------|
|                         | $N$          | $m$ | $t$  |      |       |                   |
| stvary                  | 3000000      |     |      | 5    |       | 1                 |
| suest                   | 400000       |     |      | 200  |       | 1                 |
| summarize               | 4500000      |     |      | 200  |       | 1                 |
| sunflower               | 1000000      |     |      | 2    |       | 1                 |
| sureg                   | 300000       |     |      | 100  | 2     | 1                 |
| svar                    | 40000        |     |      | 2    | 10    | 1                 |
| svmat                   | 3000         |     |      | 3000 |       | 1                 |
| svy brr: logit          |              | 128 | 200  | 20   |       | 1                 |
| svy brr: poisson        |              | 16  | 4000 | 20   |       | 1                 |
| svy brr: regress        |              | 16  | 6000 | 200  |       | 1                 |
| svy jackknife: logit    |              | 5   | 400  | 20   | 20    | 1                 |
| svy jackknife: poisson  |              | 5   | 300  | 20   | 20    | 1                 |
| svy jackknife: regress  |              | 3   | 3000 | 10   | 20    | 1                 |
| svy linearized: logit   | 200000       |     |      | 200  |       | 1                 |
| svy linearized: poisson | 200000       |     |      | 200  |       | 1                 |
| svy linearized: regress | 400000       |     |      | 200  |       | 1                 |
| swilk                   | 150000       |     |      | 20   |       | 1                 |
| symmetry                | 800000       |     |      | 2    | 50    | 1                 |
| table (one-way)         | 4000000      |     |      | 20   |       | 1                 |
| table (two-way)         | 3000000      |     |      | 20   |       | 1                 |
| tabodds                 | 300000       |     |      |      | 20    | 1                 |
| tabodds (adjusted)      | 50000        |     |      | 10   | 20    | 1                 |
| tabstat                 | 2000000      |     |      | 50   |       | 1                 |
| tabstat, by()           | 2000000      |     |      | 20   |       | 1                 |
| tabulate (one-way)      | 6000000      |     |      | 20   |       | 1                 |

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.

Table 3. Problem sizes

| Command                                | Observations |       |     | $k$ | $d_1$ | $n_{\text{iter}}$ |
|----------------------------------------|--------------|-------|-----|-----|-------|-------------------|
|                                        | $N$          | $m$   | $t$ |     |       |                   |
| tabulate (two-way)                     | 10000000     |       |     | 20  |       | 1                 |
| teffects aipw (linear)                 | 10000        |       |     | 50  |       | 1                 |
| teffects aipw (probit)                 | 10000        |       |     | 50  |       | 1                 |
| teffects ipw (logit)                   | 20000        |       |     | 100 |       | 1                 |
| teffects ipwra (linear)                | 10000        |       |     | 50  |       | 1                 |
| teffects ipwra (probit)                | 10000        |       |     | 50  |       | 1                 |
| teffects nnmatch                       | 20000        |       |     | 100 |       | 1                 |
| teffects psmatch, logit                | 10000        |       |     | 50  |       | 1                 |
| teffects ra (linear)                   | 10000        |       |     | 100 |       | 1                 |
| teffects ra (probit)                   | 10000        |       |     | 100 |       | 1                 |
| tetrachoric                            | 1200000      |       |     | 4   | 2     | 1                 |
| threshold, threshvar()                 | 1000         |       |     | 20  | 0     | 1                 |
| threshold, threshvar()<br>regionvars() | 1000         |       |     | 10  | 10    | 1                 |
| tnbreg                                 | 300000       |       |     | 10  |       | 1                 |
| tobit                                  | 300000       |       |     | 200 |       | 1                 |
| tostring                               |              | 10000 | 200 |     |       | 1                 |
| total                                  | 600000       |       |     | 200 |       | 1                 |
| tpoisson                               | 1000000      |       |     | 50  |       | 1                 |
| truncreg                               | 150000       |       |     | 200 |       | 1                 |
| tsfilter bk                            | 1000000      |       |     | 1   |       | 1                 |
| tsfilter bw                            | 1500         |       |     | 1   |       | 1                 |
| tsfilter cf                            | 1000000      |       |     | 1   |       | 1                 |
| tsfilter hp                            | 1500         |       |     | 1   |       | 1                 |
| tsrevar                                | 1100000      |       |     | 20  |       | 1                 |
| tsset                                  | 4000000      |       |     |     |       | 1                 |

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.

Table 3. Problem sizes

| Command             | Observations |     |     | $k$ | $d_1$ | $n_{\text{iter}}$ |
|---------------------|--------------|-----|-----|-----|-------|-------------------|
|                     | $N$          | $m$ | $t$ |     |       |                   |
| tssmooth exp        | 1000000      |     |     | 1   |       | 1                 |
| tssmooth ma         | 1000000      |     |     | 1   |       | 1                 |
| ttest1              | 15000000     |     |     | 1   |       | 5                 |
| ttest2              | 35000000     |     |     | 2   |       | 1                 |
| ttest, by()         | 20000000     |     |     |     |       | 1                 |
| twoway fpfit        | 400000       |     |     | 1   |       | 1                 |
| twoway lfitci       | 6000000      |     |     | 1   |       | 1                 |
| twoway mband        | 3000000      |     |     | 1   |       | 1                 |
| twoway mspline      | 4000000      |     |     | 1   |       | 1                 |
| ucm, model(rwdrift) | 5000         |     |     | 3   |       | 1                 |
| var                 | 250000       |     |     | 2   | 5     | 1                 |
| vargranger          | 4000000      |     |     | 2   | 5     | 5                 |
| varlmar             | 80000        |     |     | 2   | 5     | 1                 |
| varnorm             | 300000       |     |     | 2   | 5     | 1                 |
| varsoc              | 200000       |     |     | 2   | 5     | 1                 |
| varstable           | 4000000      |     |     | 2   | 10    | 5                 |
| vec                 | 30000        |     |     | 2   | 10    | 1                 |
| veclmar             | 50000        |     |     | 2   | 5     | 1                 |
| vecnorm             | 150000       |     |     | 2   | 5     | 1                 |
| vecrank             | 200000       |     |     | 2   | 5     | 1                 |
| vecstable           | 1000000      |     |     | 2   | 10    | 1                 |
| vwls                | 1000000      |     |     | 200 |       | 1                 |
| wntestb             | 10000        |     |     | 1   |       | 1                 |
| wntestq             | 400000       |     |     | 1   |       | 1                 |
| xcorr               | 400000       |     |     | 1   |       | 1                 |

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.

