Bootstrap inference for fixed-effect models

Additional illustrations

Ayden Higgins*
University of Oxford

Koen Jochmans†
Toulouse School of Economics

September 18, 2023

Abstract

This document contains calculations and simulations for three examples (the many normal means problem, a dynamic logit model, and inference on the second moment of the fixed effects in the many normal means setting). It also provides full estimation results for the empirical application.

1 Examples

Many normal means In the classic problem of Neyman and Scott (1948) we observe independent variables

$$z_{it} \sim N(\eta_{i0}, \varphi_0).$$

Maximum likelihood estimates the mean parameters by the within-strata sample averages $\overline{z}_i := {}^1\!/\!_m \sum_{t=1}^m z_{it}$ and the common variance parameter by

$$\hat{\varphi} = \frac{1}{nm} \sum_{i=1}^n \sum_{t=1}^m (z_{it} - \overline{z}_i)^2.$$

It is well known that, in this case,

$$\sqrt{nm}(\hat{\varphi} - \varphi_0) \xrightarrow{L} N(-\gamma\varphi_0, 2\varphi_0^2), \tag{E.1}$$

*Address: Department of Economics, University of Oxford, 10 Manor Road, Oxford OX1 3UQ, United Kingdom. E-mail: ayden.higgins@economics.ox.ac.uk.

†Address: Toulouse School of Economics, 1 esplanade de l'Université, 31080 Toulouse, France. E-mail: koen.jochmans@tse-fr.eu.

under rectangular-array asymptotics. Here, starting from the fact that $nm\,\hat{\varphi}/\varphi_0 \sim \chi^2_{n(m-1)}$, the exact distribution of the maximum-likelihood estimator can be derived. We find that

$$\sqrt{nm}(\hat{\varphi} - \varphi_0) \sim \text{Gamma}\left(-\sqrt{nm}\varphi_0, \frac{n(m-1)}{2}, \frac{2\varphi_0}{\sqrt{nm}}\right),$$

where $\text{Gamma}(\vartheta_1, \vartheta_2, \vartheta_3)$ refers to the Gamma distribution with shape ϑ_2 and scale ϑ_3, shifted by ϑ_1. It is readily verified that the mean and variance of this distribution are equal to

$$-\sqrt{\frac{n}{m}}\varphi_0, \qquad 2\varphi_0^2\left(1 - \frac{1}{m}\right),$$

respectively.

In this example, the bootstrap independently samples $z^*_{it} \sim N(\overline{z}_i, \hat{\varphi})$. The associated maximum-likelihood estimators are $\overline{z}^*_i$ and

$$\hat{\varphi}^* = \frac{1}{nm}\sum_{i=1}^{n}\sum_{t=1}^{m}(z^*_{it} - \overline{z}^*_i)^2.$$

Conditional on the data, the latter estimator follows the same Gamma distribution as above, only with φ_0 replaced by $\hat{\varphi}$. Noting that we can write $\sqrt{nm}(\hat{\varphi} - \varphi_0) = -\sqrt{n/m}\,\varphi_0 + \epsilon$, for a mean-zero random variable $\epsilon = O_P(1)$, this implies that

$$\sqrt{nm}(\hat{\varphi}^* - \hat{\varphi}) \sim \text{Gamma}\left(-\left(\sqrt{nm}\varphi_0 - \sqrt{\frac{n}{m}}\varphi_0 + \epsilon\right), \frac{n(m-1)}{2}, \frac{2\varphi_0}{\sqrt{nm}}\left(1 - \frac{1}{m}\right) + \frac{2\epsilon}{nm}\right)$$

conditional on the sample. The mean and variance of this distribution are

$$-\sqrt{\frac{n}{m}}\varphi_0 + \frac{1}{m}\left(\sqrt{\frac{n}{m}}\varphi_0 - \epsilon\right), \qquad 2\varphi_0^2 + \frac{2}{m}\left(\epsilon^2 + \sqrt{\frac{m}{n}}\varphi_0\epsilon - 3\varphi_0^2\right) + O\left(\frac{1}{m^2}\right)$$

which, to first order, agree with the corresponding moments of the maximum-likelihood estimator.