Table 3. Problem sizes

| Command                              | Observations |        |        | $k$ | $d_1$ | $n_{\text{iter}}$ |
|--------------------------------------|--------------|--------|--------|-----|-------|-------------------|
|                                      | $N$          | $m$    | $t$    |     |       |                   |
| xpologit                             | 10000        |        |        | 40  |       | 1                 |
| xpopoisson                           | 10000        |        |        | 40  |       | 1                 |
| xporegress                           | 10000        |        |        | 40  |       | 1                 |
| xtabond                              |              | 100000 | 10     | 2   |       | 1                 |
| xtabond, twostep                     |              | 100000 | 10     | 2   |       | 1                 |
| xtcloglog, re                        |              | 20000  | 5      | 5   |       | 1                 |
| xtdata, be                           |              | 15000  | 5      | 200 |       | 1                 |
| xtdata, fe                           |              | 500000 | 5      | 5   |       | 1                 |
| xtdata, re                           |              | 300000 | 5      | 5   |       | 1                 |
| xtdpd                                |              | 40000  | 5      | 5   |       | 1                 |
| xtdpdsys                             |              | 60000  | 5      | 5   |       | 1                 |
| xtereg                               |              | 1000   | 5      | 1   | 1     | 1                 |
| xtfrontier                           |              | 4000   | 10     | 50  |       | 1                 |
| xtgee, family(gaussian)<br>corr(ar2) |              | 50000  | 5      | 10  |       | 1                 |
| xtgee, fam(gauss)<br>corr(unstruct)  |              | 60000  | 5      | 10  |       | 1                 |
| xtcloglog, pa                        |              | 100000 | 5      | 5   |       | 1                 |
| xtlogit, pa                          |              | 100000 | 5      | 5   |       | 1                 |
| xtnbreg, pa                          |              | 80000  | 5      | 5   |       | 1                 |
| xtpoisson, pa                        |              | 30000  | 10     | 5   |       | 1                 |
| xtprobit, pa                         |              | 60000  | 10     | 5   |       | 1                 |
| xtreg, pa                            |              | 100000 | 5      | 10  |       | 1                 |
| xtgls                                |              | 5      | 200000 | 5   |       | 1                 |
| xthtaylor                            |              | 100000 | 10     | 4   | 4     | 1                 |
| xtile                                | 100000       |        |        |     |       | 1                 |
| xtintreg                             |              | 15000  | 5      | 5   |       | 1                 |

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.

Table 3. Problem sizes

| Command                               | Observations |        |       | $k$ | $d_1$ | $n_{\text{iter}}$ |
|---------------------------------------|--------------|--------|-------|-----|-------|-------------------|
|                                       | $N$          | $m$    | $t$   |     |       |                   |
| xtivreg, be                           |              | 120000 | 5     | 5   | 5     | 1                 |
| xtivreg, fd                           |              | 80000  | 5     | 5   | 5     | 1                 |
| xtivreg, fe                           |              | 80000  | 5     | 5   | 5     | 1                 |
| xtivreg, re                           |              | 150000 | 5     | 5   | 5     | 1                 |
| xtlogit, fe                           |              | 20000  | 10    | 50  |       | 1                 |
| xtlogit, re                           |              | 40000  | 5     | 5   |       | 1                 |
| xtnbreg, fe                           |              | 70000  | 5     | 10  |       | 1                 |
| xtnbreg, re                           |              | 40000  | 5     | 10  |       | 1                 |
| xtologit                              |              | 8000   | 10    | 10  | 0     | 1                 |
| xtoprobit                             |              | 8000   | 10    | 10  | 0     | 1                 |
| xtpcse                                |              | 3      | 80000 | 50  |       | 1                 |
| xtpoisson, fe                         |              | 20000  | 5     | 50  |       | 1                 |
| xtpoisson, re                         |              | 30000  | 5     | 50  |       | 1                 |
| xtprobit, re                          |              | 20000  | 5     | 5   |       | 1                 |
| xtrc                                  |              | 100    | 10000 | 5   |       | 1                 |
| xtreg, be                             |              | 15000  | 5     | 200 |       | 1                 |
| xtreg, fe                             |              | 200000 | 5     | 100 |       | 1                 |
| xtreg, fe vce(robust)                 |              | 50000  | 10    | 100 |       | 1                 |
| xtreg, mle                            |              | 80000  | 10    | 5   |       | 1                 |
| xtreg, re                             |              | 20000  | 3     | 200 |       | 1                 |
| xtregar, fe                           |              | 100000 | 5     | 2   |       | 1                 |
| xtregar, re                           |              | 90000  | 5     | 2   |       | 1                 |
| xtset                                 |              | 500    | 5000  |     |       | 1                 |
| xtstreg,<br>distribution(exponential) |              | 8000   | 10    | 10  | 0     | 1                 |
| xtstreg,<br>distribution(weibull)     |              | 8000   | 10    | 10  | 0     | 1                 |

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.