The studentized maximum-likelihood estimator follows a (translated) inverse-Gamma distribution, mirrored about the origin. Moreover,

$$-\sqrt{nm}\,\frac{(\hat{\varphi} - \varphi_0)}{\sqrt{2\hat{\varphi}^2}} \sim \text{Inverse-Gamma}\left(-\sqrt{\frac{nm}{2}}, \frac{n(m-1)}{2}, \sqrt{\frac{nm}{2}}\,\frac{nm}{2}\right),$$

where $\text{Inverse-Gamma}(\vartheta_1, \vartheta_2, \vartheta_3)$ refers to the Inverse-gamma distribution with shape ϑ_2 and scale ϑ_3, shifted by ϑ_1. This distribution is pivotal and the bootstrap replicates it

Table E.1: Many normal means: bias, standard deviation, coverage, and size for φ_0

		BIAS	STD	COVERAGE					SIZE	
n	m	MLE		MLE	BC	BB	DBB	SB	LR	LR*
10	10	-0.100	0.134	0.827	0.904	0.918	0.950	0.950	0.128	0.050
20	10	-0.100	0.095	0.763	0.903	0.922	0.950	0.950	0.193	0.050
40	10	-0.100	0.067	0.637	0.902	0.926	0.950	0.950	0.323	0.050
100	10	-0.100	0.042	0.330	0.897	0.927	0.950	0.950	0.642	0.050

exactly. Thus, at least in this example, the studentized bootstrap yields confidence intervals whose probability of covering φ_0 can be controlled exactly.

A first-order correction to $\hat{\varphi}$ based on a plug-in estimator of its asymptotic bias is

$$\tilde{\varphi} \coloneqq \hat{\varphi} + \frac{\hat{\varphi}}{m}.$$

It is interesting to compare the performance of confidence intervals for φ_0 based on bias correction with those obtained via the bootstrap. The bias-correction approach uses the large-sample approximation

$$\sqrt{nm}\,\frac{(\tilde{\varphi}-\varphi_0)}{\sqrt{2\hat{\varphi}^2}} \overset{L}{\to} N(0,1).$$

Its coverage accuracy can be evaluated for any given sample size from the observation that

$$-\sqrt{nm}\,\frac{(\tilde{\varphi}-\varphi_0)}{\sqrt{2\hat{\varphi}^2}} \sim \text{Inverse-Gamma}\left(-\sqrt{\frac{nm}{2}}\left(1+\frac{1}{m}\right), \frac{n(m-1)}{2}, \sqrt{\frac{nm}{2}}\,\frac{nm}{2}\right).$$

Observe that this distribution coincides with that of the studentized maximum-likelihood estimator up to the location parameter, the current distribution being located closer to zero. We further remark that bootstrapping the bias-corrected estimator would yield exactly the same confidence interval as the one obtained by bootstrapping the uncorrected estimator. In that sense, there is no gain to be had from implementing any bias correction in this example.

Table E.1 contains the bias and standard deviation of the maximum-likelihood estimator for $\varphi_0 = 1$ and gives coverage rates of two-sided 95% confidence intervals for φ_0. The rates are invariant to the value of φ_0. The bias is not small relative to the standard deviation. Consequently, confidence intervals constructed by means of the naive normal approximation to maximum likelihood (MLE) perform poorly but bootstrapping the maximum-likelihood estimator, both using the basic bootstrap (BB) and the studentized bootstrap (SB), yields reliable inference. Here, the latter gives exact coverage. Iterating the basic bootstrap (DBB) also yields exact coverage. Confidence intervals based on bias correcting (BC) the maximum-likelihood estimator improve considerably on MLE but still undercover by about 5 percentage points in all designs considered.

Next, consider testing the null hypothesis that the variance parameter is equal to φ_0. The likelihood-ratio statistic is

$$nm\,\hat{\varphi}/\varphi_0 - nm\,\log(nm\,\hat{\varphi}/\varphi_0) - nm + nm\log(nm)$$

and depends on the data only through $nm\,\hat{\varphi}/\varphi_0$. The latter has a pivotal distribution and, hence, so does the test statistic. A small calculation reveals that its limit distribution is a non-central χ^2_1-distribution with non-centrality parameter $\gamma^2/2$. Consequently, while a decision rule based on critical values from the χ^2_1-distribution will not yield size control, using the quantiles of the bootstrap distribution will. Furthermore, in this example size is controlled exact in finite samples. The size distortion of the likelihood-ratio (LR) test when using the .95 quantile of the χ^2_1-distribution and the improvement when working instead with bootstrap critical values (LR*) is illustrated in Table E.1.