Table 3. Problem sizes

| Command                          | Observations |        |      | $k$  | $d_1$ | $n_{\text{iter}}$ |
|----------------------------------|--------------|--------|------|------|-------|-------------------|
|                                  | $N$          | $m$    | $t$  |      |       |                   |
| <code>xtsum</code>               |              | 100000 | 10   | 10   |       | 1                 |
| <code>xttab</code>               | 1500000      |        |      | 2    | 50    | 1                 |
| <code>xttobit</code>             |              | 50000  | 5    | 5    |       | 1                 |
| <code>xtunitroot breitung</code> |              | 200    | 3000 |      |       | 1                 |
| <code>xtunitroot fisher</code>   |              | 50     | 1000 |      |       | 1                 |
| <code>xtunitroot hadri</code>    |              | 50     | 1000 |      |       | 1                 |
| <code>xtunitroot ht</code>       |              | 300    | 2000 |      |       | 1                 |
| <code>xtunitroot ips</code>      |              | 1000   | 20   |      |       | 1                 |
| <code>xtunitroot llc</code>      |              | 100    | 500  |      |       | 1                 |
| <code>zinb</code>                | 150000       |        |      | 50   | 50    | 1                 |
| <code>zioprobit</code>           | 200000       |        |      | 30   | 30    | 1                 |
| <code>zip</code>                 | 250000       |        |      | 50   | 50    | 1                 |
| <code>_predict, xb</code>        | 5000000      |        |      | 1000 |       | 1                 |
| <code>_rmcoll</code>             | 6000000      |        |      | 100  |       | 1                 |
| <code>_robust</code>             | 3000000      |        |      | 200  |       | 1                 |

$N$ , number of observations;  $m$ , number of panels;  $t$ , number of time periods within each panel;  $k$ , number of regressors;  $d_1$ , miscellaneous dimension; and  $n_{\text{iter}}$ , number of iterations.

## E Commands not assessed

Some commands were not explicitly assessed and thus do not appear in table 1 or in the performance graphs in appendix A. These commands fall into several categories, as detailed below.

The presented results for `generate` and `replace` apply for floating point variables only. Some other uses of `generate` and `replace` are not parallelized, so their performance was not assessed. Specifically, for string (`str#` and `strL`) variables, `generate` is not parallelized. For string (`str#` and `strL`) and integer (`byte`, `int`, and `long`) variables, `replace` is not parallelized. In these cases, `generate` and `replace` will promote the variable type when the expression returns a value that does not fit in the current storage type. Type promotion is a global change to the variable and can be triggered at any observation for these commands, so parallelization is not possible.

Replication-based prefix commands, such as `bootstrap`, `fracpoly`, `jackknife`, `mfp`, `permute`, `rolling`, `simulate`, `statsby`, and `stepwise`, were not explicitly assessed. These commands run another target command repeatedly; to the extent the target command's performance is improved for a particular problem size, a similar improvement will be obtained when it is run repeatedly by the prefix command.

Commands that do not process data or otherwise involve lengthy computations and are therefore inherently fast are not parallelized and so their performance was not assessed. These commands include `camat`, `clear`, `clonevar`, `confirm`, `describe`, `estat`, `estimates`, `factormat`, `fvexpand`, `fvunab`, `lincom`, `nlcom`, `pcamat`, `roccomp`, `rocgold`, `sampsi`, `search`, `stpower`, `svydes`, `test`, `testnl`, `unabbrev`, and `varabbrev`.

Commands that involve file I/O or Internet access are not parallelized and so were not assessed. These include `adoupdate`, `append`, `cf`, `fdadefine`, `fdasave`, `fdause`, `filefilter`, `hsearch`, `icd9`, `icd10`, `infile`, `insheet`, `merge`, `odbc`, `outfile`, `outsheet`, `rmdir`, `save`, `search`, `snapshot`, `use`, `xmlsave`, `xmluse`, `zipfile`, and `unzipfile`.

Only a subset of prediction options were assessed. If all predictions were included, they would unduly dominate the timings. Most other predictions have performances similar to the predictions presented in table 1 and in appendix A. Two prediction-like commands whose results are not obtained from `predict` but whose timings are similar to `predict` are `fracpred` and `dfbeta`.

Some commands are partially parallelized, but their degree of parallelization is extremely variable with respect to the size and characteristics of the data. These commands were not assessed and include `bcskew0`, `lnskew0`, `fracplot`, `fracgen`, `mkmat`, `stbase`, and `stjoin`.

`ac` and `pac` are two time-series commands that are not parallelized and so their performance was not assessed.

`graph twoway` is not parallelized although a few of its plottypes that involve data management or estimation are parallelized, such as `histogram`, `lowess`, `lfit`, and `qfit`. Most statistical graphs in Stata are based on `graph twoway`. Graphs that involve data management or estimation were assessed and appear in table 1 and appendix A. Graphs that do not involve data management or estimation are not parallelized and so their performance was not assessed. These include `acprplot`, `avplot`, `avplots`, `cabiplot`, `caprojection`, `cchart`, `cluster tree`, `cprplot`, `graph twoway`, `lvr2plot`, `mdsshepard`,

`pchart`, `procoverlay`, `qbys`, `qchi`, `qnorm`, `qqplot`, `quantile`, `rchart`, `rocplot`, `rvfplot`, `rvpplot`, `shewhart`, `spikeplt`, `stcoxkm`, `stcurve`, `stphplot`, and `sympplot`.

A number of commands perform similarly to related commands that were assessed, but these commands were not themselves assessed. `bsqreg`, `iqreg`, and `sqreg` perform similarly to `qreg`. `gladder` and `qladder` perform similarly to `ladder`. `gnbreg` is similar to `nbreg`. `xttrans` is similar to `xttab`.



## F Mata

Mata is Stata's optimized matrix programming language. It is fully integrated with every aspect of Stata. Some parts of Mata are parallelized and some parts are not. As with Stata, you do not need to change anything to obtain the parallelization speedups; they are automatic.

Those parts of Mata that are parallelized are fully parallelized, meaning that on large enough problems, their speedups will be close to the best theoretical speedups discussed in section 6.