Dynamic logit For our next example we consider the Markov process

$$y_{it} = \begin{cases} 1 & \text{if } \eta_{i0} + \varphi_0 y_{it-1} > \varepsilon_{it} \\ 0 & \text{if not} \end{cases},$$

where the ε_{it} are independent and identically distributed logistic random variables, i.e., $\mathbb{P}(\varepsilon_{it} \leq a) = (1+e^{-a})^{-1} =: F(a)$. The initial conditions, y_{i0}, are observed and held fixed throughout.

The maximum-likelihood estimator is not available in closed form. Nonetheless, the log-likelihood function is globally concave and numerical optimization is straightforward, exploiting the sparsity of the Hessian matrix (see, e.g., Chamberlain 1980). Further, an excellent starting value for the bootstrap maximum-likelihood estimator comes in the form of the maximum-likelihood estimator based on the original data, as the latter is used to generate the bootstrap samples. Given $\hat{\varphi}$ and $\hat{\eta}_1, \ldots, \hat{\eta}_n$ we generate bootstrap samples for the dynamic logit model by recursively drawing y_{it}^* from a Bernoulli distribution with success probability $F(\hat{\eta}_i + \hat{\varphi} y_{it-1}^*)$. Each bootstrap iteration starts at the initial condition y_{i0}.

The exact distribution of $\hat{\varphi}$ is not known so we resort to simulations. We draw y_{i0} with

$$\mathbb{P}(y_{i0} = 1) = \frac{F(\eta_{i0})}{1 - F(\eta_{i0} + \varphi_0) + F(\eta_{i0})},$$

set $\eta_{i0} = 0$ for all the strata, and consider autoregressive parameters $\varphi_0 \in \{1/2, 1, 3/2\}$. Table E.2 provides the bias and standard deviation of the maximum-likelihood estimator, the coverage rates and average length of various (two-sided) 95% confidence intervals for φ_0, and the size of the likelihood-ratio test with a (theoretical) size of 5% for different choices of critical value.

We report coverage and length for confidence intervals based on (the naive normal approximation to) the maximum-likelihood estimator (MLE), the basic bootstrap (BB) and studentized bootstrap (SB) and their iterated version (DBB and DSB, respectively), as well as on two procedures that adjust the maximum-likelihood estimator for its bias. The first of these adjustments (BC1) is the analytical correction of Hahn and Kuersteiner (2011). The second adjustment (BC2) is due to Fernández-Val (2009) and exploits the model structure to implement a refined correction that replaces certain sample averages by expected quantities. Both these approaches require a bandwidth choice. We report results for a bandwidth equal to one, which we found was the choice that performed best. For the likelihood-ratio test we report size for the decision rule based on the .95 quantile of the χ^2_1-distribution (LR), the .95 quantile of the bootstrap distribution (LR*) and quantiles set according to the double bootstrap (LR**). All (single) bootstrap results are based on the use of 999 bootstrap replications. For the double bootstrap, we use 999 replications in the outer iteration and 316 replications in the inner iteration (following Booth and Hall 1994).

Table E.2: Dynamic logit: bias, standard deviation, coverage and length of confidence interval, and size of likelihood-ratio test for φ_0