The following Mata functions are parallelized: `Cofc()`, `Cofd()`, `F()`, `Fden()`, `Ftail()`, `acos()`, `arg()`, `asin()`, `atan()`, `atan2()`, `betaden()`, `binomial()`, `binomialtail()`, `binormal()`, `ceil()`, `chi2()`, `chi2den()`, `chi2tail()`, `cofC()`, `cofd()`, `comb()`, `cos()`, `cross()`, `crossdev()`, `day()`, `dgamma``mapda()`, `dgamma``mapdada()`, `dgamma``mapdadx()`, `dgamma``mapdx()`, `dgamma``mapdx``dx()`, `digamma()`, `dofC()`, `dofc()`, `dofh()`, `dofm()`, `dofq()`, `dofw()`, `dofy()`, `dow()`, `doy()`, `dunnettprob()`, `exp()`, `exponential()`, `exponentialden()`, `exponentialtail()`, `factorial()`, `floatround()`, `floor()`, `gammaden()`, `gamma``map()`, `gamma``maptail()`, `halfyear()`, `hh()`, `hhC()`, `hofd()`, `hours()`, `ibeta()`, `ibetatail()`, `invF()`, `invFtail()`, `invbinomial()`, `invbinomialtail()`, `invchi2()`, `invchi2tail()`, `invdunnettprob()`, `invexponential()`, `invexponentialtail()`, `invgamma``map()`, `invgamma``maptail()`, `invibeta()`, `invibetatail()`, `invlogistic()`, `invlogistic``tail()`, `invnF()`, `invnFtail()`, `invnchi2()`, `invnibeta()`, `invnormal()`, `invnt()`, `invnttail()`, `invt()`, `invttail()`, `invtukeyprob()`, `invweibull()`, `invweibullph()`, `invweibullphtail()`, `invweibulltail()`, `ln()`, `lnfactorial()`, `lngamma()`, `lnigammaden()`, `lnnormal()`, `lnnormalden()`, `logistic()`, `logisticden()`, `logistic``tail()`, `mdy()`, `minutes()`, `mm()`, `mmC()`, `mod()`, `mofd()`, `month()`, `msofhours()`, `msofminutes()`, `msofseconds()`, `nF()`, `nFden()`, `nFtail()`, `nbetaden()`, `nchi2()`, `nibeta()`, `normal()`, `normalden()`, `npnF()`, `npnchi2()`, `npnt()`, `nt()`, `ntden()`, `nttail()`, `qofd()`, `quadcross()`, `quadcrossdev()`, `quarter()`, `round()`, `seconds()`, `sin()`, `sqrt()`, `ss()`, `st_data()`, `t()`, `tan()`, `tden()`, `trigamma()`, `trunc()`, `ttail()`, `tukeyprob()`, `week()`, `weibull()`, `weibullden()`, `weibullph()`, `weibullphden()`, `weibullphtail()`, `weibulltail()`, `wofd()`, `year()`, `yh()`, `ym()`, `yq()`, and `yw()`.

In addition, matrix multiplication in Mata is fully parallelized, as are Mata's colon operators for performing elementwise computations. All other parts of Mata are either not parallelized or are functions of a mixture of the two.

## G GLLAMM

Table 4 below shows results for a few models fit using `gllamm`. This is but a small subset of the models that `gllamm` can fit. Each command is described briefly in table 5.

The user-written command `gllamm` (generalized linear latent and mixed models) adds to Stata the ability to fit multilevel, mixed, or hierarchical regression models that have continuous, count, binary, or ordinal dependent variables. In addition, the model may have latent (unobserved) variables, endogenous covariates, and random coefficients or intercepts at any level. Among the many models that `gllamm` can fit, some important special cases include generalized linear mixed models, multilevel regression models, factor models, item response models, structural equation models, latent-class models, generalized linear models with covariate measurement error, endogenous switching and sample selection models, and Rasch models (including multidimensional marginally sufficient Rasch models).

`gllamm`'s authors, Sophia Rabe-Hesketh with contributions from Anders Skrondal and Andrew Pickles, maintain a web site—<http://www.gllamm.org/>—with complete documentation (140 pages), tutorials, worked examples, wrapper commands to ease estimation of special models, dates of upcoming courses on `gllamm`, and references (often with links) to more than 150 papers published on using `gllamm` to fit models.

`gllamm` uses full maximum likelihood to estimate the parameters of models and uses Gauss–Hermite quadrature or adaptive quadrature to evaluate the integrals of the likelihood. This common computation engine is one reason `gllamm` is so flexible and can fit so many models. It is, however, exceedingly computationally intensive, with the effect that `gllamm` can require substantial time to fit models. `gllamm` users are interested in seeing it run faster.

`gllamm` uses many Stata commands that have been parallelized, and some of `gllamm`'s algorithms, sections of which have been parallelized, are written in C. Even so, `gllamm` incorporates many algorithms, and these algorithms are triggered differently when fitting different models. It is difficult to say anything definitive about performance gains for `gllamm` when run under Stata/MP. Many `gllamm` models are highly parallelized, some not parallelized at all, and others lie somewhere in between.

Table 4. Stata/MP performance, command by command

| Command                     | Speed relative to a single core <sup>a</sup> |     |     |     | Percentage<br>parallelized <sup>b</sup> |
|-----------------------------|----------------------------------------------|-----|-----|-----|-----------------------------------------|
|                             | Number of cores                              |     |     |     |                                         |
|                             | 2                                            | 4   | 8   | 16  |                                         |
| Finite mixture model        | 2.3                                          | 3.5 | 4.5 | 5.6 | 86                                      |
| Item response model         | 1.4                                          | 2.1 | 2.7 | 3.1 | 74                                      |
| Latent class model          | 1.3                                          | 1.9 | 2.5 | 2.9 | 71                                      |
| Measurement error model     | 1.8                                          | 2.7 | 3.8 | 5.1 | 86                                      |
| Rank-outcome latent class   | 1.8                                          | 2.6 | 3.2 | 3.8 | 77                                      |
| MIMIC model                 | 1.2                                          | 1.8 | 2.5 | 2.9 | 72                                      |
| Random-effects logistic     | 1.4                                          | 2.0 | 2.7 | 3.2 | 73                                      |
| RE regression               | 1.5                                          | 2.2 | 3.0 | 3.7 | 79                                      |
| Two-level RE logistic       | 1.2                                          | 1.8 | 2.3 | 2.7 | 70                                      |
| Random-coefficients Poisson | 0.9                                          | 0.9 | 0.9 | 0.9 | 0                                       |
| RE logistic with constant   | 1.6                                          | 2.3 | 3.2 | 4.1 | 82                                      |

All values are expressed as the speed relative to the speed of a single core.

a. Bigger is better; 2 is perfect for 2 cores, 4 is perfect for 4 cores, 8 is perfect for 8 cores, and 16 is perfect for 16 cores.

b. Bigger is better; 100 is perfect.

Table 5. Command descriptions

| Command                     | Description                                                                                            |
|-----------------------------|--------------------------------------------------------------------------------------------------------|
| Finite mixture model        | Gaussian finite mixture model with two point masses                                                    |
| Item response model         | Two-parameter logistic item response model                                                             |
| Latent class model          | Gaussian latent class model with two levels in the latent class                                        |
| Measurement error model     | Logistic regression with measurement error in a covariate                                              |
| Rank-outcome latent class   | Latent class model for rank outcomes                                                                   |
| MIMIC model                 | Multiple-indicator, multiple-cause (MIMIC) latent variables structural equation model—ordered logistic |
| Random-effects logistic     | Random-effects (random-intercepts) logistic regression—same as <code>xtlogit, re</code>                |
| RE regression               | Continuous (Gaussian distribution) model with random intercepts—same as <code>xtreg, re</code>         |
| Two-level RE logistic       | Logistic regression with two levels of random intercepts                                               |
| Random-coefficients Poisson | Poisson count-data model with random intercepts and a random coefficient                               |
| RE logistic with constant   | Random-effects (random-intercepts) logistic regression, fewer observations                             |

The graphs below show the observed performances from table 4 in graphical form. Those graphs are followed by graphs showing performance through 40 cores.

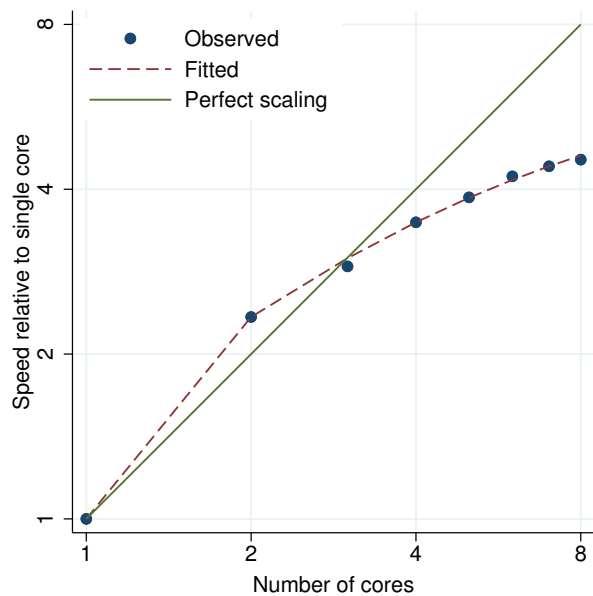


Figure 676. Finite mixture model performance plot.

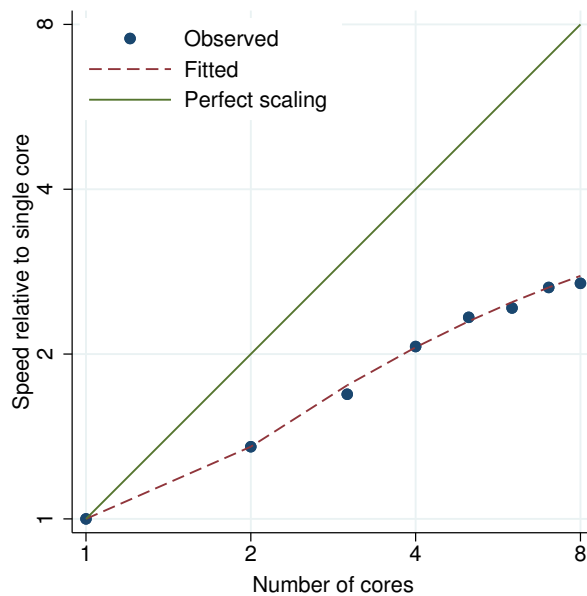


Figure 677. Item response model performance plot.

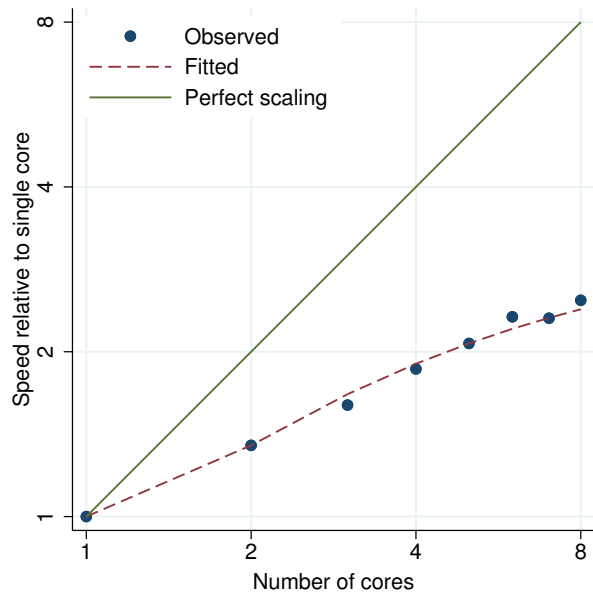


Figure 678. Latent class model performance plot.

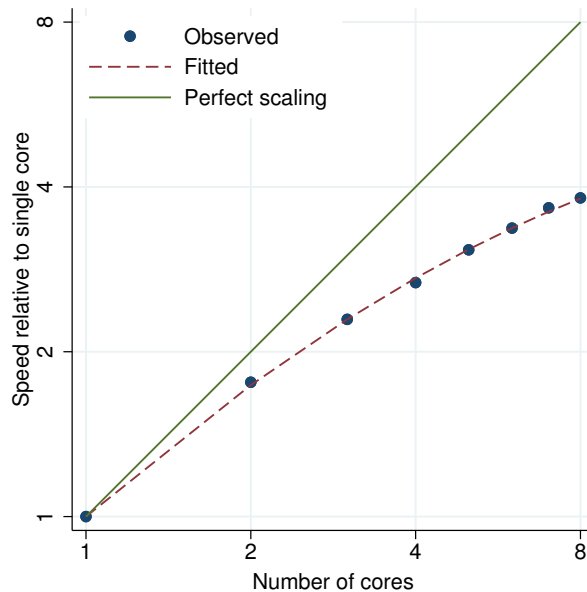


Figure 679. Measurement error model performance plot.

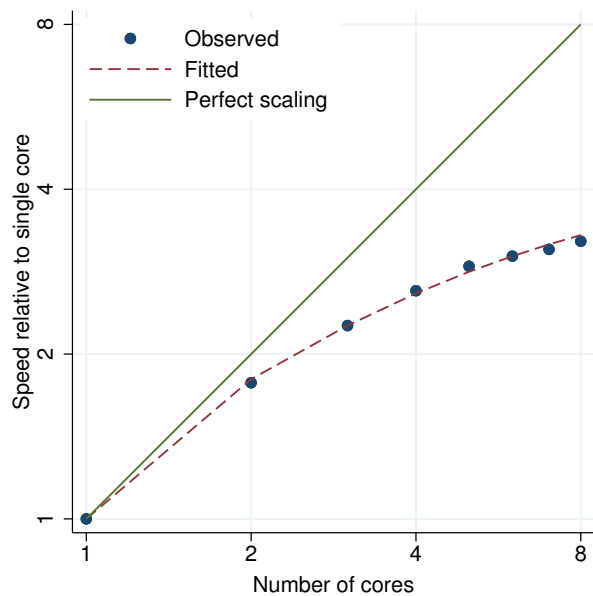


Figure 680. Rank-outcome latent class performance plot.

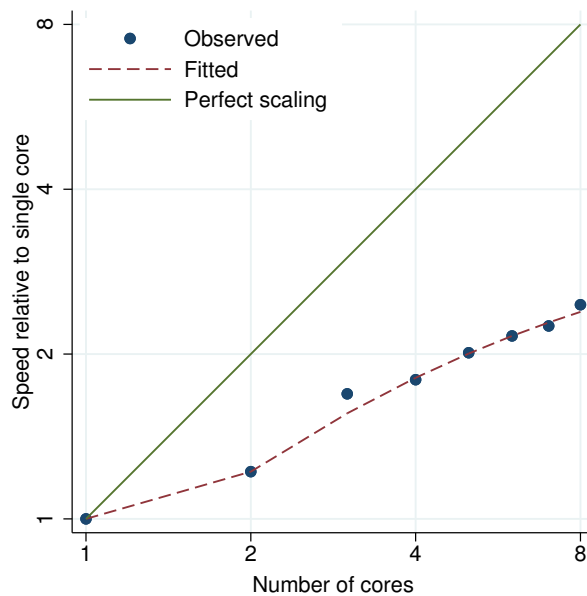


Figure 681. MIMIC model performance plot.

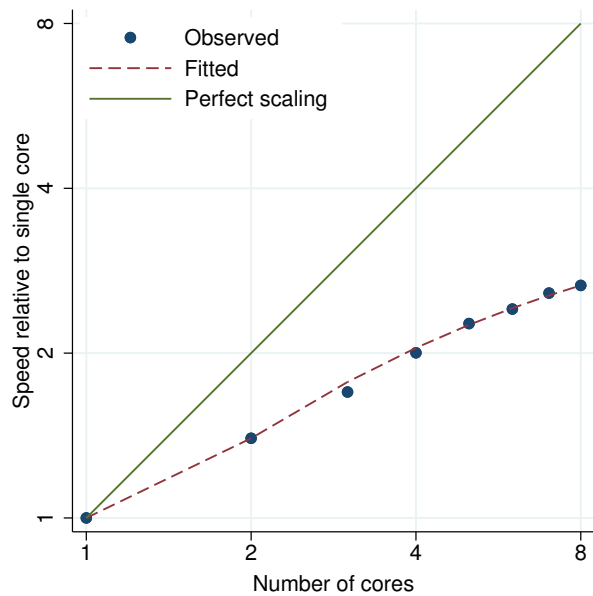


Figure 682. Random-effects logistic performance plot.

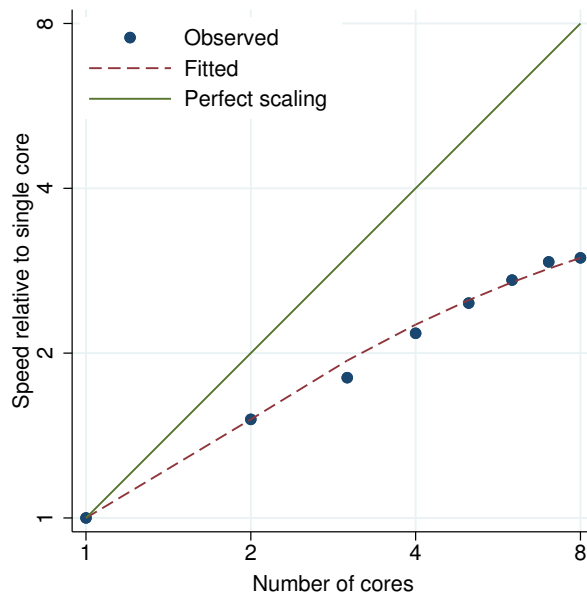


Figure 683. RE regression performance plot.

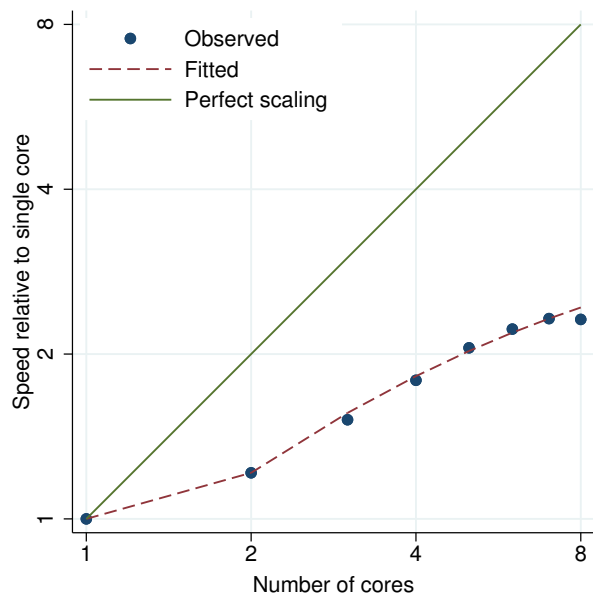


Figure 684. Two-level RE logistic performance plot.

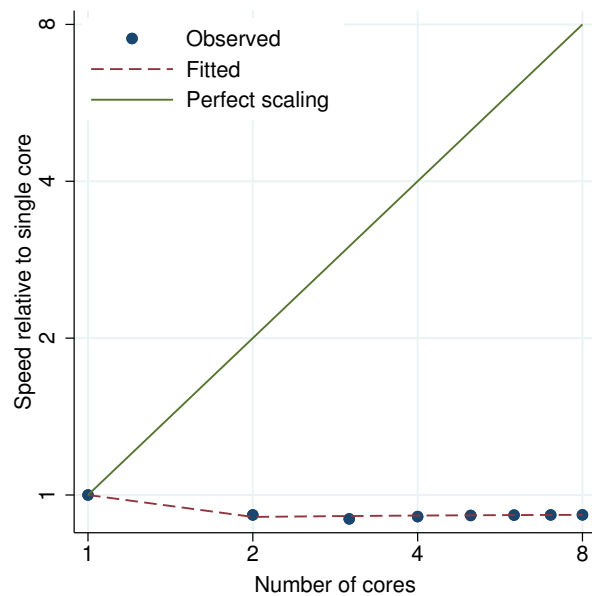


Figure 685. Random-coefficients Poisson performance plot.

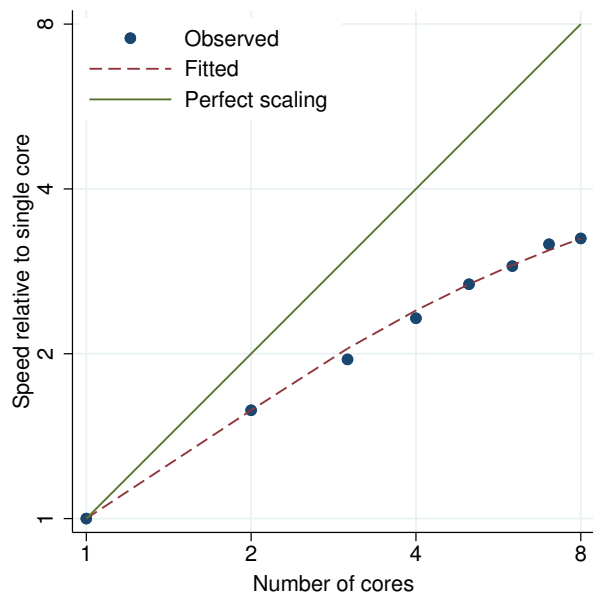


Figure 686. RE logistic with constant performance plot.

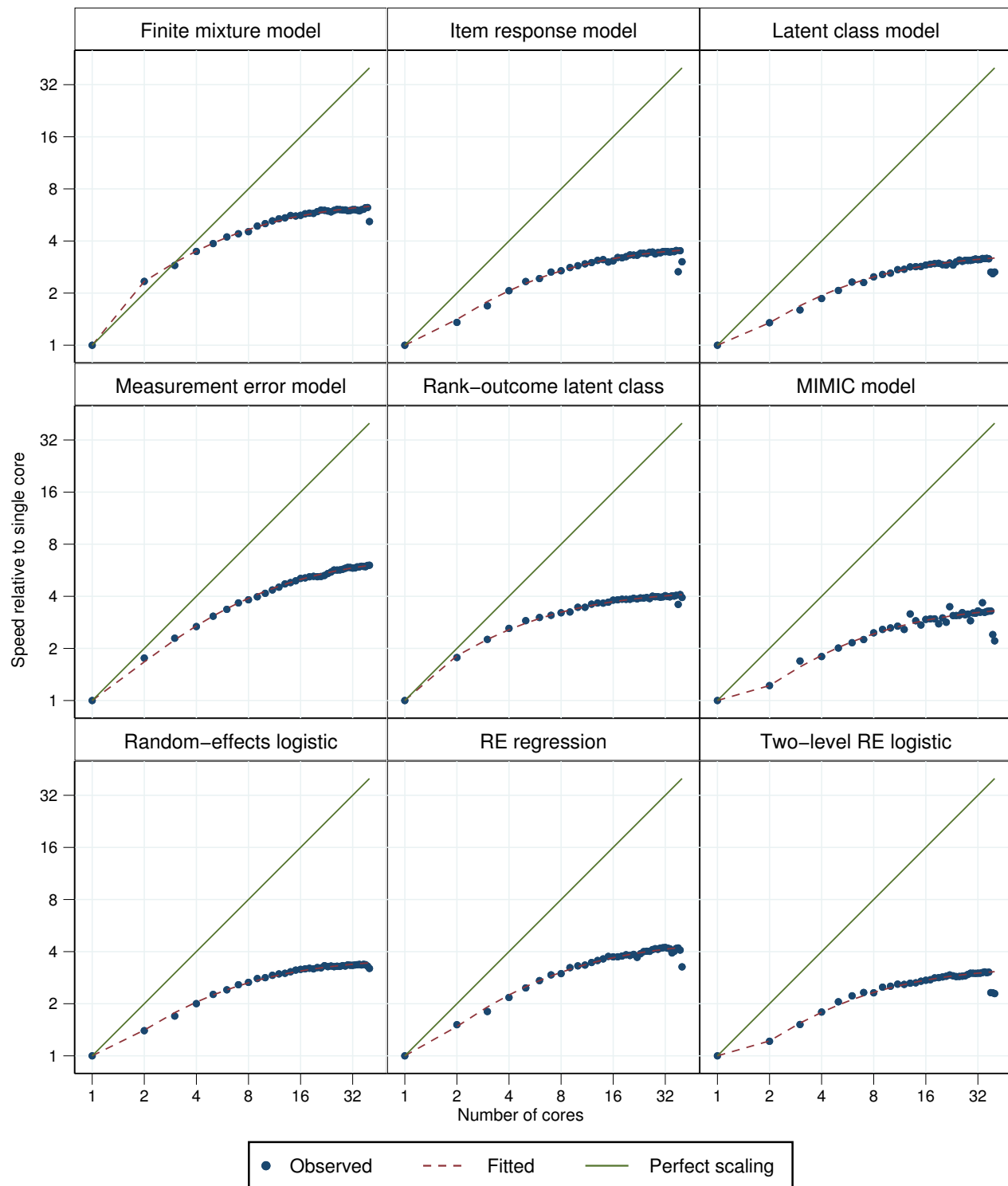


Figure 687. Parallelization performance plots.



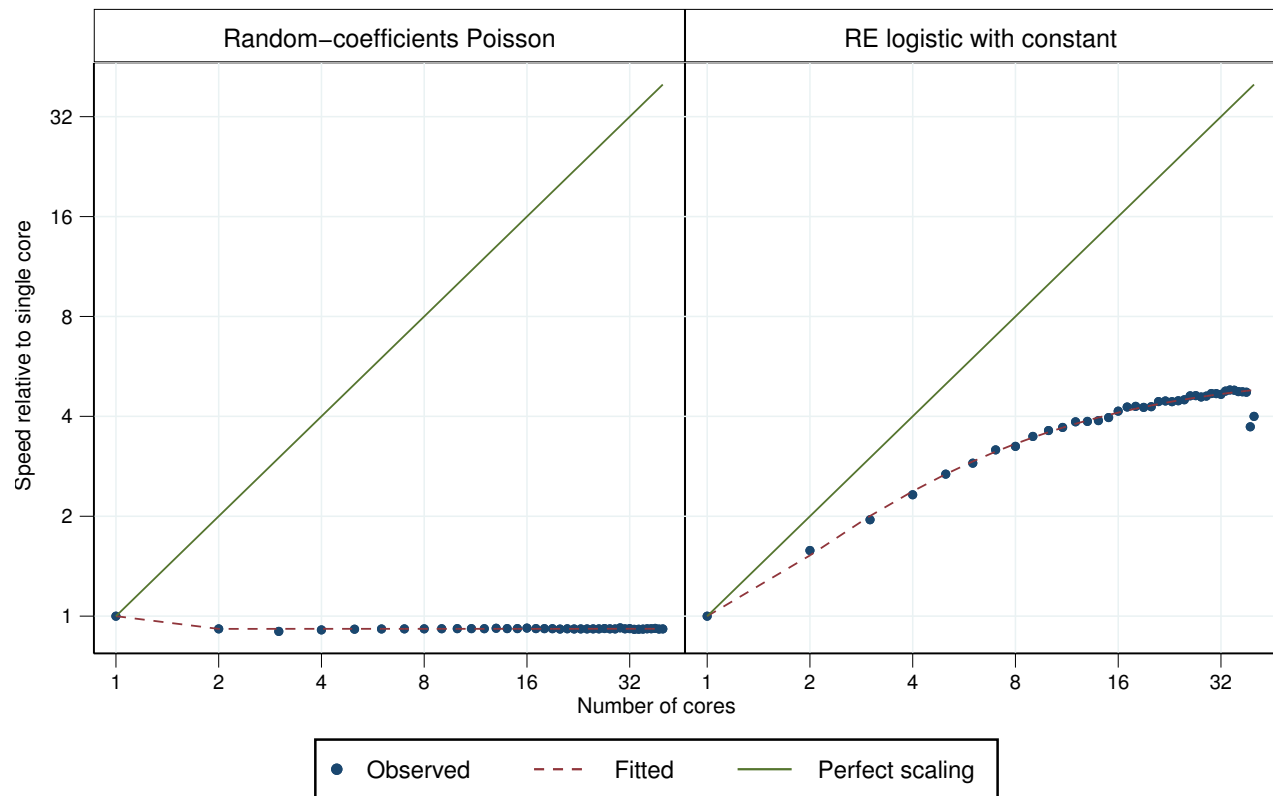


Figure 688. **Parallelization performance plots.**

## References

- Culler, D. E., J. P. Singh, and A. Gupta. 1999. *Parallel Computer Architecture: A Hardware/Software Approach*. San Francisco: Morgan Kaufmann.
- Dagum, L., and R. Menon. 1998. OpenMP: An industry-standard API for shared memory programming. *IEEE Computational Science & Engineering* 5: 46–55.
- Grama, A., A. Gupta, G. Karypis, and V. Kumar. 2003. *Introduction to Parallel Computing*. 2nd ed. Boston: Addison–Wesley.
- Moore, G. E. 1965. Cramming more components onto integrated circuits. *Electronics* 38: 114–117.