			BIAS	STD	COVERAGE							LENGTH							SIZE		
φ_0	n	m	MLE		MLE	BC1	BC2	BB	DBB	SB	DSB	MLE	BC1	BC2	BB	DBB	SB	DSB	LR	LR*	LR**
1/2	100	10	-0.461	0.145	0.111	0.942	0.970	0.964	0.958	0.928	0.940	0.567	0.573	0.575	0.630	0.616	0.543	0.571	0.887	0.064	0.041
1/2	100	20	-0.219	0.095	0.378	0.952	0.962	0.958	0.955	0.943	0.949	0.378	0.380	0.381	0.396	0.394	0.373	0.381	0.640	0.057	0.048
1/2	250	10	-0.459	0.090	0.001	0.895	0.968	0.957	0.949	0.928	0.932	0.358	0.362	0.363	0.397	0.389	0.343	0.364	0.999	0.092	0.051
1/2	250	20	-0.217	0.062	0.054	0.937	0.952	0.958	0.961	0.942	0.945	0.239	0.241	0.241	0.250	0.250	0.236	0.240	0.957	0.062	0.049
1	100	10	-0.514	0.150	0.086	0.880	0.941	0.964	0.949	0.916	0.937	0.605	0.620	0.623	0.657	0.629	0.577	0.614	0.919	0.113	0.065
1	100	20	-0.244	0.103	0.332	0.907	0.921	0.948	0.948	0.931	0.940	0.404	0.410	0.410	0.418	0.416	0.398	0.408	0.627	0.055	0.040
1	250	10	-0.513	0.095	0.000	0.745	0.898	0.970	0.952	0.907	0.952	0.383	0.392	0.394	0.414	0.393	0.366	0.396	0.999	0.144	0.053
1	250	20	-0.244	0.065	0.039	0.881	0.922	0.959	0.950	0.954	0.944	0.256	0.259	0.259	0.264	0.263	0.251	0.258	0.957	0.069	0.053
3/2	100	10	-0.623	0.165	0.034	0.695	0.837	0.942	0.930	0.867	0.926	0.678	0.700	0.706	0.722	0.675	0.642	0.693	0.966	0.178	0.091
3/2	100	20	-0.299	0.113	0.269	0.835	0.883	0.940	0.930	0.932	0.932	0.454	0.462	0.463	0.463	0.460	0.443	0.458	0.712	0.068	0.047
3/2	250	10	-0.624	0.104	0.000	0.304	0.593	0.917	0.913	0.776	0.936	0.428	0.442	0.446	0.453	0.419	0.405	0.451	1.000	0.274	0.109
3/2	250	20	-0.302	0.071	0.010	0.636	0.741	0.953	0.949	0.932	0.948	0.286	0.292	0.292	0.292	0.290	0.280	0.291	0.981	0.101	0.057

The naive normal approximation to the sampling distribution of the maximum-likelihood estimator again yields unreliable inference in this problem. Bias correction yields a large improvement in coverage rates and comes with only minor increases in the length of the confidence intervals (which is informative about efficiency). Confidence intervals based on the correction underlying BC2 tend to give better coverage than those based on BC1, with the difference sometimes being considerable (up to 30 percentage points in the table). This highlights the sensitivity of bias-corrected inference to how the bias is being estimated. The performance of both BC1 and BC2 also deteriorates substantially as the value of φ_0 increases, highlighting the sensitivity of bias estimators to relatively minor design changes. Both these issues are not accounted for by first-order theory. Confidence interval based on the bootstrap, both in its basic and in its studentized form, are competitive with those based on bias correction and their performance is stable across different values of φ_0. BB does at least as well as BC2 in terms of coverage, and its iterated version DBB gives very similar coverage. SB and SDB yield somewhat shorter confidence intervals and, especially in the shortest panels, iterating gives improved coverage. For the likelihood-ratio test we observe a similar pattern as for the studentized bootstrap. LR shows large over-rejection rates while LR* and, even more so, LR** yield tests with size close to nominal size.

Many normal means (cont'd) In our third example we reconsider the setup of Neyman and Scott (1948) but change the parameter of interest to

$$\mathit{\Delta} = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \eta_{i0}^2,$$

the second moment of the fixed effects. The plug-in estimator is ${}^1\!/_n \sum_{i=1}^n \overline{z}_i^2$. Using the fact that $\overline{z}_i \sim N(\eta_{i0}, {}^{\varphi_0}\!/_m)$ by normality of the data it is easy to verify that the plug-in bias due to the estimation of the fixed effects is ${}^{\varphi_0}\!/_m$, while the estimator's sampling variance equals

$$\frac{2\varphi_0}{nm}\left(2\frac{\sum_{i=1}^n \eta_{i0}^2}{n} + \frac{\varphi_0}{m}\right).$$

The second component in the variance expression is of smaller order and asymptotically negligible.

Table E.3: Many normal means: bias, standard deviation and coverage and length of confidence interval for $\lim_{n\to\infty} {}^1\!/n \sum_{i=1}^n \eta_{i0}^2$

		BIAS	STD	COVERAGE						LENGTH					
n	m	MLE		MLE	BC	BB	DBB	SB	DSB	MLE	BC	BB	DBB	SB	DSB
50	10	0.100	0.055	0.550	0.953	0.954	0.952	0.917	0.945	0.232	0.232	0.232	0.231	0.210	0.234
50	20	0.050	0.038	0.702	0.958	0.942	0.934	0.915	0.929	0.156	0.156	0.156	0.153	0.145	0.153
50	50	0.020	0.023	0.782	0.936	0.939	0.932	0.916	0.929	0.095	0.095	0.095	0.094	0.091	0.094
100	10	0.100	0.039	0.256	0.951	0.956	0.962	0.922	0.965	0.163	0.163	0.163	0.165	0.147	0.178
100	20	0.050	0.027	0.517	0.955	0.947	0.942	0.918	0.937	0.110	0.110	0.110	0.109	0.101	0.110
100	50	0.020	0.017	0.702	0.941	0.949	0.945	0.935	0.943	0.067	0.067	0.067	0.066	0.064	0.066

The bootstrap again independently samples $z_{it}^* \sim N(\bar{z}_i, \hat{\varphi})$ and subsequently constructs the estimator ${}^1\!/n \sum_{i=1}^n \overline{z}_i^{*2}$. The exact distribution of the estimator is a complicated mixture and so we once more resort to simulations to evaluate the performance of the bootstrap. In our simulations we set $\eta_{i0} = i/n$ so that, in large samples, the distribution of the fixed effects is uniform on $[0, 1]$; hence, $\Delta = {}^1\!/3$. Data were generated with $\varphi_0 = 1$.

We report results for several choices of (n, m) in Table E.3. The bootstrap confidence intervals are again found to yield a large improvement in coverage rates relative to the ones based on the naive plug-in approach and are competitive with those based on bias correction. Again the basic bootstrap does better than the studentized version and has actual coverage very close to theoretical coverage for all designs. Iterating the former does little in terms of coverage rates. Iterating the latter gives further improvement, especially in the shorter panels. The average length of the confidence intervals is very similar across the different methods.

2 Empirical illustration

For our empirical example we use data from the Panel Study of Income Dynamics (PSID) to look at determinants of labor-force participation decisions of married woman. We follow Hyslop (1999) and specify the participation decision as a dynamic probit model with unitspecific intercepts. We included the number of children of at most two years of age (# children 0–2), between 3 and 5 years of age (# children 3–5), and between 6 and 17 years of age (# children 6–17), as well as the log of the husband's earnings (log husband income; expressed in thousands of 1995 U.S. dollars), and a quadratic function of age. Carro (2007), Fernández-Val (2009), and Dhaene and Jochmans (2015) have previously estimated the same specification using various bias-corrected estimators. To ensure comparability with their results we use the same data (Fernández-Val, 2022), which concern the period 1979–1988. The sample consists of 1461 women aged between 18 and 60 in 1985 who, throughout the sampling period, were married to men who were in the active labor force the whole time. During the sampling period, 664 of these individuals changed participation status at least once. The time series for the others can be fitted perfectly by setting their fixed effect to either $-\infty$ (if they never worked) or $+\infty$ (if they always worked). These observations do not carry any information about the common parameters and do not contribute to the (concentrated) likelihood function.

Table E.4 contains points estimates, standard errors, and 95% confidence intervals for the coefficients of the probit model. As before, we provide results for maximum likelihood, the bias-corrected estimators of Hahn and Kuersteiner (2011) and Fernández-Val (2009), the basic bootstrap and the studentized bootstrap, and the iterated version of the latter two. For maximum likelihood and for the two bias-corrected estimators the standard errors and confidence intervals are based on the conventional normal approximation, using the Hessian matrix evaluated at the point estimates to estimate the Fisher information. For the bootstrap we provide a single point estimate, obtained by subtracting the median of the bootstrap distribution of $\hat{\varphi}^* - \hat{\varphi}$ from $\hat{\varphi}$, and a single standard error, calculated as the standard deviation of the bootstrap distribution of $\hat{\varphi}^* - \hat{\varphi}$ (without winsorization). Bias correction using the mean rather than the median (not reported) yielded very similar point estimates.

The difference between standard maximum likelihood and the other approaches is most pronounced in the coefficients that capture state dependence and the impact of having young children. Adjusting the point estimates for bias leads to an upward revision in each

Table E.4: Female labor-force participation

	MLE	BC1	BC2	BB	SB
Lagged participation	0.756	0.992	1.031	1.163	
	(0.043)	(0.043)	(0.043)	(0.045)	
	[0.672 0.840]	[0.908 1.075]	[0.948 1.115]	[1.077 1.250]	[1.053 1.210]
				[1.105 1.262]	[1.105 1.279]
# Children 0-2	-0.554	-0.477	-0.436	-0.365	
	(0.057)	(0.057)	(0.057)	(0.074)	
	[-0.667 -0.442]	[-0.590 -0.365]	[-0.548 -0.324]	[-0.507 -0.219]	[-0.511 -0.263]
				[-0.487 -0.219]	[-0.489 -0.278]
# Children 3-5	-0.279	-0.213	-0.193	-0.140	
	(0.053)	(0.053)	(0.053)	(0.068)	
	[-0.384 -0.175]	[-0.317 -0.109]	[-0.297 -0.089]	[-0.274 -0.006]	[-0.274 -0.035]
				[-0.247 -0.006]	[-0.249 -0.051]
# Children 6-17	-0.075	-0.056	-0.050	-0.036	
	(0.042)	(0.042)	(0.042)	(0.055)	
	[-0.158 0.008]	[-0.140 0.027]	[-0.134 0.033]	[-0.144 0.074]	[-0.138 0.059]
				[-0.113 0.039]	[-0.115 0.042]
Log husband income	-0.246	-0.232	-0.209	-0.185	
	(0.055)	(0.055)	(0.055)	(0.069)	
	[-0.354 -0.139]	[-0.339 -0.124]	[-0.317 -0.101]	[-0.317 -0.049]	[-0.312 -0.075]
				[-0.292 -0.095]	[-0.293 -0.094]
Age	2.050	1.844	1.616	1.407	
	(0.387)	(0.387)	(0.387)	(0.500)	
	[1.292 2.809]	[1.086 2.602]	[0.858 2.374]	[0.424 2.370]	[0.587 2.346]
				[0.805 2.147]	[0.768 2.159]
Age squared	-0.250	-0.224	-0.196	-0.169	
	(0.051)	(0.051)	(0.051)	(0.066)	
	[-0.351 -0.149]	[-0.325 -0.123]	[-0.297 -0.095]	[-0.296 -0.038]	[-0.293 -0.060]
				[-0.268 -0.038]	[-0.269 -0.083]

of these coefficients for all methods. This revision is especially large (relative to its standard error) for the coefficient on lagged participation. Its maximum-likelihood estimate is well outside any of the theoretically-justified confidence intervals. The bootstrap bias correction is somewhat larger than that of the two analytical corrections

The bootstrap-based confidence intervals are located somewhat further away from the maximum-likelihood point estimates, as are those obtained by centering a conventional confidence interval around a bias-corrected estimator. Confidence intervals based on the studentized bootstrap are somewhat shorter and more asymmetric than those based on the basic bootstrap. Iterating the bootstrap leads to some revision of the confidence intervals, especially for the studentized version of the bootstrap. The double-bootstrap confidence intervals based on the basic version of the bootstrap and the studentized bootstrap are very similar, and typically closer to those of the basic bootstrap than to those of the studentized bootstrap that had been obtained prior to iterating. All these observations are in line with what has been observed in the simulation results for the dynamic logit model reported on above.

References

Booth, J. G. and P. Hall (1994). Monte Carlo approximation and the iterated bootstrap. *Biometrika 81*, 331–340.

Carro, J. M. (2007). Estimating dynamic panel data discrete choice models with fixed effects. *Journal of Econometrics 140*, 503–528.

Chamberlain, G. (1980). Analysis of covariance with qualitative data. *Review of Economic Studies 47*, 225–238.

Dhaene, G. and K. Jochmans (2015). Split-panel jackknife estimation of fixed-effect models. *Review of Economic Studies 82*, 991–1030.

Fernández-Val, I. (2009). Fixed effects estimation of structural parameters and marginal effects in panel probit models. *Journal of Econometrics 150*, 71–85.

Fernández-Val, I. (2022). Datasets for Fernández-Val (2009), Journal of Econometrics

150(1), pp. 71-85. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2022-06-10. https://doi.org/10.3886/E172521V1.

Hahn, J. and G. Kuersteiner (2011). Bias reduction for dynamic nonlinear panel models with fixed effects. *Econometric Theory 27*, 1152–1191.

Hyslop, D. R. (1999). State dependence, serial correlation and heterogeneity in intertemporal labor force participation of married women. *Econometrica 67*, 1255–1294.

Neyman, J. and E. L. Scott (1948). Consistent estimates based on partially consistent observations. *Econometrica 16*, 1–32.

PSID. Panel Study of Income Dynamics, public use dataset. Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI.