Appendix to "The Nature of Firm Growth"

Benjamin W. Pugsley   Petr Sedláček   Vincent Sterk

For Online Publication

Contents

A Data appendix

Our starting point is the U.S. Census Bureau Longitudinal Business Database first created by Jarmin and Miranda (2002) and maintained by the Census Bureau Center for Economic Studies (CES). This establishment-level administrative database provides near comprehensive coverage of all businesses with paid employees. Each establishment, i.e., physical location of economic activity, has a unique identifier that may be used to link them across any time span. Each establishment is also assigned a firm identifier, which may be used, in the case of multiple locations, to group establishments within a year at the firm level.

For this paper we use data from 1976 to 2012 for the entire nonfarm private sector. Both our firm and establishment level results rely on a measure of age. For establishments, we measure entry (i.e. assign age 0) in the year in which it hires its first employee, after which it ages naturally. This measure of age is tied to the physical location and is not affected by business sales or reorganizations. When a new firm is identified, it is assigned the age of the oldest establishment. Because new entrants are identified as not operating in any previous years, we use data from 1979 to 2012. This ensures that any establishment or firm in the 1979 birth cohort were not present in at least 3 prior years.

For each establishment, we assign a longitudinally consistent NAICS6 industry using the methodology and concordance developed by Fort, Klimek, et al. (2016). This allows a consistent measure of industry for establishments created prior to the replacement of the SIC system with NAICS. To assign an industry to each firm we apply an establishment-payroll weighted hierarchal system that first assigns the highest payroll 2-digit industry, then the 3-digit industry within the matching 2-digit group and so on until a 6 digit industry is assigned to each firm.

We use for each establishment and aggregated at the firm level as described above, the total paid employment (annually for the week including March 12), NAICS6 industry and age. For each birth cohort from 1979 to 1993, we link establishment and firms across each age pair for which it operates.

A.1 Linking firms and establishments

While the LBD is designed to link physical establishments over time, linking at the enterprise or firm level poses unique challenges. The Census Bureau staff links establishments, physical locations of business activity, across years and provides each establishment a unique identifier that is preserved across establishment reorganizations or ownership changes. Although taxes are paid at the level of a Employer Identification Number (EIN), the Census Bureau uses additional data from the quinquennial Economic Census and Annual Company Organization survey to both identify individual establishments and their locations and to identify in the cases of multiple locations and subsidiaries the highest level of operational control, known as an enterprise. The measure of enterprises corresponds to our notion of firms, which are defined by the span of managerial control over their inputs. The creation of the LBD and process of identifying establishments, linking them over time, and grouping at the enterprise level are all described by Jarmin and Miranda (2002) and references therein.

We use the unique physical establishment identifier to link establishments across any time span for which they are operating. The firm identifiers however, are only unique within an annual cross section. Identifiers are tied to the EIN, which may change in the case or reorganizations or when the business is officially recognized in an Economic Census year as having multiple locations. When used to link firms across time periods, the quality of the match degrades with the length of the time period, and the matched sample is no longer representative, since it only contains firms without an organizational change.

To construct firm-level linkages, we first create firm level longitudinal identifiers that are robust to organizational changes. For any establishment that undergoes a change in firm id or any establishment at a firm where an establishment undergoes a change in firm id, we assign new firm identifiers. The new measure works backwards and preserves the most recent firm id when the firms should be linked. The matching algorithm relies on administrative records and is detailed in Pugsley and Wheeler (2018). There are some certainty cases, such as when a "multiunit" firm id is assigned where the new firm id can be raked backwards along the previous firm id spell. In other cases, we rely on the first and last appearance of the firm id to determine when the firms should be linked and when the break should be preserved. In cases of mergers, we give preference to the firm with the largest historical payroll.

A.2 Estimated autocovariance matrices

We pool the 1979 to 1993 birth cohorts linked at the firm level across all firm ages for which the firm is operating. We refer to this as the unbalanced panel. We also construct another sample that restricted to include firms that survival at least 19 years. We refer to this as the balanced panel. We estimate the autocovariance structure separately for each panel.

Letting $\ln n_{i,a,j,t}$ represent the log employment for firm i of age a in 6-digit NAICS industry j and year t , we first project log employment on industry j and birth cohort $t-a$ fixed effects. Then, we construct log employment residuals $\widetilde{\ln n}_{i,a,j,t} = \ln n_{i,a,j,t} - \hat{\mu}_j - \hat{\lambda}_{t-a}$ by subtracting off the estimated fixed effects. This removes effects for all ages on employment due purely to the industry-specific technology and long-lasting effects on firms of all ages from business conditions at birth, as documented e.g., by Sedláček and Sterk (2017). Removing a joint industry×birth-cohort fixed effect instead has little effect on our results.

For each panel, we estimate the cross sectional autocovariance by taking age pair $h \geq 0$ and $a \geq h$ for $a \leq 19$:

$$\widehat{\text{Cov}}[\widetilde{\ln n}_{i,a,j,t}, \widetilde{\ln n}_{i,h,j,t}] = \frac{1}{N_{a,h}} \sum_{i=1}^{N_{a,h}} \left(\widetilde{\ln n}_{i,a,j,t} - \overline{\widetilde{\ln n}}_{i,a,j,t} \right) \left(\widetilde{\ln n}_{i,h,j,t} - \overline{\widetilde{\ln n}}_{i,h,j,t} \right) \quad \text{(A.1)}$$

Here $N_{a,h}$ refers to the number of firms with employment observed both at ages a and $h \leq a$, and $\bar{\cdot}$ taking the sample mean (of either age a or h residual log employment) over this group. In terms of composition, for the unbalanced panel, the composition and number $N_{a,h}$ of firms may change across age different (a, h) age group pairs. For example, the (1,0) pair will contain a larger set of firms than the (10,0) pair since more than half of the firms in the (1,0) pair will have exited before they reach age 10. For the balanced panel, the composition of age group pairs and size $N_{a,h}$ may still slightly fluctuate. They will not be constant because on occasion even surviving incumbent firms may be missing an employment report.

For reference, we report the estimated autocovariance matrices for firms for both the unbalanced and balanced panels in Tables A.1 and A.2

Table A.1: Autocovariance matrix of firms for unbalanced panel

	Age $a \geq 0$																			
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
0	0.964																			
1	0.771	1.022																		
2	0.712	0.863	1.047																	
3	0.668	0.805	0.904	1.065																
4	0.629	0.758	0.845	0.928	1.079															
5	0.606	0.730	0.811	0.883	0.960	1.099														
6	0.587	0.707	0.784	0.851	0.918	0.985	1.117													
7	0.570	0.686	0.760	0.824	0.886	0.943	1.005	1.133												
8	0.555	0.669	0.740	0.801	0.861	0.913	0.965	1.026	1.150											
9	0.542	0.654	0.723	0.781	0.839	0.888	0.935	0.986	1.045	1.165										
10	0.531	0.640	0.707	0.764	0.818	0.865	0.909	0.955	1.004	1.061	1.179									
11	0.517	0.623	0.689	0.745	0.799	0.843	0.885	0.929	0.973	1.020	1.076	1.192								
12	0.507	0.611	0.675	0.731	0.783	0.826	0.866	0.907	0.949	0.990	1.037	1.091	1.204							
13	0.498	0.600	0.661	0.716	0.768	0.809	0.848	0.888	0.927	0.966	1.008	1.052	1.105	1.216						
14	0.490	0.590	0.651	0.702	0.753	0.794	0.831	0.869	0.907	0.944	0.984	1.024	1.067	1.119	1.228					
15	0.483	0.581	0.640	0.691	0.734	0.775	0.811	0.848	0.884	0.920	0.958	0.995	1.034	1.077	1.128	1.235				
16	0.476	0.572	0.630	0.679	0.722	0.763	0.798	0.834	0.869	0.904	0.940	0.975	1.013	1.051	1.093	1.143	1.250			
17	0.471	0.564	0.621	0.669	0.712	0.751	0.787	0.823	0.857	0.890	0.925	0.959	0.994	1.031	1.069	1.110	1.160	1.264		
18	0.466	0.558	0.614	0.661	0.702	0.740	0.776	0.812	0.845	0.878	0.912	0.944	0.978	1.013	1.048	1.085	1.127	1.174	1.278	
19	0.462	0.554	0.608	0.655	0.695	0.733	0.768	0.803	0.837	0.869	0.902	0.933	0.966	0.999	1.033	1.068	1.105	1.144	1.192	1.294

Table A.2: Autocovariance matrix of firms for balanced panel

	Age $a \geq 0$																			
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
0	0.927																			
1	0.736	0.937																		
2	0.672	0.794	0.939																	
3	0.629	0.740	0.817	0.952																
4	0.597	0.702	0.772	0.843	0.966															
5	0.572	0.674	0.739	0.802	0.864	0.983														
6	0.552	0.651	0.713	0.772	0.825	0.886	0.997													
7	0.537	0.634	0.694	0.749	0.799	0.851	0.907	1.015												
8	0.525	0.620	0.679	0.732	0.779	0.827	0.875	0.931	1.035											
9	0.514	0.608	0.666	0.717	0.762	0.809	0.853	0.901	0.956	1.056										
10	0.503	0.598	0.654	0.704	0.748	0.792	0.834	0.878	0.925	0.977	1.076									
11	0.495	0.588	0.644	0.693	0.735	0.777	0.817	0.858	0.901	0.946	0.998	1.094								
12	0.486	0.579	0.634	0.683	0.724	0.765	0.803	0.842	0.882	0.923	0.968	1.017	1.111							
13	0.480	0.572	0.627	0.675	0.716	0.756	0.793	0.831	0.869	0.907	0.948	0.989	1.037	1.130						
14	0.473	0.566	0.620	0.668	0.708	0.748	0.785	0.822	0.859	0.895	0.934	0.971	1.012	1.061	1.152					
15	0.468	0.559	0.613	0.661	0.701	0.740	0.776	0.812	0.848	0.884	0.921	0.956	0.993	1.034	1.082	1.173				
16	0.463	0.553	0.607	0.654	0.693	0.732	0.768	0.803	0.839	0.874	0.909	0.943	0.978	1.016	1.056	1.104	1.197			
17	0.458	0.547	0.599	0.645	0.685	0.723	0.758	0.793	0.828	0.863	0.898	0.930	0.964	1.000	1.037	1.077	1.126	1.217		
18	0.453	0.540	0.592	0.638	0.676	0.713	0.748	0.783	0.817	0.851	0.885	0.917	0.950	0.985	1.019	1.057	1.099	1.146	1.239	
19	0.447	0.534	0.585	0.630	0.668	0.705	0.739	0.773	0.807	0.840	0.874	0.905	0.938	0.972	1.006	1.042	1.080	1.121	1.170	1.271

B Statistical model

B.1 Derivation of the autocovariance function

Consider the employment process given in Section 2.3 in the main text. It is helpful to write each of its components in its moving average representation:

$$\begin{aligned} u_{i,a} &= \rho_u^{a+1} u_{i,-1} + \sum_{k=0}^{a} \rho_u^k \theta_i \\ v_{i,a} &= \rho_v^{a+1} v_{i,-1} \\ w_{i,a} &= \sum_{k=0}^{a} \rho_w^k \varepsilon_{i,a-k} = \sum_{k=0}^{j-1} \rho^k \varepsilon_{i,a-k} + \rho_v^j \sum_{k=0}^{a-j} \rho_v^k \varepsilon_{i,a-j-k} \quad 0 \leq j \leq a. \end{aligned}$$

The last equality splits the moving average into terms before and after $a-j$ and will be helpful when computing the covariance. The level of log employment of firm i at age a can be written as:

$$\ln n_{i,a} = \rho_u^{a+1} u_{i,-1} + \sum_{k=0}^{a} \rho_u^k \theta_i + \rho_v^{a+1} v_{i,-1} + \sum_{k=0}^{j-1} \rho^k \varepsilon_{i,a-k} + \rho_v^j \sum_{k=0}^{a-j} \rho_v^k \varepsilon_{i,a-j-k} + z_{i,a}. \quad \text{(B.1)}$$

The autocovariance of log employment at age a and $h = a - j$ for $j \geq 0$ is:

$$\begin{aligned} \text{Cov}\left[\log n_{i,a}, \log n_{i,a-j}\right] = &\left(\sum_{k=0}^{a} \rho_u^k\right) \sigma_\theta^2 \left(\sum_{k=0}^{a-j} \rho_u^k\right) + \rho_u^{a+1} \sigma_{\tilde{u}}^2 \rho_u^{a-j+1} + \rho_v^{a+1} \sigma_{\tilde{v}}^2 \rho_v^{a-j+1} \\ &+ \text{Cov}\left[\rho_v^j \sum_{k=0}^{a-j} \rho_v^k \varepsilon_{i,a-j-k}, \sum_{k=0}^{a-j} \rho_v^k \varepsilon_{i,a-j-k}\right] + \mathbf{1}_{\{j=0\}} \sigma_z^2 \\ = &\sigma_\theta^2 \left(\sum_{k=0}^{a} \rho_u^k\right) \left(\sum_{k=0}^{a-j} \rho_u^k\right) + \sigma_{\tilde{u}}^2 \rho_u^{2(a+1)-j} + \sigma_{\tilde{v}}^2 \rho_v^{2(a+1)-j} + \sigma_\varepsilon^2 \rho_w^j \sum_{k=0}^{a-j} \rho_w^{2k} + \sigma_z^2 0^j \end{aligned}$$

This gives Equation (2) in the main text.[1]

[1] Because age a and by extension $h = a - j$ are finite, this covariance is well-defined even in the presence of unit roots. If we had restricted each persistence parameter to be strictly less than one in absolute value then the covariance expression would further simplify to:

$$\sigma_\theta^2 \frac{\left(1 - \rho_u^{a+1}\right)\left(1 - \rho_u^{a-j+1}\right)}{\left(1 - \rho_u\right)^2} + \sigma_{\tilde{u}}^2 \rho_u^{2a-j+2} + \sigma_{\tilde{v}}^2 \rho_v^{2a-j+2} + \sigma_\varepsilon^2 \rho_w^j \frac{1 - \rho_w^{2(a-j+1)}}{1 - \rho_w^2} + \sigma_z^2 0^j.$$

B.2 Estimation details

B.2.1 Estimation on the microdata

The reduced-form model is estimated using a minimum distance procedure, following Chamberlain (1984). We formulate its estimation and inference using GMM. Let $\vartheta = (\rho_u, \rho_v, \rho_w, \sigma_\theta, \sigma_{\tilde{u}}, \sigma_{\tilde{v}}, \sigma_\varepsilon, \sigma_z)$ be an arbitrary parameter vector in a compact parameter space. We define a random variable for firm i when observed at ages a and $a-j$:

$$f(n_{i,a}, n_{i,a-j}, \vartheta) \equiv (\ln n_{i,a} - E[\ln n_{i,a}])(\ln n_{i,a-j} - E[\ln n_{i,a-j}]) - \text{Cov}[\ln n_{i,a}, \ln n_{i,a-j}; \vartheta],$$

where $\text{Cov}[\ln n_{i,a}, \ln n_{i,a-j}; \vartheta]$ is a scalar computed from the statistical model, Equation (2) of the main text, with parameters ϑ. Then, for each $a = 0, \ldots, A$ and $0 \leq j \leq a$ we stack these random variables to form the random vector of length $K = \frac{A(A-1)}{2}$:

$$f(n_i, \vartheta) \equiv [f(n_{i,a}, n_{i,a-j})],$$

where n_i is a vector of firm i (residual) employment at each age $a = 0, \ldots, A$.[2] The moment conditions we exploit in the estimation are $E[f(n_i; \vartheta)] = 0$. These are satisfied when the autocovariance matrix of $\ln n_{i,a}$ for $a = 0, \ldots, A$ is equal to the autocovariance matrix of the statistical model for a particular combination of parameters $\vartheta = \vartheta_0$.

To operationalize the estimator we define $\tilde{f}(n_{i,a}, \vartheta)$ that replaces in f each $E[\ln n_{i,a}]$ with its sample average, and then let

$$\tilde{g}_N(\vartheta) \equiv \frac{1}{N} \sum_i \tilde{f}(n_i, \vartheta).$$

Then, the sample average $\tilde{g}_N(\vartheta)$ of $\tilde{f}$ is the empirical autocovariance structure less the statistical model's autocovariance structure for parameter ϑ.[3] The estimator solves $\min_\vartheta \tilde{g}_N(\vartheta)' W \tilde{g}_N(\vartheta)$, where W is a $K \times K$ weighting matrix. We set $W = I$, the identity matrix, so this is an equally-weighted minimum distance (EWMD) estimator.[4]

[2] When estimating the model, we use residualized log employment $\ln \tilde{n}_{ia}$ as described in Section A.2.

[3] We write $\tilde{g}$ this way for simplicity. However, N will actually vary across each element of $\tilde{g}$, since the number of firm observations may vary across each a and $a-j$ pair.

[4] A common alternative, c.f., Guvenen (2009) and Blundell, Pistaferri, and Preston (2008), is diagonally-weighted minimum distance (DWMD) that adjusts only for the heteroskedasticity induced by the different number of observations used to calculate each moment. We include the EWMD

The estimator $\widehat{\vartheta}$ follows, asymptotically, a normal distribution with a mean equal to the true value of ϑ and a covariance matrix given by $\Sigma = (D'D)\, D'\Omega D\, (D'D)^{-1}$, where $D = E[\frac{\partial f(n_i,\vartheta)}{\partial \vartheta}]$ and $\Omega = E[f\,(n_i,\vartheta)\, f\,(n_i,\vartheta)']$. We estimate Σ using the sample analogues $\widetilde{D} = \frac{1}{N}\sum_i \frac{\partial \widetilde{f}(n_i,\vartheta)}{\partial \vartheta}$ and $\widetilde{\Omega} = \frac{1}{N}\sum_i \widetilde{f}\,(n_i,\vartheta)'\, \widetilde{f}\,(n_i,\vartheta)$.

B.2.2 Replication using the included autocovariance matrices

Finally, note that the point estimates can be computed using just the empirical autocovariance matrices we provide without accessing the underlying microdata.[5] With $W = I$, the estimator simply chooses parameters to minimize the squared distance between the model's autocovariance matrix and the empirical autocovariance matrix. This can also be implemented for any other statistical model whose parameters are identified from its covariance structure, and we explore some alternatives in the next section. The codes to estimate the benchmark and alternative models from just the reported autocovariance matrices are included in the replication files.

Inference, however, requires estimating the Jacobian matrix $\widetilde{D}$ at the optimum and the variance of the GMM model $\widetilde{\Omega}$ on the microdata, both for the particular statistical model f. Thus, the standard errors in Tables 1 and G.12 can only be replicated using the microdata. These programs are archived by the Census Bureau and may be requested for any sworn researcher with an approved LBD project.

B.3 Alternative statistical models

B.3.1 Details on the restricted models in the main text

In the main text and its discussion of Figure 3, we consider 4 restricted versions of the baseline model. The purpose of these restrictions is to show the importance of each component of the baseline model. For reference, in Table B.1 we report the EWMD parameter estimates under these restrictions using the autocovariance of both the full and balanced panels.

results because we cannot release the exact number of observations used in each moment. In practice, there is little difference between the EWMD and DWMD estimates.

[5] Tables A.1 and A.2 provide the empirical autocovariance of *firm-level* (residual) log employment, unbalanced and balanced panels, respectively. Tables G.10 and G.11 are the *establishment-level* counterparts.

Table B.1: EWMD parameter estimates of restricted models in main text Figure 3

	(1) Base	(2) Case I. $u = z = 0$	(3) Case II. $w = z = 0$	(4) Case III. $v = z = 0$	(5) Case IV. $z = 0$
A. Unbalanced Panel					
ρ_u	0.2741	0.9718	0.7138	0.1695	0.2439
ρ_v	0.8536	0.9718	—	—	0.8009
ρ_w	0.9458	0.9718	—	0.9182	0.9303
σ_θ	0.5339	—	0.2939	0.6408	0.5627
σ_u	1.4837	—	1.555	3.7862	1.6747
σ_v	0.6928	0.8865	—	—	0.7608
σ_ε	0.2858	0.2902	—	0.3199	0.316
σ_z	0.2863	—	—	0.2663	—
$RMSE$	0.0151	0.0427	0.1212	0.0347	0.0239
B. Balanced Panel					
ρ_u	0.2184	0.9771	0.9476	0.123	0.1973
ρ_v	0.8323	0.9771	—	—	0.8045
ρ_w	0.9625	0.9771	—	0.9435	0.9521
σ_θ	0.5545	—	0.0853	0.6465	0.5678
σ_u	1.7425	—	1.0277	5.0117	2.0373
σ_v	0.6951	0.8304	—	—	0.7446
σ_ε	0.2548	0.2676	—	0.2766	0.2778
σ_z	0.2716	—	—	0.2754	—
$RMSE$	0.012	0.0368	0.0839	0.0333	0.0212

Note: EWMD estimation of the restricted models presented in Figure 3 (panels I. to IV.) using 210 empirical moments from Census LBD. Column (1) baseline estimates; column (2) I. no ex-ante profiles, just single initial condition; column (3) II. no ex-post shocks; (4) III. restricted ex-ante profile with single initial condition; (5) IV. no i.i.d. component ($\sigma_z = 0$). $RMSE$ reports $\sqrt{SSR/210}$ where SSR is the equally weighted sum of the 210 squared residuals.

B.3.2 Additional statistical models summary

Our baseline model captures the information in the empirical autocovariance function well and has the benefit of parsimony and a structural interpretation in an industry dynamics equilibrium under some fairly rigid assumptions. However, there are numerous other statistical models that can also be estimated, and some have structural shock counterparts that have been suggested in the prior literature. Like our baseline, given the orthogonality restrictions on the shocks and their functional forms, the parameters of these alternative statistical models are identified from the empirical autocovariance function.

We estimate 8 additional alternative models using the same equally weighted minimum distance (EWMD) procedure that minimizes the squared distance between the statistical model and the empirical autocovariance function (separately for the unbalanced and balanced panel). Including the benchmark, we consider the following models:

1. Benchmark process (Generalized AR(1) in main text)
2. Generalized AR(1) with unit root in w process
3. Generalized AR(1) with stationary w process and additional random walk x term
4. Generalized AR(1) with age dependent shocks ε_a
5. AR(1) with non stationary initial condition as in Hopenhayn and Rogerson (1993)[6]
6. Separate AR(1) and FE terms in $\ln n$ with non stationary initial condition for AR(1) term
7. AR(1) dynamic panel data model
8. AR(2) dynamic panel data model
9. ARMA(1,1) dynamic panel data model

We report the estimation results in Table B.2, where the column numbers correspond to the model number in the list of alternatives. For reference, we report each model's autocovariance function required for the minimum distance estimation below in Section B.3.3 where we also include in Figures B.2 to B.4 plots of each additional model's estimated autocovariance structure plotted against the data.

As a parsimonious representation of the full autocovariance function, our baseline process performs better than many common alternatives in terms of model fit. Several observations are in order.

First, our benchmark process does not estimate any permanent ex-post shocks. This is a feature of the data, not the process. Because it matches the autocovariance

[6] Note that this version is identical to model I in Section 2.4., Figure 3. We repeat it here because it is a popular specification in the literature.

Table B.2: EWMD estimation of alternative firm employment processes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
							Dynamic Panel Data Models		
	Base	RW	RW+Base	Age Dep.	AR(1)	AR(1) + FE	AR(1)	AR(2)	ARMA(1,1)
A. Unbalanced Panel									
ρ_u	0.2741	0.6567	0.2741	0.0001	0.9718	—	—	—	—
ρ_v	0.8536	0.9485	0.8536	0.8299	0.9718	0.9493	—	—	—
ρ_w/ρ	0.9458	1	0.9457	0.9585	0.9718	0.9493	0.9718	0.6518	0.9754
ρ_2	—	—	—	—	—	—	—	0.3117	—
σ_θ	0.5339	0.2061	0.5339	0.6752	—	0.5932	0.0000	0.0000	0.0035
σ_u	1.4837	0.3389	1.4837	—	—	—	—	—	—
σ_v	0.6928	0.8631	0.6928	0.7615	0.8865	0.6168	0.8865	1.1749	0.8410
σ_ε	0.2858	0.1968	0.2858	0.2716	0.2902	0.3140	0.2902	0.3763	0.4749
σ_x	—	—	0.0020	—	—	—	—	—	—
σ_z	0.2863	0.3867	0.2862	0.2984	—	—	—	—	—
γ	—	—	—	—	—	—	—	—	-0.4237
$RMSE$	0.0151	0.0292	0.0151	0.0082	0.0427	0.0405	0.0427	0.0353	0.0346
Ex-ante var (%)									
5 years	57.12	64.71	57.12	19.33	55.88	54.53	55.88	54.33	49.42
10 years	46.90	50.45	46.90	19.89	37.18	41.02	37.18	36.74	34.45
20 years	41.21	32.75	41.21	24.67	19.50	31.38	19.50	19.67	19.49
50 years	38.79	14.62	38.79	31.86	2.87	26.27	2.87	3.03	4.15
B. Balanced Panel									
ρ_u	0.2184	0.5853	0.2199	0.0002	0.9771	—	—	—	—
ρ_v	0.8323	0.9608	0.8245	0.8123	0.9771	0.9716	—	—	—
ρ_w/ρ	0.9625	1	0.9491	0.9694	0.9771	0.9716	0.9749	0.684	0.9756
ρ_2	—	—	—	—	—	—	—	0.2817	—
σ_θ	0.5545	0.2142	0.5572	0.6669	—	0.3781	0.0179	0.0316	0.0181
σ_u	1.7425	0.7402	1.7305	—	—	—	—	—	—
σ_v	0.6951	0.7709	0.6992	0.7605	0.8304	0.7309	0.8420	1.1019	0.8004
σ_ε	0.2548	0.2020	0.2407	0.2476	0.2676	0.2732	0.2641	0.3313	0.4429
σ_x	—	—	0.0945	0.2760	—	—	—	—	—
σ_z	0.2716	0.3313	0.2660	—	—	—	—	—	—
γ	—	—	—	—	—	—	—	—	-0.4394
$RMSE$	0.0120	0.0191	0.0119	0.0083	0.0368	0.0367	0.0367	0.032	0.0289
Ex-ante var (%)									
5 years	58.30	63.42	58.47	17.83	57.63	57.22	59.03	58.82	52.37
10 years	47.09	47.83	47.64	18.12	39.60	40.53	41.97	43.53	39.05
20 years	39.72	29.48	40.11	21.54	22.20	25.48	27.32	31.59	27.25
50 years	34.93	11.22	31.5	27.39	4.32	11.96	19.82	30.32	21.84
# Params	8	7	9	27	3	4	4	5	5

Note: EWMD estimation of each process using 210 empirical moments from Census LBD unbalanced and balanced panels. Column (1) presents the estimates from the main text; column (2) imposes a unit root in the w term; column (3) imposes stationary (in the long run) w term with $\rho_w < 1$ and adds a unit root shock $x' = x + \xi$; column (4) allows the "ex-post" shocks ε_{ia} to depend on age—reported σ_ε is the average of $\sigma_{\varepsilon a}$ for $a = 0, \ldots, 19$, see Figure B.5—and normalizes $\sigma_u = 0$; column (5) presents the estimates for an AR(1) with an initial condition as in Hopenhayn and Rogerson (1993); column (6) presents the estimates for a process with an AR(1) and a FE term where the FE term is not included in the AR(1); columns (7) to (9) present the estimates for alternative dynamic panel data models of log employment allowing for a non stationary initial condition $\ln n_{-1}$. $RMSE$ reports $\sqrt{SSR/210}$ where SSR is the equally weighted sum of the 210 squared residuals. It does not adjust for the number of estimated parameters. Number of parameters reports the number of unrestricted parameters estimated in each process.

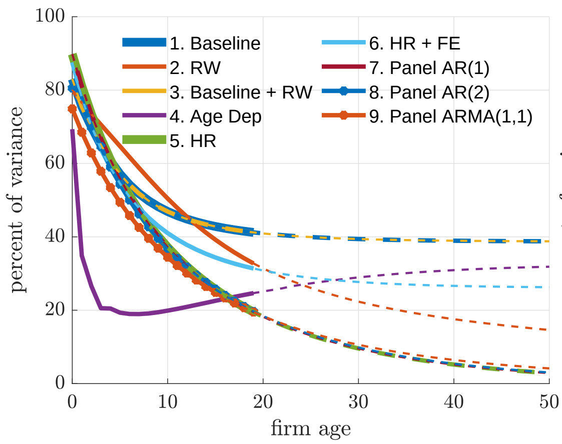
function over a finite horizon, there is no need to require a-priori a bounded autocovariance function in the limit. While our estimated shocks are persistent, the data strongly reject a unit root. Imposing a unit root (column 2) significantly degrades the model fit. Even when augmented with both a unit root and a stationary persistent process for w_a (column 3), matching the autocovariance requires near zero volatility unit root term. This finding may be of independent interest since it suggests a trade-off when identifying firm dynamics by matching "early" life cycle dynamics as we do using the autocovariance function and long run restrictions such as the matching the ergodic size distribution.

Next, the canonical process in the literature struggles to match the covariance structure. This process (column 5) fit is considerably worse compared to the baseline, with a root mean squared error that is two to four times as high as that in our baseline). In this process, any ex-ante heterogeneity imparted by the initial condition vanishes over time, so all firms expect to be the same in the long run. Without any permanent ex-ante differences, matching long-run autocovariances with an AR(1) requires a very persistent process that imposes a nearly linear increase in variance and decay in autocovariance.

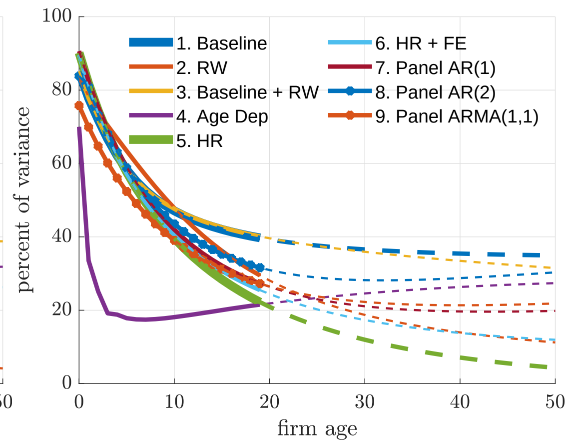
Interestingly, even when augmented with a separate fixed effect term (column 6) or specified as a dynamic panel (column 7), both of which allow long-run heterogeneity, there is little improvement in fit. The reason is there is not enough flexibility to match the age dependence in autocovariance we observe in the data. For example, the correlation of employment now with employment 5 years prior is much higher for a 10 year old firm than for a 5 year old firm. Higher order terms, such as the AR(2) (column 8) recommended by Lee and Mukoyama (2015) to match the dynamics of plants in the manufacturing sector, allow some additional flexibility in the rate of decay in autocovariance, but do little to improve the overall fit. Allowing the ex-post shocks to be serially correlated (column 9) using an ARMA(1) also does little to improve the fit.

One obvious way to improve on our model fit, at the cost of introducing many free parameters, is to relax the restriction on homoskedastic ex-post shocks ε. Introducing age-dependent shocks (column 4) can, unsurprisingly, produce a near perfect model fit. Yet, even with this flexibility in the ex-post component, the model still needs ex-ante profile heterogeneity. After 50 years, the ex-ante components are explaining about 30 percent of the variation.

The importance of ex-ante profile heterogeneity is a robust feature of the data. To better compare across models, in Figure B.1, we plot for each model the fraction of variance explained by the ex-ante component over the life cycle. This plot fills the in horizons in between the ages listed in Table B.2 and is the counterpart to Figure 4 from the main text. We observe that regardless of process, ex-ante characteristics explain a significant fraction of early lifecycle employment dispersion. For the processes that are best able to match the shape of the autocovariance function (see Figures B.2 to B.4 and discussion below), they also attribute a nearly identical share of long-run variance to ex-ante characteristics reinforcing our findings for the benchmark generalized AR(1) process.



(a) Unbalanced panel



(b) Balanced panel

Figure B.1: Each model's "ex-ante" share of variance for each firm age

B.3.3 Expressions and discussion of each statistical model

We derive the autocovariance function for each statistical model. First, we consider alternative sets of restrictions on the generalized AR(1) process. Next, we consider alternative dynamic panel data models. Since not all dynamic panel data models are restricted cases of our generalized AR

1. Benchmark model. The autocovariance function of benchmark model is derived in Section A.2 and is repeated here:

$$\text{Cov}\left[\log n_{i,a}, \ln n_{i,a-j}\right] = \sigma_\theta^2 \left(\sum_{k=0}^{a} \rho_u^k\right)\left(\sum_{k=0}^{a-j} \rho_u^k\right) + \sigma_{\tilde{u}}^2 \rho_u^{2(a+1)-j} + \sigma_{\tilde{v}}^2 \rho_v^{2(a+1)-j} + \sigma_\varepsilon^2 \rho_w^j \sum_{k=0}^{a-j} \rho_w^{2k} + \sigma_z^2 0^j.$$

Note that this function allows for random walks in each of the terms u, v, and w, i.e., ρ_w, ρ_v and ρ_w are not required to be strictly less than 1. These estimates are in column (1) in Table B.2 and the implied autocovariance function is plotted against the empirical autocovariance function in Figure B.2 panels (a) and (b).

2. Permanent shocks. If we restrict $\rho_w = 1$ so that the ex-post term is a random walk then the MA representation from Equation (B.1) is:

$$\ln n_{i.a} = \sum_{k=0}^{a} \rho_u^k \theta_i + \rho_u^{a+1} u_{i,-1} + \rho_v^{a+1} v_{i,-1} + \sum_{k=0}^{j-1} \varepsilon_{i,a-k} + \sum_{k=0}^{a-j} \varepsilon_{i,a-j-k} + z_{i,a}.$$

The autocovariance function is:

$$\begin{aligned}\text{Cov}\left[\log n_{i,a}, \log n_{i,a-j}\right] =& \sigma_\theta^2 \left(\sum_{k=0}^{a} \rho_u^k\right)\left(\sum_{k=0}^{a-j} \rho_u^k\right) + \sigma_{\tilde{u}}^2 \rho_u^{2(a+1)-j} + \sigma_{\tilde{v}}^2 \rho_v^{2(a+1)-j} \\ &+ (a-j+1)\,\sigma_\varepsilon^2 + \sigma_z^2 0^j.\end{aligned}$$

We report the estimated parameters of this process in column (2) of Table B.2. The implied autocovariance function is plotted against the empirical autocovariance function in Figure B.2 panels (c) and (d).

3. Permanent and persistent shocks. We consider a variant on this process (not just a restriction) where $\ln n_{ia}$ has both a persistent w and permanent x ex-post component

$$\ln n_{i,a} = u_{i,a} + v_{i,a} + w_{i,a} + x_{i,a} + z_{i,a}$$

with $\rho_w < 1$ and

$$x_{i,a} = x_{i,a-1} + \xi_{i,a}.$$

Then the moving average representation is:

$$\ln n_{i,a} = \sum_{k=0}^{a} \rho_u^k \theta_i + \rho_u^{a+1} u_{i,-1} + \rho_v^{a+1} v_{i,-1} + \sum_{k=0}^{j-1} \rho_w^k \varepsilon_{i,a-k} + \rho_w^j \sum_{k=0}^{a-j} \rho_w^k \varepsilon_{i,a-j-k} + \sum_{k=0}^{a} \xi_{i,a-k} + z_{i,a}.$$

The autocovariance function is:

$$\begin{aligned}\mathrm{Cov}\left[\log n_{i,a}, \log n_{i,a-j}\right] = \sigma_\theta^2 \left(\sum_{k=0}^{a} \rho_u^k\right)\left(\sum_{k=0}^{a-j} \rho_u^k\right) + \sigma_{\tilde{u}}^2 \rho_u^{2(a+1)-j} + \sigma_{\tilde{v}}^2 \rho_v^{2(a+1)-j} \\ + \rho_w^j \frac{1-\rho_w^{2(a-j+1)}}{1-\rho_w^2} \sigma_\varepsilon^2 + (a-j+1)\,\sigma_\xi^2 + \sigma_z^2 0^j.\end{aligned}$$

We report the estimated parameters for this process in column (3) of Table B.2. The implied autocovariance function is plotted against the empirical autocovariance function in Figure B.2 panels (e) and (f).

4. Age dependent shocks $\sigma_{\varepsilon a}$. We relax the assumption in the benchmark model that σ_ε^2 does not vary by age and instead estimate the model with the following alternative parameters $\sigma_{\varepsilon 0}^2, \ldots, \sigma_{\varepsilon A}^2$. This additional flexibility also replicates some of the features of the "ex-ante" profile—the first three terms of Equation (B.1)—so we normalize $\sigma_u = 0$. With these changes, the autocovariance process is:

$$\mathrm{Cov}\left[\log n_{i,a}, \log n_{i,a-j}\right] = \sigma_\theta^2 \left(\sum_{k=0}^{a} \rho_u^k\right)\left(\sum_{k=0}^{a-j} \rho_u^k\right) + \sigma_{\tilde{v}}^2 \rho_v^{2(a+1)-j} + \rho_w^j \sum_{k=0}^{a-j} \rho_w^{2k} \sigma_{\varepsilon,a-j-k}^2 + \sigma_z^2 0^j.$$

We report the estimated parameters for this process in column (4) of Table B.2. In the table we report for σ_ε, the mean over all ages $\overline{\sigma_\varepsilon} = \sum_{a=0}^{19} \sigma_{\varepsilon a}$. Figure B.5 plots this volatility by firm age. The implied autocovariance function is plotted against the empirical autocovariance function in Figure B.3 panels (a) and (b).

5. AR(1) with non stationary initial condition (Hopenhayn and Rogerson, 1993). This AR(1) is a special case of Equation (B.1) with constant σ_ε, no i.i.d. component, i.e. $\sigma_z = 0$, no permanent ex-ante heterogeneity $\theta_i = \theta$ ($\sigma_\theta = 0$), $u_{i,-1} = 0$ ($\sigma_{\tilde{u}} = 0$), and common persistence $\rho_u = \rho_v = \rho_w = \rho < 1$:

$$\ln n_{i,a} = \theta + \rho \ln n_{i,a-1} + \varepsilon_{i,a}$$

for $a \geq 0$ with $\ln n_{i,-1} = v_{i,-1}$. Ex-ante heterogeneity is present only through this initial condition and fades over time. The autocovariance function is:

$$\mathrm{Cov}\left[\log n_{i,a}, \log n_{i,a-j}\right] = \sigma_{\tilde{n}}^2 \rho^{2(a+1)-j} + \sigma_{\varepsilon}^2 \rho^j \frac{1-\rho^{2(a-j+1)}}{1-\rho^2}.$$

We report the estimated parameters for this process in column (5) of Table B.2. The implied autocovariance function is plotted against the empirical autocovariance function in Figure B.3 panels (c) and (d). For the balanced panel, the estimated autocovariance function is also plotted as case I of Figure 3 in the main text.

6. Separate AR(1) and FE terms. We also estimate a restricted version of Equation (B.1) that includes an AR(1) term as above and a separate fixed effect term. Relative to above, the instead sets $\rho_u = 0$ and $\sigma_\theta \geq 0$, with the autocovariance function

$$\mathrm{Cov}\left[\log n_{i,a}, \log n_{i,a-j}\right] = \sigma_\theta^2 + \sigma_{\tilde{n}}^2 \rho^{2(a+1)-j} + \sigma_{\varepsilon}^2 \rho^j \frac{1-\rho^{2(a-j+1)}}{1-\rho^2}.$$

We report the estimated parameters for this process in column (6) of Table B.2. The implied autocovariance function is plotted against the empirical autocovariance function in Figure B.3 panels (e) and (f). Note that this process is *not* a dynamic panel data model since the FE is not part of the AR term. Next we consider dynamic panel data models, which incorporate the fixed effect in the AR or ARMA term.

Dynamic panel data models. A dynamic panel data model is

$$\ln n_{i,a} = \sum_{l=1}^{L} \rho_l \ln n_{i,a-l} + \theta_i + \sum_{k=0}^{K} \gamma_k^k \varepsilon_{i,a-k} \tag{B.2}$$

for $a \geq 0$ with initial condition $\ln n_{-1} \equiv \ln \tilde{n}$ and where $\ln \theta_i$ and $\ln \tilde{n}$ are iid with $\mathrm{Var}\left[\theta_i\right] = \sigma_\theta^2$, $\mathrm{Var}\left[\ln \tilde{n}\right] = \sigma_{\tilde{n}}^2$, $\mathrm{Var}\left[\varepsilon_a\right] = \sigma_\varepsilon^2$, and $\gamma_k \geq 0$. Note that in some special cases of this dynamic panel data model, namely $L = 1$ and $K = 0$, then Equation (B.2) is a special case of Equation (B.1) with $u_{i,-1} + v_{i,-1} = \ln \tilde{n}$ and $\rho_u = \rho_v = \rho_w = \rho_1 < 1$. We estimate several common specifications of Equation (B.2). Although there are a variety of linear estimators for some of these models, to be consistent with our main results, we continue to use the nonlinear minimum distance estimator that chooses parameters to match the empirical autocovariance function.

7. AR(1) dynamic panel. In this specification $L = 1$ and $K = 0$, so that it is as noted above a special case of (B.1). The autocovariance function is

$$\mathrm{Cov}\left[\ln n_{i,a}, \ln n_{i,a-j}\right] = \sigma_\theta^2 \frac{1-\rho^{a+1}}{1-\rho}\frac{1-\rho^{a-j+1}}{1-\rho} + \rho^{2(a+1)-j}\sigma_{\tilde{n}}^2 + \sigma_\varepsilon^2 \rho^j \frac{1-\rho^{2(a-j+1)}}{1-\rho^2}.$$

We report the estimated parameters for this process in column (7) of Table B.2. The implied autocovariance function is plotted against the empirical autocovariance function in Figure B.4 panels (a) and (b).

8. AR(2) dynamic panel. We relax the AR(1) restriction on the dynamics and now set $L = 2$, then

$$\ln n_{i,a} = \rho_1 \ln n_{i,a-1} + \rho_2 \ln n_{i,a-2} + \theta_i + \varepsilon_{i,a}, \tag{B.3}$$

for $a \geq 0$ with $\ln n_{-1} = \ln \tilde{n}$ and $\ln n_{-2} = 0$. Inverting to the MA representation with the initial condition is tricky. Instead we express it an AR(1) in $z_a = (\ln n_a, \ln n_{a-1})^\intercal$. Then for $a \geq 0$

$$z_{i,a} = A z_{i,a-1} + B\theta_i + C\varepsilon_{i,a}$$

where

$$A = \begin{bmatrix} \rho_1 & \rho_2 \\ 1 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad C = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

with $z_0 = (\ln n_0, \ln \tilde{n})^\intercal$ and initial condition $z_{-1} = (\ln \tilde{n}, 0)^\intercal$. Then

$$\begin{aligned} (I - AL)\, z_{ia} =& B\theta_i + C\varepsilon_{ia} \\ (I + AL + \ldots + A^a L^a)(I - AL)\, z_{ia} =& (I + AL + \ldots + A^a L^a)(B\theta_i + C\varepsilon_{ia}) \\ z_{ia} =& A^{a+1} z_{i,-1} + \theta_i \sum_{k=0}^{a} A^k B + \sum_{k=0}^{a} A^k C \varepsilon_{ia-k} \end{aligned}$$

Using this MA representation, the autocovariance function may be determined by

$$\mathrm{Cov}\left[z_{i,a}, z_{i,a-j}\right] = A^{a+1} \begin{bmatrix} \sigma_{\tilde{n}}^2 & 0 \\ 0 & 0 \end{bmatrix} A^{\intercal a+1-j} + \sigma_\theta^2 \left(\sum_{k=0}^{a} A^k\right) BB^\intercal \left(\sum_{k=0}^{a-j} A^{\intercal k}\right) + \sigma_\varepsilon^2 \sum_{k=0}^{a-j} A^k CC^\intercal A^{\intercal k+j}.$$

We are specifically interested in the upper left element of this matrix

$$\text{Cov}\left[\ln n_{i,a}, \ln n_{i,a-j}\right] = e' \text{Cov}\left[z_{i,a}, z_{i,a-j}\right] e$$

where $e = (1, 0)^\intercal$. We report the estimated parameters for this process in column (8) of Table B.2. The implied autocovariance function is plotted against the empirical autocovariance function in Figure B.4 panels (c) and (d).

9. ARMA(1,1) dynamic panel. We next relax the assumption of iid persistent shocks and allow for serial correlation, so $L = 1$ and $K = 1$. Then

$$\ln n_{i,a} = \rho \ln n_{i,a-1} + \theta_i + \varepsilon_{i,a} + \gamma \varepsilon_{i,a-1} \tag{B.4}$$

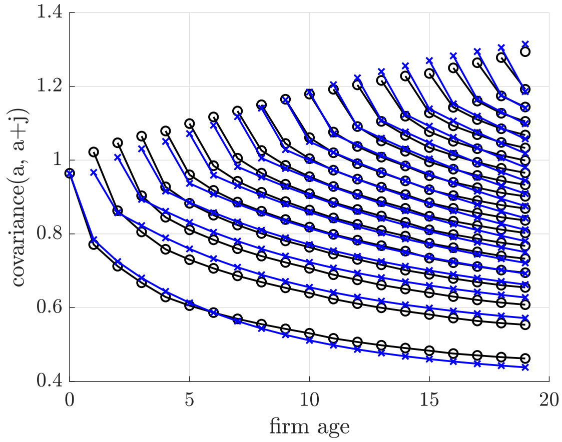
for $a \geq 0$ with $\varepsilon_{i,-1} = 0$. The moving average representation is:

$$\begin{aligned} \ln n_{i,a} &= \sum_{k=0}^{a} \rho^k \theta_i + \rho^{a+1} \ln \tilde{n}_i + \sum_{k=0}^{a} \rho^k \left(\varepsilon_{ia-k} + \gamma \varepsilon_{i,a-k-1}\right) \\ &= \sum_{k=0}^{a} \rho^k \theta_i + \rho^{a+1} \ln \tilde{n}_i + \sum_{k=0}^{a} \left(\rho^{k-1} \left(\rho + \gamma\right)\right)^{k>0} \varepsilon_{i,a-k} \end{aligned}$$

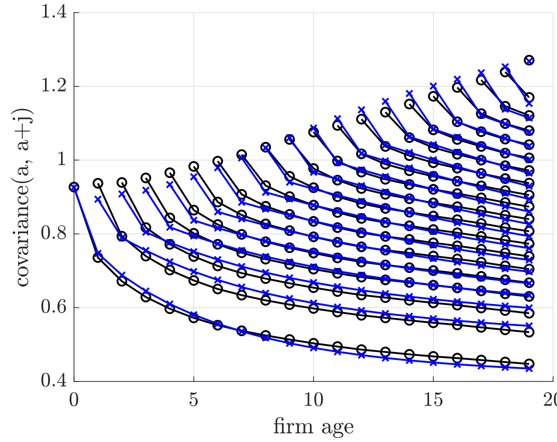
Then the autocovariance function for $a, j \geq 0$ is:

$$\begin{aligned} \text{Cov}\left[\ln n_{i,a}, \ln n_{i,a-j}\right] = \sigma_\theta^2 \frac{1 - \rho^{a+1}}{1 - \rho} \frac{1 - \rho^{a+1-j}}{1 - \rho} + \sigma_{\tilde{n}}^2 \rho^{2(a+1)-j} \\ + \left[\left(\rho^{j-1} \left(\rho + \gamma\right)\right)^{j>0} + \mathbf{1}_{a \geq 1} \left(\gamma + \rho\right)^2 \rho^j \frac{1 - \rho^{2(a-j)}}{1 - \rho^2} \right] \sigma_\varepsilon^2 \end{aligned} \tag{B.5}$$

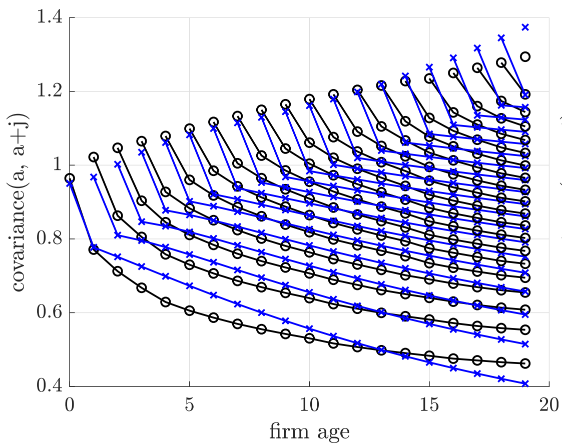
We report the estimated parameters for this process in column (9) of Table B.2. The implied autocovariance function is plotted against the empirical autocovariance function in Figure B.4 panels (e) and (f).



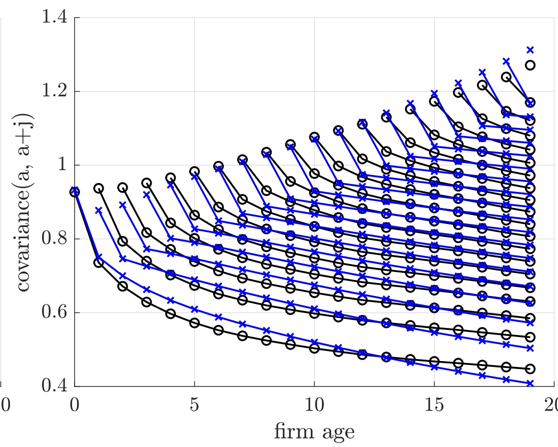
(a) 1. Benchmark for unbalanced panel



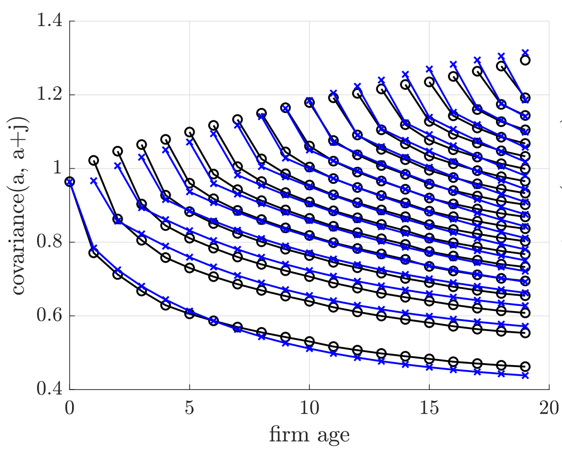
(b) 1. Benchmark for balanced panel



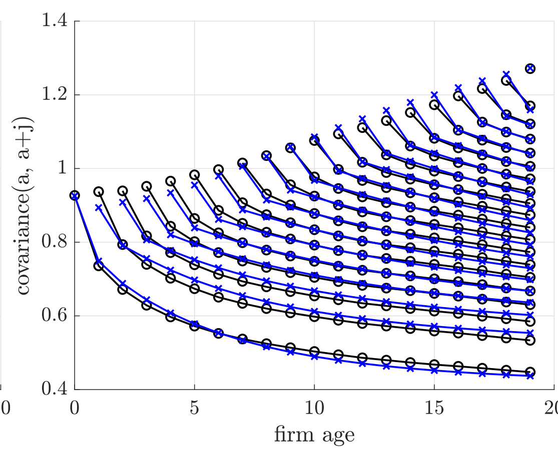
(c) 2. "Ex-post" random walk for unbalanced panel



(d) 2. "Ex-post" random walk for balanced panel

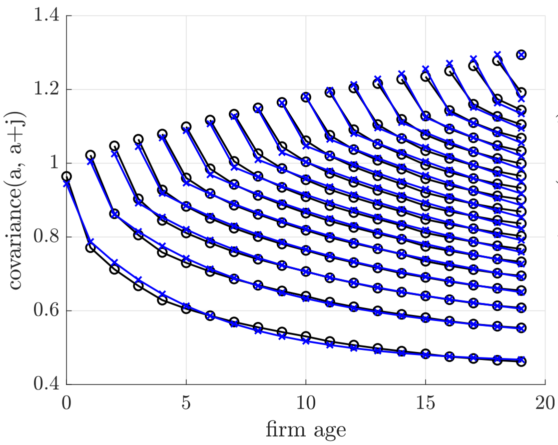


(e) 3. Benchmark plus random walk for unbalanced panel

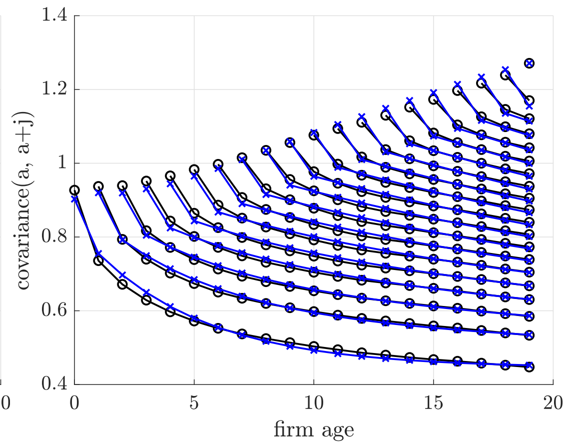


(f) 3. Benchmark plus random walk for balanced panel

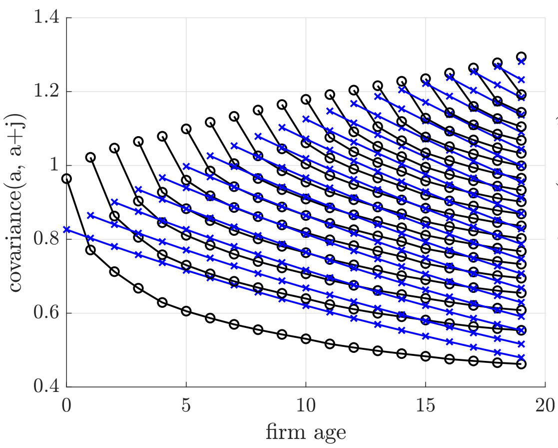
Figure B.2: Empirical autocovariance and estimated model fit for models 1 to 3



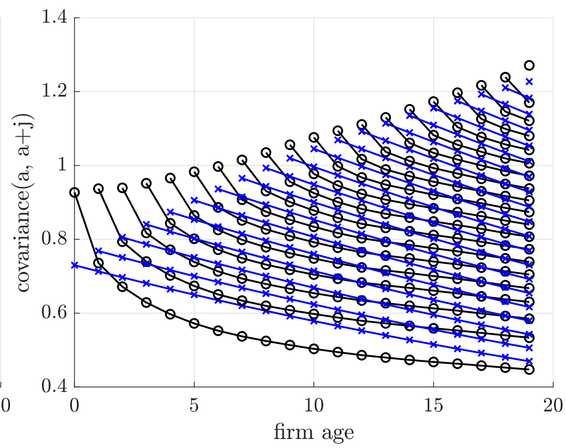
(a) 4. Age dependent shocks for unbalanced panel



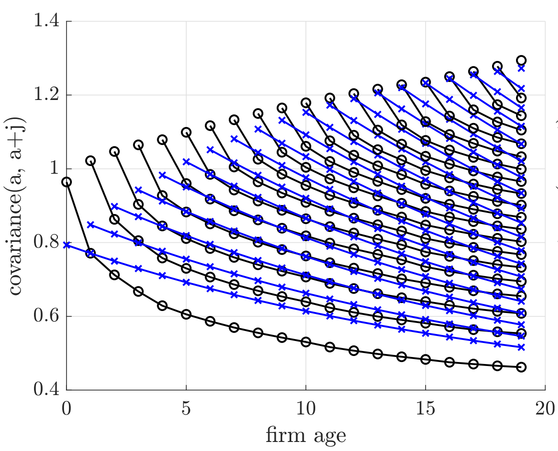
(b) 4. Age dependent shocks for balanced panel



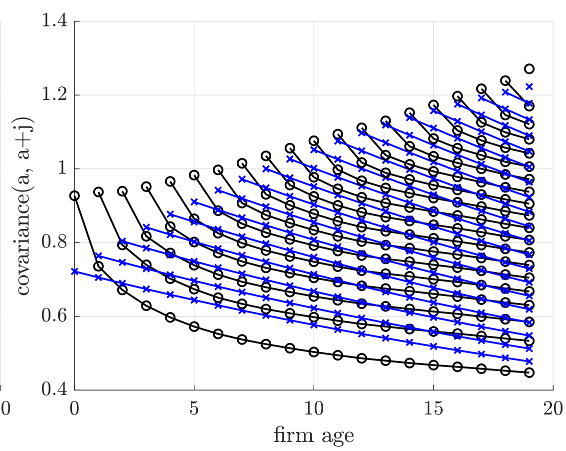
(c) 5. AR(1) for unbalanced panel



(d) 5. AR(1) for balanced panel

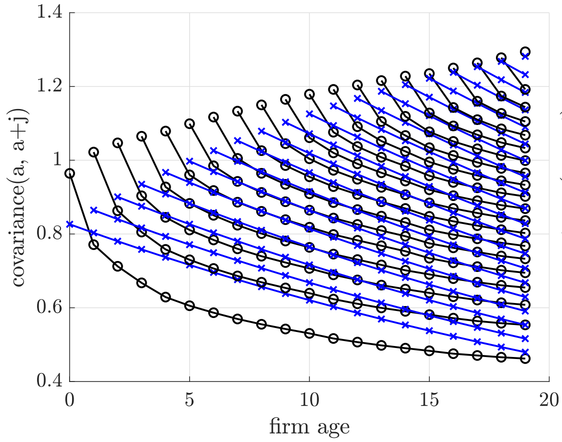


(e) 6. Plus separate FE term unbalanced panel

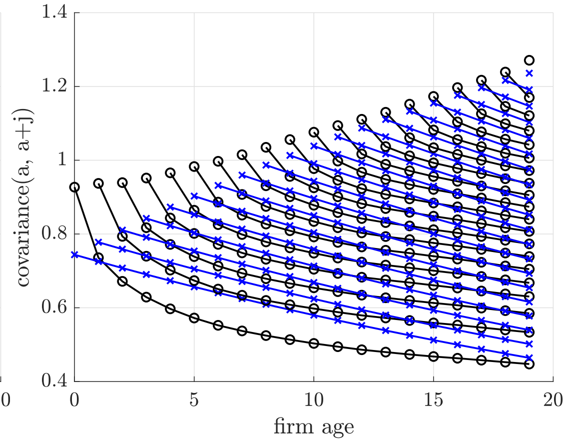


(f) 6. Plus random walk for balanced panel

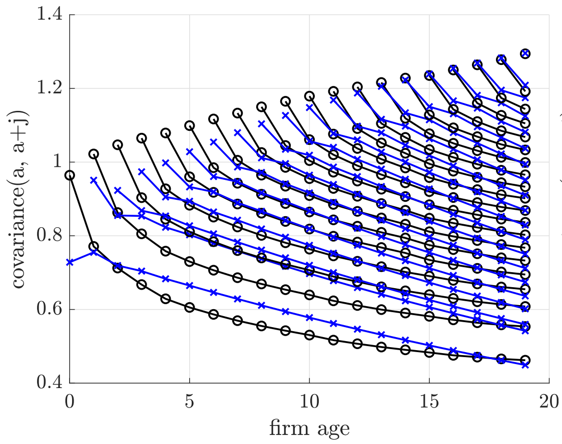
Figure B.3: Empirical autocovariance and estimated model fit for models 4 to 6



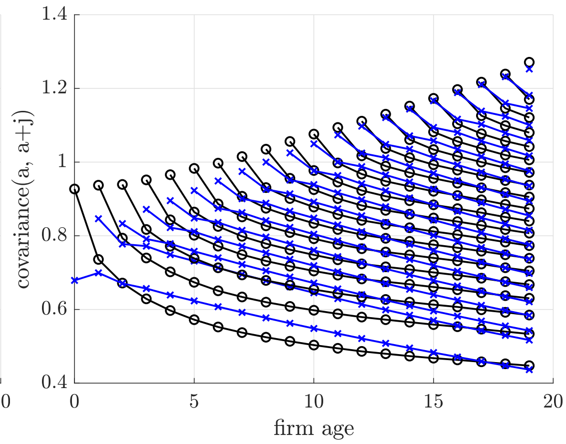
(a) 7. AR(1) dynamic panel model for unbalanced



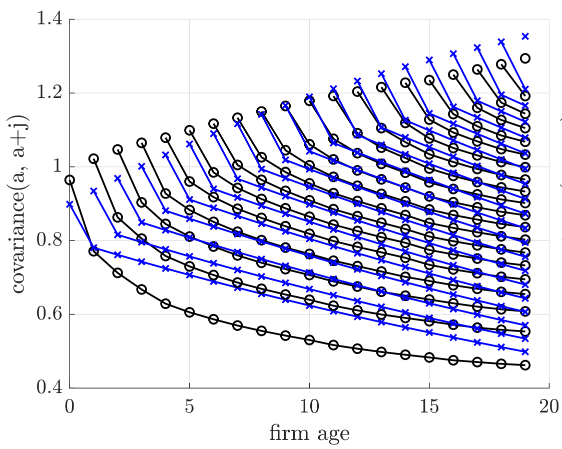
(b) 7. AR(1) dynamic panel model for balanced



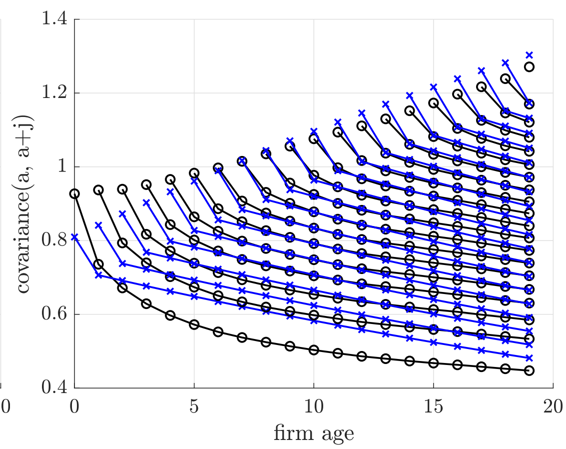
(c) 8. AR(2) dynamic panel model for unbalanced



(d) 8. AR(2) dynamic panel model for unbalanced

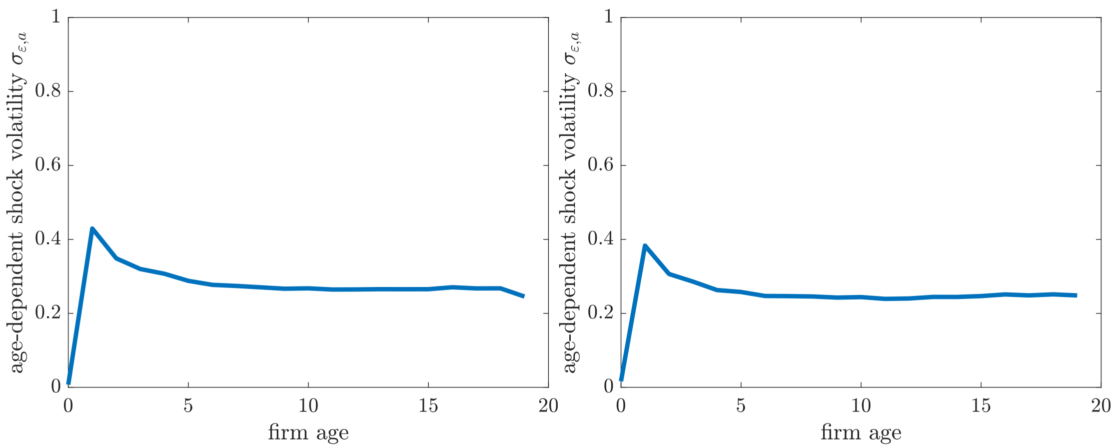


(e) 9. ARMA(1,1) dynamic panel model for unbalanced



(f) 9. ARMA(1,1) dynamic panel model for unbalanced

Figure B.4: Empirical autocovariance and estimated model fit for models 7 to 9



(a) Unbalanced (b) Balanced

Figure B.5: Alternative estimates of shock volatility by firm age $\sigma_{\varepsilon a}$

C Sectoral heterogeneity in the statistical model

We now repeat by sector the estimation for a large set of roughly 2-digit NAICS sectors and for a high-tech sector composed of industries with a high intensity in STEM occupations (drawing from manufacturing, information, and administrative support NAICS sectors, see Heckler, 2005). For each sector, we follow the procedure in Appendix Section A.2 to estimate the sector-specific residual autocovariance matrix, and then follow the procedure in Section B.2 to estimate the parameters of the sector-specific model. We present these estimates in Tables C.3 and C.4 for firms and establishments, respectively.

The importance of ex-ante heterogeneity in dispersion of employment by firm age is widespread across all sectors. Tables C.3 and C.4 also report the fraction of variance in log employment at ages 5, 10, 20, and 50 accounted for by the ex-ante component; Figures C.6 and C.7 show the share for all ages in between. Across all sectors, the sector-level patterns are consistent with the overall pattern in Figure 4. Even quantitatively the results at each horizon are very similar across sectors, with the exception of the Arts/Entertainment (NAICS 71) and Accomodationa/Food (NAICS 72) service sectors in which ex-ante heterogeneity plays a more pronounced role at all horizons. The distinction between firm and establishment also makes little difference except for the manufacturing (NAICS 31-33) and retail trade (NAICS 44-45) sectors where ex-ante heterogeneity plays a significantly more important role at

Table C.3: EWMD estimation of *firm* employment processes by sector

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	CON	MFG	WHL	RET	TWH	INF	FIR	PRO	ADW	HLT	AES	AFS	OTH	HTC
A. Unbalanced Panel														
ρ_u	0.2579	0.2197	0.3082	0.1368	0.2170	0.3476	0.3289	0.5579	0.3249	0.4252	0.2026	0.0876	0.3064	0.5430
	(0.0055)	(0.0074)	(0.0095)	(0.0050)	(0.0107)	(0.0274)	(0.0083)	(0.0099)	(0.0111)	(0.0078)	(0.0145)	(0.0048)	(0.0064)	(0.0168)
ρ_v	0.8118	0.8616	0.8816	0.8495	0.8348	0.8699	0.8575	0.9660	0.8367	0.8918	0.7878	0.8249	0.8954	0.9207
	(0.0084)	(0.0052)	(0.0080)	(0.0052)	(0.0221)	(0.0271)	(0.0069)	(0.0012)	(0.0124)	(0.0035)	(0.0327)	(0.0087)	(0.0039)	(0.0100)
ρ_w	0.9585	0.9553	0.9568	0.9495	0.9697	0.9484	0.9455	0.9564	0.9424	0.9324	0.9593	0.9657	0.9503	0.9599
	(0.0007)	(0.0009)	(0.0010)	(0.0006)	(0.0007)	(0.0032)	(0.0012)	(0.0010)	(0.0014)	(0.0011)	(0.0018)	(0.0010)	(0.0007)	(0.0016)
σ_θ	0.5274	0.6406	0.4940	0.5803	0.6666	0.5184	0.4432	0.2664	0.5414	0.3748	0.7278	0.8528	0.4211	0.3636
	(0.0042)	(0.0071)	(0.0084)	(0.0040)	(0.0109)	(0.0269)	(0.0063)	(0.0068)	(0.0096)	(0.0063)	(0.0141)	(0.0049)	(0.0053)	(0.0153)
$\sigma_{\tilde{u}}$	1.0730	2.4100	1.4090	2.2690	1.5670	1.3520	1.0880	0.8131	1.1600	1.0830	1.9110	4.7060	1.1410	0.8727
	(0.0921)	(0.0867)	(0.0760)	(0.0937)	(0.2078)	(0.3235)	(0.0770)	(0.0265)	(0.1896)	(0.0316)	(0.3346)	(0.2760)	(0.0373)	(0.0816)
$\sigma_{\tilde{v}}$	0.6394	0.8147	0.6624	0.5988	0.5704	0.7845	0.6850	0.6435	0.8274	0.8258	0.6698	0.6408	0.6287	0.6764
	(0.0172)	(0.0126)	(0.0165)	(0.0088)	(0.0411)	(0.0703)	(0.0166)	(0.0054)	(0.0368)	(0.0081)	(0.0728)	(0.0172)	(0.0064)	(0.0205)
σ_ε	0.2678	0.3149	0.3014	0.2561	0.2904	0.3653	0.2710	0.2702	0.3426	0.2932	0.2553	0.2306	0.2564	0.3324
	(0.0012)	(0.0014)	(0.0016)	(0.0004)	(0.0006)	(0.0064)	(0.0014)	(0.0014)	(0.0019)	(0.0011)	(0.0023)	(0.0019)	(0.0008)	(0.0025)
σ_z	0.3758	0.3178	0.2828	0.2891	0.2993	0.3212	0.2637	0.2869	0.3266	0.2608	0.3507	0.3266	0.2587	0.3028
	(0.0019)	(0.0036)	(0.0037)	(0.0018)	(0.0063)	(0.0142)	(0.0028)	(0.0021)	(0.0046)	(0.0016)	(0.0078)	(0.0029)	(0.0014)	(0.0058)
$RMSE$	0.0154	0.0155	0.0139	0.0128	0.0186	0.0204	0.0151	0.0171	0.0219	0.0168	0.0188	0.0160	0.0117	0.0206
Ex-ante var (%)														
5 years	52.19	57.5	53.71	56.18	59.17	50.58	55.48	58.49	52.98	57.29	65.92	70.83	55.63	54.76
10 years	42.95	46	41.81	46.25	48.23	39.37	44.65	47.36	42.98	45.07	57.69	62.43	43.49	43.25
20 years	36.61	38.71	34.17	40.15	39.41	32.97	38.55	36.32	37.58	38.39	51.09	55.05	35.88	34.2
50 years	33.29	35.49	30.85	37.71	33.46	30.71	36.47	28.77	35.8	36.95	47.29	50.02	33.21	30.02
B. Balanced Panel														
ρ_u	0.2125	0.1873	0.2606	0.1102	0.1747	0.2818	0.2609	0.2410	0.2447	0.2614	0.1967	0.0953	0.2460	0.3501
	(0.0056)	(0.0073)	(0.0086)	(0.0047)	(0.0114)	(0.0224)	(0.0084)	(0.0049)	(0.0116)	(0.0065)	(0.0151)	(0.0048)	(0.0054)	(0.0127)
ρ_v	0.8002	0.8406	0.8215	0.8271	0.8371	0.8604	0.8627	0.7949	0.8122	0.8406	0.7516	0.8239	0.8591	0.7697
	(0.0055)	(0.0050)	(0.0062)	(0.0039)	(0.0102)	(0.0097)	(0.0046)	(0.0056)	(0.0087)	(0.0033)	(0.0257)	(0.0061)	(0.0034)	(0.0132)
ρ_w	0.9642	0.9784	0.9793	0.9574	0.9734	0.9806	0.9558	0.9576	0.9576	0.9644	0.9436	0.9482	0.9533	0.9777
	(0.0008)	(0.0011)	(0.0010)	(0.0009)	(0.0018)	(0.0031)	(0.0012)	(0.0009)	(0.0015)	(0.0008)	(0.0035)	(0.0016)	(0.0008)	(0.0014)
σ_θ	0.5546	0.6578	0.5236	0.5938	0.6823	0.5626	0.4712	0.5513	0.5768	0.4567	0.7413	0.8533	0.4663	0.5148
	(0.0043)	(0.0070)	(0.0068)	(0.0038)	(0.0117)	(0.0222)	(0.0066)	(0.0041)	(0.0095)	(0.0045)	(0.0146)	(0.0053)	(0.0041)	(0.0102)
$\sigma_{\tilde{u}}$	1.5170	2.5780	1.5120	2.9790	2.0710	1.5300	1.3010	1.4010	1.5650	1.3500	2.0660	4.7560	1.3710	1.0310
	(0.0475)	(0.0933)	(0.0537)	(0.1145)	(0.1337)	(0.1092)	(0.0481)	(0.0373)	(0.0879)	(0.0361)	(0.2003)	(0.2075)	(0.0325)	(0.0823)
$\sigma_{\tilde{v}}$	0.6770	0.7726	0.7154	0.6398	0.7331	0.7393	0.6663	0.6685	0.7938	0.7182	0.7199	0.6771	0.6187	0.8322
	(0.0085)	(0.0083)	(0.0105)	(0.0054)	(0.0160)	(0.0090)	(0.0079)	(0.0088)	(0.0166)	(0.0061)	(0.0447)	(0.0092)	(0.0048)	(0.0294)
σ_ε	0.2591	0.2648	0.2476	0.2418	0.2805	0.2837	0.2560	0.2552	0.3126	0.2322	0.2844	0.2606	0.2476	0.2892
	(0.0013)	(0.0022)	(0.0020)	(0.0012)	(0.0033)	(0.0054)	(0.0017)	(0.0013)	(0.0028)	(0.0012)	(0.0051)	(0.0021)	(0.0011)	(0.0030)
σ_z	0.3555	0.2698	0.2694	0.2539	0.2944	0.3234	0.2491	0.2441	0.3138	0.2523	0.3285	0.2676	0.2438	0.2827
	(0.0015)	(0.0033)	(0.0029)	(0.0017)	(0.0052)	(0.0087)	(0.0025)	(0.0020)	(0.0042)	(0.0018)	(0.0075)	(0.0030)	(0.0016)	(0.0053)
$RMSE$	0.0156	0.0125	0.0126	0.0099	0.0153	0.0193	0.0122	0.0142	0.0194	0.0114	0.0182	0.0146	0.0106	0.0199
Ex-ante var (%)														
5 years	53.17	62.06	57.59	58.21	59.86	56.4	55.93	59.54	52.44	57.08	64.49	70.6	55.72	55.3
10 years	43.34	49.47	44.87	47.63	47.83	43.06	43.98	49.47	41.77	44.68	57.12	62.59	44.35	43.38
20 years	36.19	39.31	34.62	40.56	38.28	32.45	36.57	42.39	34.96	36.48	52.15	56.96	37.4	33.58
50 years	31.92	31.72	26.49	37.03	31.45	24.26	33.34	38.73	31.61	31.92	50.23	54.54	34.49	26

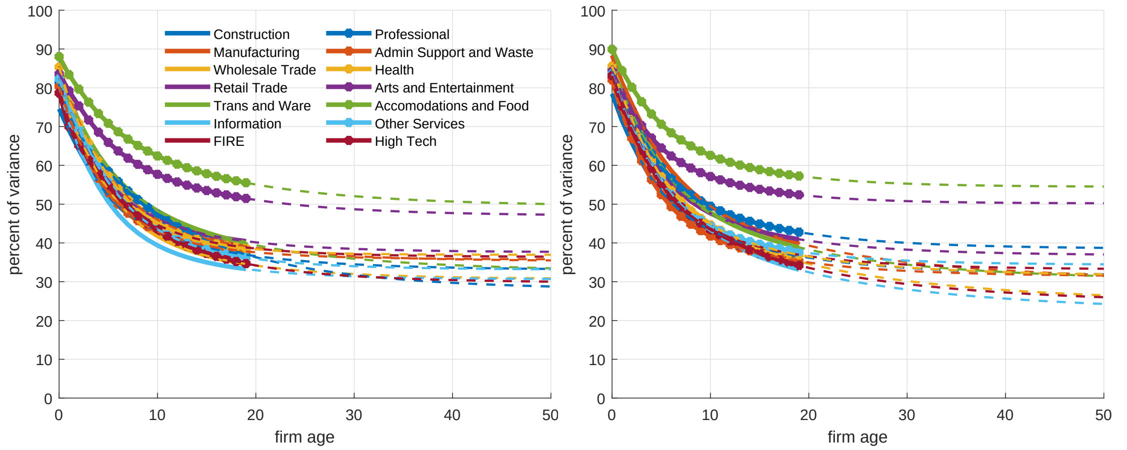
Note: EWMD estimated parameters using residual log *firm* employment autocovariance matrix for each of the following sectors (NAICS codes): construction (23), manufacturing (31-33), wholesale trade (42), retail trade (44-45), transportation and warehousing (48-49), information (51), finance insurance and real estate (52-53), professional/business/technical services (54), administrative/support/waste management services (56), health care and social assistance services (62), arts/entertainment/recreation (71), accommodation and food services (72), other services (81); high tech sector spans multiple NAICS sectors based on intensity of STEM occupations (see Hecker 2005) and includes NAICS 3341, 3342, 3344, 3345, 5112, 5161, 5179, 5181, 5182, 5415, 3254, 3364, 5413, 5417. GMM standard errors in parenthesis. $RMSE$ reports $\sqrt{SSR/210}$ where SSR is the equally weighted sum of the 210 squared residuals.

Table C.4: EWMD estimation of *establishment* employment processes by sector

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	CON	MFG	WHL	RET	TWH	INF	FIR	PRO	ADW	HLT	AES	AFS	OTH	HTC
							A. Unbalanced Panel							
ρ_u	0.2540	0.1987	0.2144	0.1894	0.3741	0.2435	0.3002	0.4845	0.2818	0.3430	0.1989	0.1047	0.2640	0.4610
	(0.0057)	(0.0052)	(0.0066)	(0.0044)	(0.0277)	(0.0298)	(0.0061)	(0.0094)	(0.0096)	(0.0061)	(0.0127)	(0.0039)	(0.0058)	(0.0160)
ρ_v	0.8498	0.8564	0.9058	0.9142	0.9664	0.9497	0.8710	0.9572	0.8667	0.8947	0.8613	0.9161	0.9030	0.9508
	(0.0074)	(0.0049)	(0.0058)	(0.0032)	(0.0023)	(0.0095)	(0.0058)	(0.0013)	(0.0102)	(0.0028)	(0.0296)	(0.0047)	(0.0035)	(0.0034)
ρ_w	0.9579	0.9492	0.9482	0.9150	0.9678	0.9430	0.9231	0.9587	0.9282	0.9170	0.9496	0.8787	0.9474	0.9599
	(0.0006)	(0.0009)	(0.0011)	(0.0014)	(0.0013)	(0.0031)	(0.0013)	(0.0008)	(0.0017)	(0.0011)	(0.0031)	(0.0033)	(0.0008)	(0.0014)
σ_θ	0.5461	0.8086	0.6204	0.5640	0.4335	0.6476	0.5041	0.3308	0.6150	0.4360	0.7742	0.8076	0.4567	0.4437
	(0.0048)	(0.0061)	(0.0087)	(0.0050)	(0.0320)	(0.0607)	(0.0058)	(0.0082)	(0.0109)	(0.0056)	(0.0141)	(0.0058)	(0.0055)	(0.0189)
$\sigma_{\tilde{u}}$	1.1880	2.9750	2.0410	2.2710	0.7248	2.0390	1.3070	0.8554	1.3920	1.4430	2.3090	4.8360	1.2660	1.0440
	(0.0821)	(0.0827)	(0.0799)	(0.0531)	(0.1174)	(0.2674)	(0.0791)	(0.0329)	(0.2064)	(0.0320)	(0.2749)	(0.1800)	(0.0396)	(0.0749)
$\sigma_{\tilde{v}}$	0.6213	0.7767	0.7158	0.6502	0.9314	0.9299	0.7776	0.7564	0.8931	0.8504	0.5627	0.6502	0.6488	0.8611
	(0.0135)	(0.0093)	(0.0098)	(0.0043)	(0.0206)	(0.0396)	(0.0141)	(0.0066)	(0.0295)	(0.0062)	(0.0501)	(0.0071)	(0.0053)	(0.0154)
σ_ε	0.2690	0.2950	0.2959	0.2588	0.3086	0.3614	0.2810	0.2744	0.3651	0.3002	0.2527	0.2777	0.2545	0.3333
	(0.0010)	(0.0011)	(0.0012)	(0.0010)	(0.0026)	(0.0049)	(0.0012)	(0.0013)	(0.0023)	(0.0010)	(0.0039)	(0.0022)	(0.0007)	(0.0027)
σ_z	0.3727	0.3063	0.2631	0.2726	0.3003	0.3204	0.2627	0.2816	0.3261	0.2495	0.3453	0.2864	0.2588	0.2947
	(0.0019)	(0.0029)	(0.0024)	(0.0011)	(0.0051)	(0.0073)	(0.0018)	(0.0019)	(0.0037)	(0.0013)	(0.0066)	(0.0023)	(0.0012)	(0.0050)
$RMSE$	0.0149	0.0121	0.0119	0.0094	0.0137	0.0202	0.0145	0.0178	0.0206	0.0154	0.0149	0.0126	0.0098	0.0192
Ex-ante var (%)														
5 years	54.5	68.87	61.97	64.61	64.54	63.16	61.38	62.23	56.46	59.49	70.16	73.38	57.97	62.73
10 years	44.71	59.95	50.52	55.94	50.77	51.34	51.74	49.53	46.52	48.02	62.78	68.59	45.89	49.89
20 years	38.06	53.83	42.81	50.93	36.13	41.31	47.12	37.12	41.72	42.15	57.35	66.3	38.14	38.08
50 years	34.76	51.26	40.06	49.92	24.74	36.57	46.26	29.55	40.69	41.2	54.95	65.93	35.59	31.56
							B. Balanced Panel							
ρ_u	0.2101	0.2034	0.1703	0.1124	0.1761	0.2759	0.2791	0.2336	0.2137	0.2339	0.2050	0.0641	0.2312	0.2728
	(0.0054)	(0.0055)	(0.0057)	(0.0037)	(0.0089)	(0.0342)	(0.0066)	(0.0043)	(0.0091)	(0.0053)	(0.0130)	(0.0039)	(0.0048)	(0.0090)
ρ_v	0.7928	0.8370	0.8318	0.8370	0.8624	0.9539	0.8825	0.7875	0.8169	0.8336	0.8015	0.8445	0.8642	0.7823
	(0.0055)	(0.0036)	(0.0041)	(0.0025)	(0.0066)	(0.0056)	(0.0037)	(0.0051)	(0.0068)	(0.0027)	(0.0177)	(0.0043)	(0.0029)	(0.0096)
ρ_w	0.9634	0.9665	0.9703	0.9296	0.9594	0.9683	0.9397	0.9505	0.9492	0.9445	0.9317	0.8850	0.9471	0.9586
	(0.0008)	(0.0012)	(0.0010)	(0.0010)	(0.0017)	(0.0026)	(0.0010)	(0.0009)	(0.0016)	(0.0009)	(0.0035)	(0.0025)	(0.0008)	(0.0016)
σ_θ	0.5794	0.7940	0.6671	0.6105	0.7506	0.5742	0.4910	0.6082	0.6534	0.4941	0.7735	0.8479	0.4928	0.6998
	(0.0043)	(0.0063)	(0.0057)	(0.0030)	(0.0106)	(0.0588)	(0.0057)	(0.0039)	(0.0087)	(0.0038)	(0.0132)	(0.0043)	(0.0038)	(0.0099)
$\sigma_{\tilde{u}}$	1.5360	2.6660	2.3820	3.4540	2.7970	2.1030	1.5940	1.5410	1.8330	1.6040	2.2930	7.3850	1.5060	1.6560
	(0.0479)	(0.0686)	(0.0744)	(0.1034)	(0.1312)	(0.2590)	(0.0378)	(0.0372)	(0.0886)	(0.0387)	(0.1572)	(0.4094)	(0.0332)	(0.0718)
$\sigma_{\tilde{v}}$	0.6923	0.8595	0.7421	0.6609	0.8159	1.0260	0.6842	0.7062	0.8403	0.7700	0.6765	0.7252	0.6392	0.8730
	(0.0088)	(0.0071)	(0.0069)	(0.0041)	(0.0119)	(0.0355)	(0.0057)	(0.0087)	(0.0140)	(0.0055)	(0.0318)	(0.0071)	(0.0045)	(0.0192)
σ_ε	0.2597	0.2551	0.2404	0.2282	0.2974	0.3056	0.2538	0.2631	0.3183	0.2441	0.2797	0.2500	0.2485	0.3083
	(0.0013)	(0.0020)	(0.0017)	(0.0010)	(0.0031)	(0.0046)	(0.0012)	(0.0014)	(0.0027)	(0.0011)	(0.0045)	(0.0013)	(0.0010)	(0.0030)
σ_z	0.3562	0.2638	0.2441	0.2431	0.2724	0.2527	0.2449	0.2294	0.3164	0.2442	0.3267	0.2524	0.2408	0.2314
	(0.0015)	(0.0030)	(0.0026)	(0.0013)	(0.0052)	(0.0072)	(0.0016)	(0.0022)	(0.0041)	(0.0016)	(0.0063)	(0.0015)	(0.0014)	(0.0059)
$RMSE$	0.0148	0.0118	0.0093	0.0073	0.0135	0.0167	0.0084	0.0138	0.0174	0.0110	0.0149	0.0087	0.0093	0.0170
Ex-ante var (%)														
5 years	54.94	72.93	66.29	64.98	64.9	69.27	61.9	63.41	56.65	59.13	68.7	75.72	58.25	64.99
10 years	45.34	63.06	54.92	56.86	53.65	54.87	51.21	54.17	46.68	48.42	62.31	71.49	47.3	54.91
20 years	38.23	54.75	45.69	52.71	45.62	39.73	44.99	47.92	40.64	42.57	58.58	70.1	40.85	47.41
50 years	34	49.17	39.23	51.69	41.5	29.76	43.21	45.13	38.18	40.54	57.52	69.99	38.54	43.35

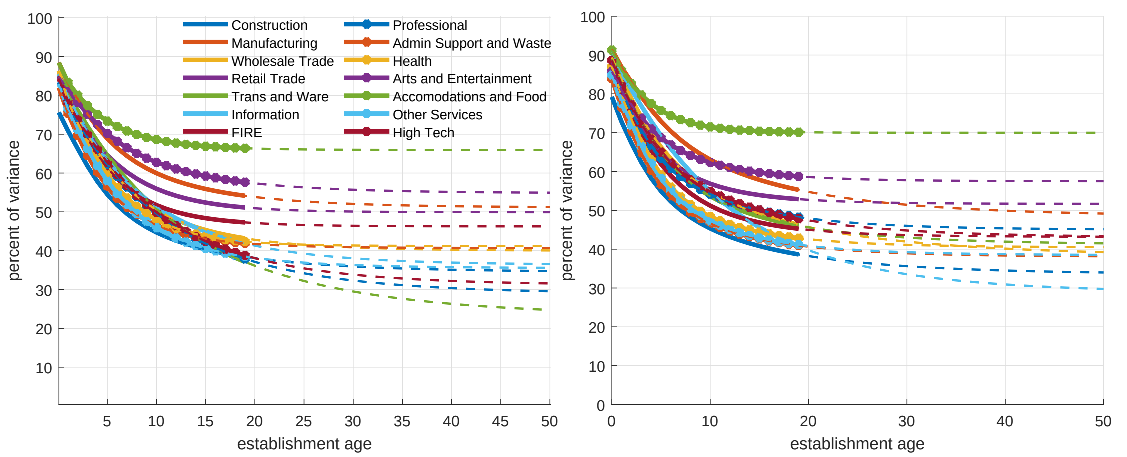
Note: EWMD estimated parameters using residual log *establishment* employment autocovariance matrix for each of the following sectors (NAICS codes): construction (23), manufacturing (31-33), wholesale trade (42), retail trade (44-45), transportation and warehousing (48-49), information (51), finance insurance and real estate (52-53), professional/business/technical services (54), administrative/support/waste management services (56), health care and social assistance services (62), arts/entertainment/recreation (71), accommodation and food services (72), other services (81); high tech sector spans multiple NAICS sectors based on intensity of STEM occupations (see Hecker 2005) and includes NAICS 3341, 3342, 3344, 3345, 5112, 5161, 5179, 5181, 5182, 5415, 3254, 3364, 5413, 5417. GMM standard errors in parenthesis. $RMSE$ reports $\sqrt{SSR/210}$ where SSR is the equally weighted sum of the 210 squared residuals.

the establishment level, which is consistent with these sectors' higher concentration of multi-establishment firms.



(a) Unbalanced (b) Balanced

Figure C.6: Ex-ante share of *firm*-level variance in log employment by sector



(a) Unbalanced (b) Balanced

Figure C.7: Ex-ante share of *establishment*-level variance in log employment by sector

D Structural model: additional details

D.1 Numerical solution of the structural model

Let us define $\widehat{\mu}(\mathbf{S}) \equiv \frac{\mu(\mathbf{S})}{M^e}$, which evolves as:

$$\widehat{\mu}(\mathbf{S}') = \int (1 - x(\mathbf{s})) F(\mathbf{S}'|\mathbf{s}) ((1-\delta)\widehat{\mu}(d\mathbf{s}) + G(d\mathbf{s})).$$

and note that in the stationary equilibrium $\mu(\mathbf{S}') = \mu(\mathbf{S})$. The labor market clearing condition in the stationary equilibrium can now be written as:

$$\bar{N} = M^e \left(\frac{\eta}{\eta - 1}\right)^{-\eta} w^{-\eta} Y \widetilde{\varphi} + M^e \widetilde{f} + M^e f^e,$$

where $\widetilde{\varphi} \equiv \int \varphi(\mathbf{s}) \widehat{\mu}(d\mathbf{s})$ and $\widetilde{f} \equiv \int f(1 - x(\mathbf{s}))((1-\delta)\widehat{\mu}(d\mathbf{s}) + G(d\mathbf{s}))$. Note further that $p_i = \frac{\eta}{\eta - 1}$ and that the wage is given as

$$w = P^{-1} = \frac{\eta - 1}{\eta} (M^e \widetilde{\varphi})^{\frac{1}{\eta - 1}}$$

We solve the model using the following algorithm (following Hopenhayn and Rogerson, 1993):

1. Solve for $Q \equiv w^{-\eta} Y$ from the free entry condition (i.e. guess Q, solve for the firm value functions, evaluate the free-entry condition, update the guess for Q and iterate until the condition holds with equality).

2. Normalize $M^e = 1$, simulate the model and compute $\widehat{\mu}(S)$, $\widetilde{\varphi}$ and $\widetilde{f}$.

3. Solve for M^e from the labor market clearing condition. Compute w, Y, and $\frac{Y}{N}$.

Note further that aggregate output can be written as:

$$Y = Qw^{\eta} = \chi (M^e \widetilde{\varphi})^{\frac{\eta}{\eta - 1}} = \chi \left(\int \varphi(\mathbf{s}) \mu(d\mathbf{s})\right)^{\frac{\eta}{\eta - 1}} = \Omega^{\frac{\eta}{\eta - 1}} \chi \left(\int \varphi(\mathbf{s}) \widetilde{\mu}(d\mathbf{s})\right)^{\frac{\eta}{\eta - 1}}$$

where $\widetilde{\mu}(d\mathbf{s}) \equiv \mu(d\mathbf{s}) / \Omega$ is the density of firms at state $\mathbf{s}$. It now follows that

$$Y = \Omega^{\frac{\eta}{\eta - 1}} \chi^{\frac{1}{1 - \eta}} \overline{n}^{\frac{\eta}{\eta - 1}}$$

where $\overline{n} = \int n(\mathbf{s})\,\widetilde{\mu}(d\mathbf{s})$ is average firm size and where we have used that $n_i = \chi\varphi_i$.

The state variables for an individual firm, in addition to its type θ_i, consist of the separate components of its demand fundamental: $u_{i,t}$, $v_{i,t}$, $w_{i,t}$, and $z_{i,t}$.[7] As mentioned in the main text, we restrict $\rho_v = \rho_w$, in which case the firm only needs to keep track of the sum $v_{i,t} + w_{i,t}$, rather than the two terms separately.

We allow for 31 grid points (equally spaced between -3 and 4) for the permanent component of the demand fundamental, θ. Similarly, we allow for 31 grid points (equally spaced between -5 and 7) for the initial condition $u_{i,-1}$. Finally, the process for w_i is discretized using the method of Rouwenhorst (1995) allowing for 31 grid points. The related initial condition $v_{i,-1}$, is drawn from those gridpoints. We use value function iteration to solve the firm's maximization problem on the grid specified above.

In simulating the economy, we use $100,000$ startups (i.e. firms which endogenously decide to remain in operation in the first period) and we follow these until the age of 20, consistent with the autocovariance data. Aggregate model variables are constructed using all surviving firms in the model.

D.2 Details on the restricted version of the model

In Section 4 of the main text, we introduce a restricted version of the baseline model. The restricted model imposes $\rho_u = \rho_v = \rho_w = \rho$ along with $\theta_i = \mu_\theta$ and $u_{i,-1} = 0$. Together these imply that the demand fundamental follows the widely-use AR(1) process with noise. As described, we reparametrize the model in the main text by targeting the same size and exit profiles, but replacing the autocovariance moments with the estimated persistence and volatility of an AR(1) for log employment as is common in the literature. We report here the details of this parametrization and its model fit.

Table D.1 shows the parametrization for both the baseline and the restricted version of the model. Figure D.1 depicts the model fit for the average size and exit profiles, which are targeted, and the autocovariance, which was not. The autocovariance is implicitly targeted through matching the persistence and dispersion parameters of the AR(1) for log employment. Reassuringly, the model-implied autocovariacne matrix is also very similar to the "reduced-form" counterpart shown in Section 2.4 when

[7] Note that $z_{i,t}$ is purely transitory and therefore its past values do not affect the decision of the firm.

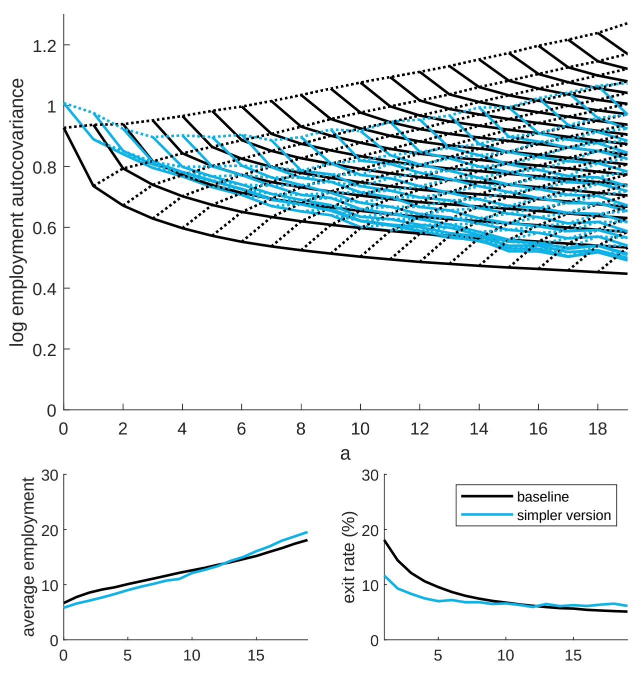
restricted to an AR(1) process. We show below that we find nearly identical results if we were instead to explicitly target the autocovariance matrix.

Table D.1: Parameter values

parameter		baseline	restricted
β	discount factor	0.96	0.96
η	elasticity of substitution	6.00	6.00
f^e	entry cost	0.44	0.20
f	fixed cost of operation	0.539	0.241
δ	exogenous exit rate	0.041	0.062
μ_θ	permanent component θ, mean	-1.762	0.284
σ_θ	permanent component θ, st. dev.	1.304	0
$\sigma_{\widetilde{u}}$	initial condition u_{-1}, st. dev.	1.572	0
$\sigma_{\widetilde{v}}$	initial condition v_{-1}, st. dev.	1.208	3.058
σ_ε	transitory shock ε, st. dev.	0.307	0.253
σ_z	noise shock z, st. dev.	0.203	0.260
ρ_u	permanent component, persistence	0.393	0.974
ρ_v	transitory component, persistence	0.988	0.974

Note: Top three parameters are calibrated as discussed in the main text. The remaining parameters are set as described above.

Figure D.1: Restricted model and data



Notes: Model fit (autocovariances of firm-level employment, life-cycle size and exit profiles) of the "restricted" model calibration. Counterpart to Figure 5 for the baseline model.

D.2.1 An alternative parametrization of the restricted model

The main text uses the life-cycle profiles of average size and exit rates together with the estimated dispersion and persistence of a simple AR(1) process for firm-level employment to parametrize the restricted model. While the calibration targets of average size and exit profiles are identical between the baseline and restricted models, one may wonder whether targeting the full autocovariance structure, as in the baseline model calibration, in place of the parameters of an AR(1) for employment would deliver similar results. In fact it does, because the restrictions on the process for the demand fundamental that rule out the heterogeneity in ex-ante profiles imply that it cannot take advantage of the additional information contained in the autocovariance matrix.

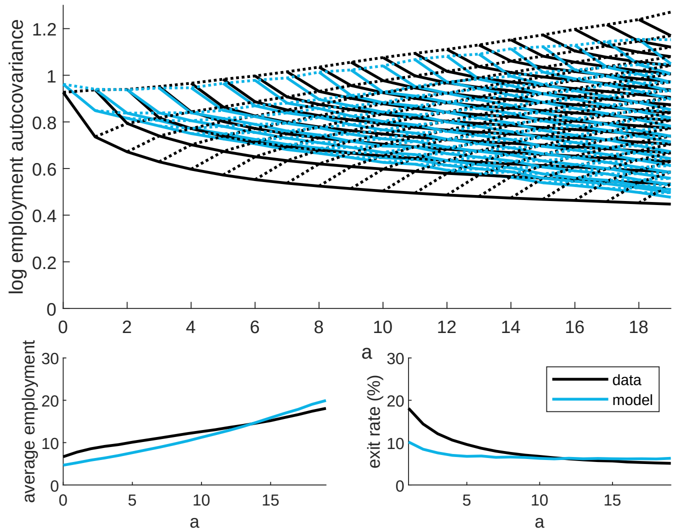
To see this, we recalibrate the restricted model using exactly the same targets as the baseline. These consist of the 210 moments in the upper triangle of the 20×20 autocovariance matrix, 20 moments in the average size profile (age 0 to 19), and 19 moments in the exit profile (age 1 to 19). The calibrated parameters change very little from the case in the main text that is calibrated to match the parameters of an AR(1) in log employment in place of the autocovariance matrix. Table D.2 compares the main text calibration of the restricted model against this alternative parametrization. The most notable change is in the dispersion of initial values, which falls somewhat.

Table D.2: Parameter values in the restricted model under two parameterizations

parameter		as in main text	alternative
β	discount factor	0.96	0.96
η	elasticity of substitution	6.00	6.00
f^e	entry cost	0.20	0.20
f	fixed cost of operation	0.241	0.241
δ	exogenous exit rate	0.062	0.062
μ_θ	permanent component θ, mean	0.284	0.284
σ_θ	permanent component θ, st. dev.	0	0
$\sigma_{\widetilde{u}}$	initial condition u_{-1}, st. dev.	0	0
$\sigma_{\widetilde{v}}$	initial condition v_{-1}, st. dev.	3.058	2.800
σ_ε	transitory shock ε, st. dev.	0.253	0.241
σ_z	noise shock z, st. dev.	0.260	0.260
ρ	persistence	0.974	0.973

Note: Top three parameters are calibrated as discussed in the main text. The remaining parameters are set such that the restricted model matches the same targets as the baseline.

Figure D.2: Model fit: restricted version under the alternative parametrization



Notes: The figure shows model fit (autocovariances of firm-level employment, life-cycle size and exit profiles) of the "restricted" model under the alternative parametrization.

Figure D.2 shows the model fit. The lower dispersion of initial conditions enables the model to match the level of the autocovariances slightly better, but it comes at a cost of missing some of the age-dependence of exit rates. Otherwise the targets are matched in a very similar manner as those under the parametrization used in the main text and displayed in Figure D.1.

Finally, Table D.3 shows the impact of financial frictions in the restricted model used in the main text and that under the alternative parametrization. Again, the results are very similar.

The similarity of the results to the parametrization in the main text follows because, conditional on a restricted shock process, the autocovariance matrix of employment provides little additional information over and above the persistence and dispersion of the AR(1) employment process. The restricted process simply does not have enough flexibility to match the autocovariance matrix. As discussed in the main text and the discussion of identification in Appendix D.4, the baseline model's ex-ante profiles feature (i) long-run ("steady state") differences between firms that allow it to match the long-horizon autocovariances (and which are identified by them); and (ii) transitory factors that allow it to adjust to fit the curvature of the autocovariances (which helps identify the dispersion and persistence of initial conditions). By contrast, lacking these features, the only way for the restricted process to match

Table D.3: Aggregate impact of financial frictions in the restricted model under two parameterizations (percent change)

	output	wage	size	exit	firms
	restricted model				
as in main text	−1.4	−0.8	+31.4	+3.8	−24.3
alternative parametrization	−1.4	−1.0	+30.9	+7.3	−23.9

Notes: Long-run impact of introducing financial frictions in the restricted model under two parametrizations: that in the main text and the alternative. The latter uses exactly the same targets as the baseline model. The former uses life-cycle size and exit profiles, but instead of the autocovariance of employment, it targets the persistence and dispersion from an AR(1) firm-level employment estimation. The financial constraint (ζ) is set to zero in both cases. Reported values are relative to the model without frictions. Output refers to aggregate production, wage is the real wage rate, size is average firm size, exit is the average exit rate and firms refers to the number of incumbent firms.

the long-horizon autocovariances (and therefore the persistent and large heterogeneity across firms in the long-run) is through the persistence of ex-post shocks, but this works against the relatively fast decline in the correlation of employment over the early part of firms' lifecycles. The AR(1) in log employment, which effectively matches the *unconditional* serial correlation of employment, is already extracting the relevant (for the restricted process) information in the full autocovariance matrix.

D.3 Details on split-sample results

This Appendix presents details on the parametrization and fit of the model in the split-sample analysis presented in Section 5 of the main text. Tables D.4 and D.5 show the parameter values for the two subsamples and Figures D.3 and D.4 document the model's fit across the two subsamples. From the two tables, it is apparent that most of the parameters remain relatively stable across the two sub-samples. However, the distribution of the permanent component θ, a key determinant of long-run size, is estimated to have changed. In particular, both the mean and the dispersion have declined going from the early to the late sample.

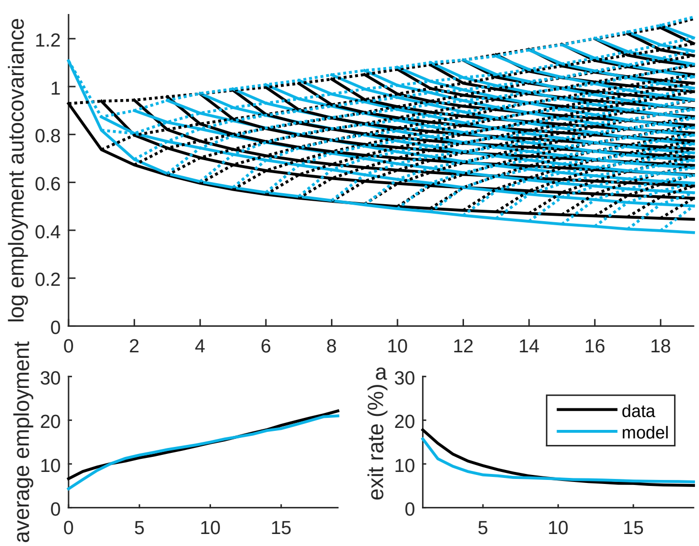
Skewness. Interestingly, despite that endogenous firm selection is the only source of skewness of firm growth rates in the model, our framework does well in accounting for the average level of skewness and especially its changes over time. In particular, Decker, Haltiwanger, Jarmin, and Miranda (2016) document an average skewness of

Table D.4: Parameter values (early sample)

parameter		value
		set a priori
β	discount factor	0.96
η	elasticity of substitution	6.00
f^e	entry cost	0.447
		estimated
f	fixed cost of operation	0.545
δ	exogenous exit rate	0.042
μ_θ	permanent component θ, mean	−1.770
σ_θ	permanent component θ, st. dev.	1.322
$\sigma_{\tilde{u}}$	initial condition u_{-1}, st. dev.	1.540
$\sigma_{\tilde{v}}$	initial condition v_{-1}, st. dev.	1.208
σ_ε	transitory shock ε, st. dev.	0.304
σ_z	noise shock z, st. dev.	0.153
ρ_u	permanent component, persistence	0.394
ρ_v	transitory component, persistence	0.987

Notes: parameter values. Top three parameters are calibrated as discussed in the main text. The remaining parameters are set such that the model matches the empirical autocovariance of employment and the age profiles of average size and exit rates from age 0 to 19.

Figure D.3: Targeted moments: data and structural model (early sample)



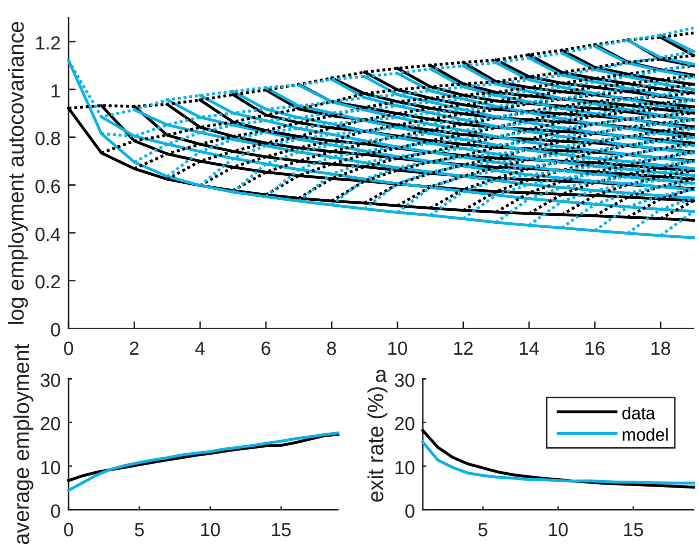
Notes: Top panel: Autocovariances of log employment between age $a = h + j$ and age $h \leq a$ in the data and the model, for a balanced panel of firms surviving up to at least age $a = 19$. Bottom left panel: Average employment by age a (unbalanced panel). Bottom right panel: exit rate by age a.

Table D.5: Parameter values (late sample)

parameter		value
		set a priori
β	discount factor	0.96
η	elasticity of substitution	6.00
f^e	entry cost	0.434
		estimated
f	fixed cost of operation	0.530
δ	exogenous exit rate	0.043
μ_θ	permanent component θ, mean	-1.846
σ_θ	permanent component θ, st. dev.	1.303
$\sigma_{\tilde{u}}$	initial condition u_{-1}, st. dev.	1.563
$\sigma_{\tilde{v}}$	initial condition v_{-1}, st. dev.	1.209
σ_ε	transitory shock ε, st. dev.	0.301
σ_z	noise shock z, st. dev.	0.195
ρ_u	permanent component, persistence	0.393
ρ_v	transitory component, persistence	0.987

Notes: parameter values. Top three parameters are calibrated as discussed in the main text. The remaining parameters are set such that the model matches the empirical autocovariance of employment and the age profiles of average size and exit rates from age 0 to 19.

Figure D.4: Targeted moments: data and structural model (late sample)



Notes: Top panel: Autocovariances of log employment between age $a = h + j$ and age $h \leq a$ in the data and the model, for a balanced panel of firms surviving up to at least age $a = 19$. Bottom left panel: Average employment by age a (unbalanced panel). Bottom right panel: exit rate by age a.

firm-level growth rates of 13 percent between 1979 and 2011.[8] Furthermore, they show that skewness declined by about 43 percent when comparing the first and last half of their sample. Our baseline specification delivers an average skewness of 5 percent over the entire sample period. Importantly, the changes in firm selection estimated in our split sample analysis imply a decline in skewness of firm-level growth rates of 35 percent.

D.4 Sources of identification of parameters in structural model

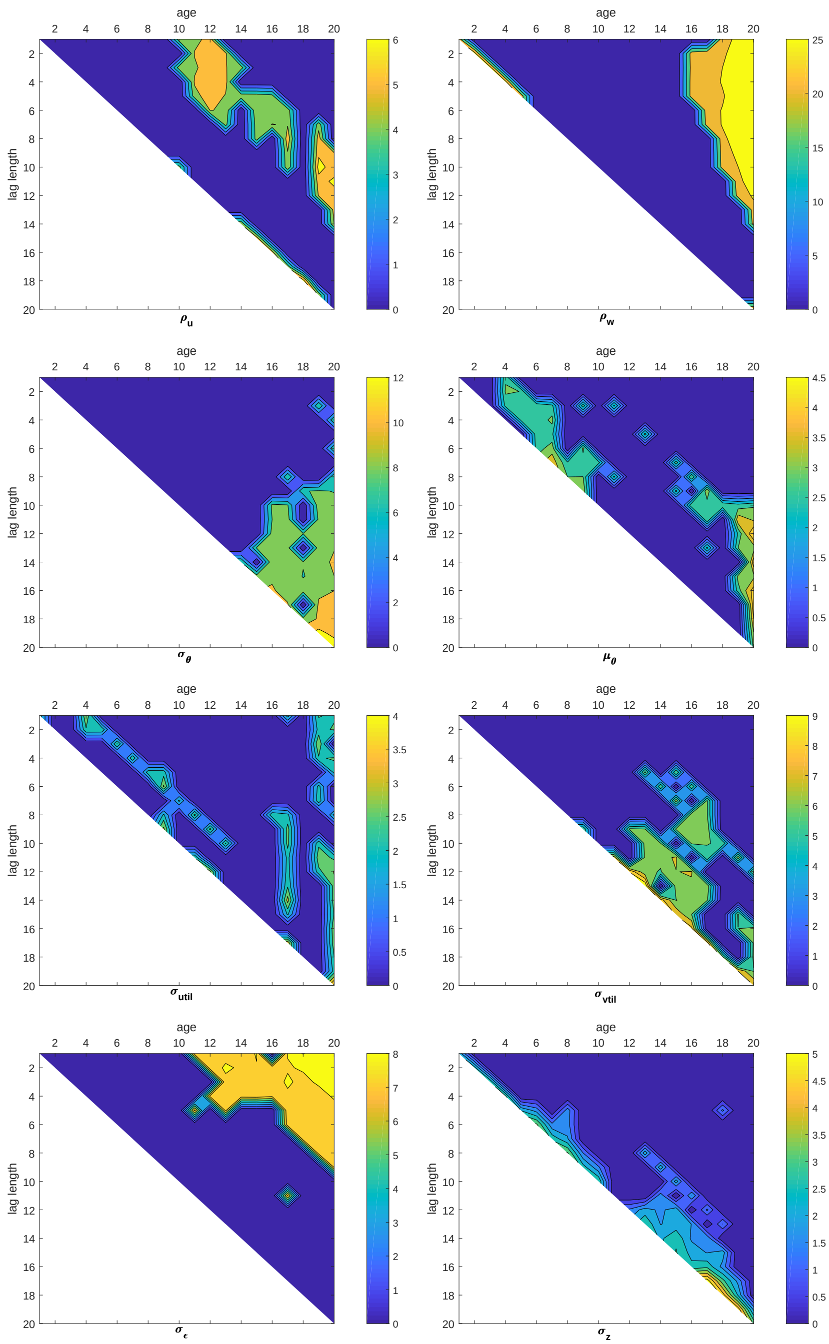
Section 2.4 in the main text provides a mapping between parameters of the shock process and the empirical autocovariance matrix. The structural model adds more complexity, including an endogenous exit decision and therefore we revisit this type of mapping in this Appendix. Specifically, starting from the parametrization described in the main text, we decrease each parameter of the shock process (one-by-one) by 5 percent. Then, we solve the model under the new parametrization and compute the implied autocovariance matrix.

Figure D.5 depicts a "heat map" of absolute changes in the autocovariance matrix of log employment relative to the baseline parametrization. To ease the exposition we set the lowest 75 percent of changes to zero. Therefore, dark blue means that a particular parameter has relatively little impact on that part of the autocovariance, while bright yellow indicates that it is important. The shape of the heat map follows that of the autocovariances depicted in the main text with age on the horizontal axis and lag length on the vertical axis. For instance, point (19,19) denotes the covariance between log employment at startup (age 19 minus lag length 19) and at age 19.

There are several points to highlight. First, long-horizon autocovariances (bottom right part of the heat map) are key for identifying the dispersion of ex-ante heterogeneity, σ_θ as highlighted in the main text. On the other hand, the persistence and variance of transitory shocks (ρ_w and σ_ε) are more important for shorter-horizon autocovariances at older ages (top right part of the heat map). Second, while there is heterogeneity in the sensitivity of the autocovariance matrix to the different parameters, there is relatively little overlap across them. This is suggestive that the arguments brought forward in Section 2.4. in the statistical framework carry over to the structural model. Third, the figure suggest that all parameters of the shock

[8] The authors measure skewness as the ratio of the 90 to 50 and 50 to 10 percentiles of the growth rate distribution for all firms in their sample.

Figure D.5: Autocovariance matrices: baseline and restricted version



Notes: "Heat map" of changes to the autocovariance of log-employment induced by changes in individual parameters. Each parameter is decreased (one-by-one) by 5 percent. Colors indicate the strength of the absolute change relative to the baseline parametrization where the smallest 75 percent of changes are set to zero for clarity. The horizontal axes denote age and the vertical axes denote lag length. For instance, point (19,19) denotes the covariance between age 19 and age 19-19=0 sizes.

process could be identified off of the autocovariance matrix, even though we utilize more information stemming from targeting also the average size and exit life-cycle profiles. For instance, the level of ex-ante heterogeneity μ_θ (dispersion of initial draws $\sigma_{\widetilde{v}}$) strongly influences average firm size (startup size).

E Structural model: extensions

E.1 Imperfect Information

A key goal of this paper is to quantify the importance of ex-ante versus ex-post heterogeneity. In the model, ex-ante heterogeneity is formed by a smooth, firm-specific demand profile, which is determined by three stochastic parameters, all drawn upon entry ($\widetilde{u}_i$, $\widetilde{w}_i$, and θ_i). Ex-post heterogeneity is induced by firm-specific stochastic shock innovations which hit the firms during subsequent years ($\varepsilon_{i,t}$ and $z_{i,t}$). In our baseline model, we assume that the various stochastic draws are contemporaneously revealed to the firm. That is, upon entry the firm learns its entire ex-ante profile, whereas shocks are revealed in the later periods, once they hit the firm. In this section, we explore the importance of the timing of information, conducting the two exercises described in Section 3.4.

Exercise 1: learning by an outside observer. Our first exercise is to quantify the amount ex-ante heterogeneity from the perspective of an outside observer who cannot observe the individual state variables driving firm-level demand, but only the firm's overall demand level, $\varphi_{i,t}$.[9] To proceed, we derive Bayesian learning formulas which the agent optimally uses to learn about the underlying state variables. To this end, let us define the following variable:

$$\begin{aligned} g_{i,t} \equiv& \ln \varphi_{i,t} - \rho_w \ln \varphi_{i,t-1} \\ =&\theta_i + \varepsilon_{i,t} + z_{i,t} - \rho_w z_{i,t-1} + (\rho_v - \rho_w) v_{i,t-1} + (\rho_u - \rho_w) u_{i,t-1} \quad \text{for } t = 1, ..., T \end{aligned} \tag{E.1}$$

We can now write the likelihood of the observed data ($\Phi_T = \{\phi_1, ..., \phi_t\}$), conditional on the (beliefs about the) unobserved draws $\chi_i = (\theta_i, \widetilde{v}_i, \widetilde{u}_i)'$ as

[9] Given that in the baseline model there are no adjustment costs, we could alternatively assume that the agent instead observes the firm's employment, and obtain the same results

$$\mathcal{L}(\Phi_T|\chi_i) \propto \exp\left\{-\frac{1}{2}\frac{1}{\sigma_g^2}\sum_{t=1}^{T}(g_{i,t}-H_t'\chi_i)^2\right\},$$

where

$$H_t = \begin{pmatrix} h_{1,t} \\ h_{2,t} \\ h_{3,t} \end{pmatrix} = \begin{pmatrix} 1+(\rho_u-\rho_w)\sum_{j=1}^{t-1}\rho_u^j \\ (\rho_v-\rho_w)\rho_v^t \\ (\rho_u-\rho_w)\rho_u^t \end{pmatrix}$$

and

$$\sigma_g^2 = \sigma_\varepsilon^2 + \sigma_z^2(1-\rho_w^2)$$

The prior distribution is the unconditional distribution of the stochastic draws, i.e. a multivariate normal given by:

$$P(\chi_i) = N(M_\chi, \Sigma_\chi)$$

where the mean is $M_\chi = (\mu_\theta, \mu_{\widetilde{v}}, \mu_{\widetilde{u}})'$ and the variance is

$$\Sigma_\chi = \begin{pmatrix} \sigma_\theta^2 & 0 & 0 \\ 0 & \sigma_{\widetilde{v}}^2 & 0 \\ 0 & 0 & \sigma_{\widetilde{u}}^2 \end{pmatrix}$$

Therefore, the *posterior distribution* of the unobserved draws, conditional on the data and the prior is also a multivariate normal given by

$$P(\chi|\Phi_T) \propto \mathcal{L}(\Phi_T|\chi_i)P(\chi_i) = \exp\left\{-\frac{1}{2\sigma_g^2}\sum_{t=1}^{T}(g_{i,t}-H_t'\chi_i)^2\right\}\exp\left\{-\frac{1}{2}(\chi_i-M_\chi)\Sigma_\chi^{-1}(\chi_i-M_\chi)'\right\},$$

which can be re-written as:

$$P(\chi|\Phi_T) \propto \begin{pmatrix} \exp\left\{-\frac{1}{2}\left[\frac{1}{\sigma_g^2}\sum_{t=1}^{T}(g_{i,t}-H_t'\chi_i)^2+\frac{1}{\sigma_\theta^2}(\theta_i-\mu_\theta)^2\right]\right\} \\ \exp\left\{-\frac{1}{2}\left[\frac{1}{\sigma_g^2}\sum_{t=1}^{T}(g_{i,t}-H_t'\chi_i)^2+\frac{1}{\sigma_{\widetilde{v}}^2}(\widetilde{v}_i-\mu_{\widetilde{v}})^2\right]\right\} \\ \exp\left\{-\frac{1}{2}\left[\frac{1}{\sigma_g^2}\sum_{t=1}^{T}(g_{i,t}-H_t'\chi_i)^2+\frac{1}{\sigma_{\widetilde{u}}^2}(\widetilde{u}_i-\mu_{\widetilde{u}})^2\right]\right\} \end{pmatrix} = \begin{pmatrix} P_1(\chi|\Phi_T) \\ P_2(\chi|\Phi_T) \\ P_3(\chi|\Phi_T) \end{pmatrix}$$

Given this expression, one can derive the optimal learning formulas the uncertainty about all three ex-ante components:

$$\sigma(\theta|\Phi_T) = \frac{\sigma_g^2 \sigma_\theta^2}{\sigma_\theta^2 T \overline{h}_{1,t}^2 + \sigma_g^2} \tag{E.2}$$

$$\sigma(\widetilde{v}|\Phi_T) = \frac{\sigma_g^2 \sigma_{\widetilde{v}}^2}{\sigma_{\widetilde{v}}^2 T \overline{h}_{2,t}^2 + \sigma_g^2} \tag{E.3}$$

$$\sigma(\widetilde{u}|\Phi_T) = \frac{\sigma_g^2 \sigma_{\widetilde{u}}^2}{\sigma_{\widetilde{u}}^2 T \overline{h}_{3,t}^2 + \sigma_g^2} \tag{E.4}$$

The standard deviation of $\ln \varphi_t^{ex-ante} = u_{i,t} + v_{i,t}$, conditional on observing demand up until period t and conditional on Bayesian learning, can now be expressed as:

$$\sigma(\ln \phi_T^{ex-ante}|\Phi_T) = \rho_u^{2T} \sigma(\widetilde{u}|\Phi_T) + \sigma(\theta|\Phi_T) \left(\sum_{j=0}^{T-1} \rho_u^j \right)^2 + \rho_v^{2T} \sigma(\widetilde{v}|\Phi_T)$$

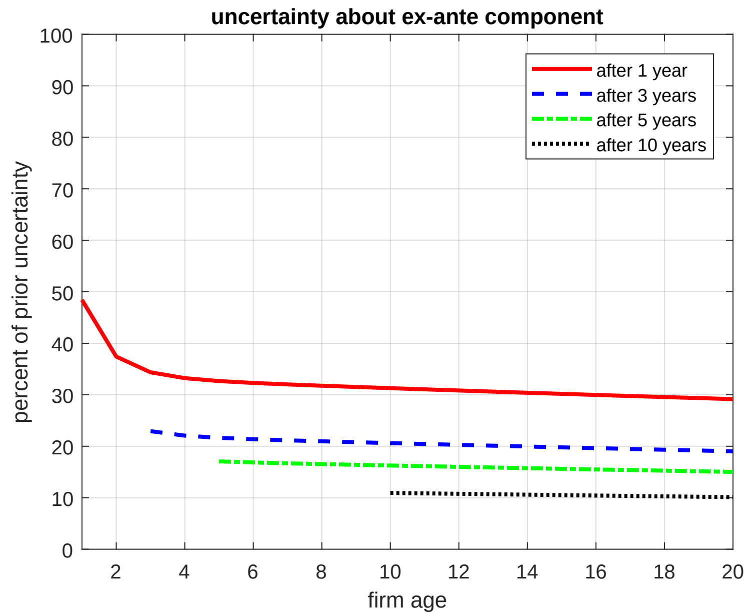
Figure E.6 shows how quickly uncertainty about the entire ex-ante growth profile gets resolved. The figure is based on the parameter estimates from the structural model. We show the uncertainty about the ex-ante growth path, in percent of prior uncertainty, for four different cases: after 1, 3, 5 and 10 years of operation.

Clearly, despite not observing the ex-ante profile directly, the outside observer nevertheless learns about it very quickly. Between 50 and 70 percent of the prior uncertainty gets resolved already after just one observation. This increases to about 85 percent after 5 years of observations. Intuitively, there is a large amount of heterogeneity in ex-ante growth profiles, which tend to dominate the dynamics during the first few years. This allows the observer to learn about the profiles very quickly.

Exercise 2: learning by firms. Next, we consider a version of the model in which firms themselves have limited information. This affects their continuation decisions, and hence all the equilibrium outcomes. We therefore re-calibrate the entire model.

Given the importance learning in the first year upon entry, we assume that firms cannot observe their individual state variables (as the outsider in Exercise 1) only at startup. Therefore, the firm uses Bayesian learning formulas similar to the ones derived above. At age 1, however, the state variables are fully revealed to the firm. This alleviates the additional computational burden brought about by additional state variables introduced because of the optimal learning formulas. We believe this is a

Figure E.6: Unresolved uncertainty about ex-ante component of growth path



reasonable assumption since Exercise 1 indicates that most learning occurs in the very first year.

We can apply the Bayesian updating formulas previously, but with one difference. Since one of the transitory shocks, ε, is persistent, firms also need to update their beliefs about this first draw. This implies a direct extension of the Bayesian updating formulas to include ε_1 as another "initial draw" to learn about.

The posterior distributions for the four components the firm needs to learn about can be written as:

$$
\begin{aligned}
P(\theta_i|\varphi_{i,1}) \sim & N(\mu(\theta_i|\varphi_{i,1}), \sigma(\theta_i|\varphi_{i,1}^2)) \\
\mu(\theta_i|\varphi_{i,1}) = & \frac{\sigma_\theta^2[\varphi_{i,1} - \rho_v\mu(\widetilde{v}_i|\varphi_{i,1}) - \rho_u\mu(\widetilde{u}_i|\varphi_{i,1}) - \mu(\varepsilon_{i,1}|\varphi_{i,1})] + \sigma_z^2\mu_\theta}{\sigma_z^2 + \sigma_\theta^2} \\
\sigma(\theta_i|\varphi_{i,1})^2 = & \frac{\sigma_z^2\sigma_\theta^2}{\sigma_z^2 + \sigma_\theta^2}
\end{aligned}
$$

$$
\begin{aligned}
P(\widetilde{u}_i|\varphi_{i,1}) \sim & N(\mu(\widetilde{u}_i|\varphi_{i,1}), \sigma(\widetilde{u}_i|\varphi_{i,1}^2)) \\
\mu(\widetilde{u}_i|\varphi_{i,1}) = & \frac{\sigma_{\widetilde{u}_i}^2[\varphi_{i,1} - \rho_v\mu(\widetilde{v}_i|\varphi_{i,1}) - \mu(\theta_i|\varphi_{i,1}) - \mu(\varepsilon_{i,1}|\varphi_{i,1})] + \sigma_z^2\mu_{\widetilde{u}}}{\sigma_z^2 + \sigma_{\widetilde{u}}^2} \\
\sigma(\widetilde{u}_i|\varphi_{i,1})^2 = & \frac{\sigma_z^2\sigma_{\widetilde{u}}^2}{\sigma_z^2 + \sigma_{\widetilde{u}}^2}
\end{aligned}
$$

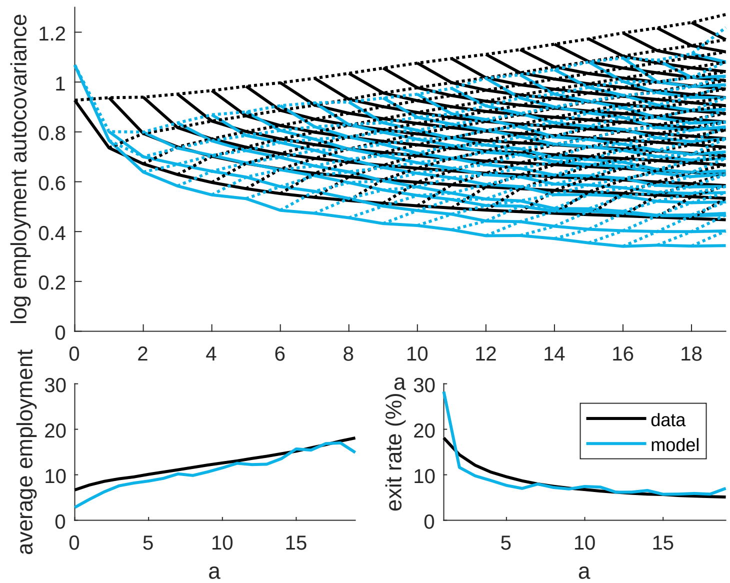
$$
\begin{aligned}
P(\widetilde{v}_i|\varphi_{i,1}) \sim & N(\mu(\widetilde{v}_i|\varphi_{i,1}), \sigma(\widetilde{v}_i|\varphi_{i,1}^2)) \\
\mu(\widetilde{v}_i|\varphi_{i,1}) = & \frac{\sigma_{\widetilde{v}}^2[\varphi_{i,1} - \rho_u\mu(\widetilde{u}_i|\varphi_{i,1}) - \mu(\theta_i|\varphi_{i,1}) - \mu(\varepsilon_{i,1}|\varphi_{i,1})] + \sigma_z^2\mu_{\widetilde{v}}}{\sigma_z^2 + \sigma_{\widetilde{v}}^2} \\
\sigma(\widetilde{v}_i|\varphi_{i,1})^2 = & \frac{\sigma_z^2\sigma_{\widetilde{v}}^2}{\sigma_z^2 + \sigma_{\widetilde{v}}^2}
\end{aligned}
$$

$$
\begin{aligned}
P(\varepsilon_i|\varphi_{i,1}) \sim & N(\mu(\varepsilon_i|\varphi_{i,1}), \sigma(\varepsilon_i|\varphi_{i,1}^2)) \\
\mu(\varepsilon_i|\varphi_{i,1}) = & \frac{\sigma_\varepsilon^2[\varphi_{i,1} - \rho_v\mu(\widetilde{v}_i|\varphi_{i,1}) - \rho_u\mu(\widetilde{u}_i|\varphi_{i,1}) - \mu(\theta_i|\varphi_{i,1})] + \sigma_z^2\mu_\varepsilon}{\sigma_z^2 + \sigma_\varepsilon^2} \\
\sigma(\varepsilon_i|\varphi_{i,1})^2 = & \frac{\sigma_z^2\sigma_\varepsilon^2}{\sigma_z^2 + \sigma_\varepsilon^2}
\end{aligned}
$$

Based on these formulas we re-solve and -calibrate the model. Figure E.7 shows the targets, including the exit profile by age. The model with learning generates more exit in year 1 than the baseline, as firms learn about their fundamentals. Overall, the fit appears slightly worse than our baseline model, supporting the full information assumption in our baseline. Table E.6 shows the parameter values in the baseline model and in the version with learning; overall these turn out to be similar.

How much does the learning assumption matter for the main results? Figure E.8

Figure E.7: Model fit: model with learning



shows the contribution of ex-ante heterogeneity in the variance decomposition, but now in the model with learning. Results are very similar to the baseline. Thus, while assumptions regarding information and learning in principle matter for our results, the quantitative impact appears very limited.

E.2 Targeting the Firm Size Distribution

When parametrizing the baseline model, we do not directly target the size distribution. Figure 6 in the main text shows that, overall, the model nonetheless provides a reasonable fit of the firm size distribution. An exception is that the model somewhat understates the importance of the right tail of the size distribution for old businesses.

In this appendix, we re-calibrate the model targeting the firm size distribution as well, in addition to the autocovariance matrix and the size and exit profiles. As shown by Figure E.9, this version provides a better fit of the size distribution. Figure E.10 shows that the model still provides a reasonable fit to the remaining targets, although the fit is somewhat inferior to the baseline. This suggests that there is a non-trivial trade-off in matching the size distribution and the remaining targets.

Figure E.11 shows the variance decomposition for the alternative calibration, as

Table E.6: Parameter values (model with learning)

	parameter	baseline	learning
		set a priori	
β	discount factor	0.96	0.96
η	elasticity of substitution	6.00	6.00
f^e	entry cost	0.44	0.44
		used to target moments	
f	fixed cost of operation	0.539	0.542
δ	exogenous exit rate	0.041	0.035
μ_θ	permanent component θ, mean	-1.762	-1.882
σ_θ	permanent component θ, st. dev.	1.304	1.319
$\sigma_{\widetilde{u}}$	initial condition u_{-1}, st. dev.	1.572	1.601
$\sigma_{\widetilde{v}}$	initial condition v_{-1}, st. dev.	1.208	1.218
σ_ε	transitory shock ε, st. dev.	0.307	0.314
σ_z	noise shock z, st. dev.	0.203	0.184
ρ_u	permanent component, persistence	0.393	0.394
ρ_v	transitory component, persistence	0.988	0.981

Notes: parameter values. Top three parameters are calibrated as discussed in the main text. The remaining parameters are set such that the model matches the empirical autocovariance of employment and the age profiles of average size and exit rates from age 0 to 19.

Figure E.8: Variance decomposition: model with learning

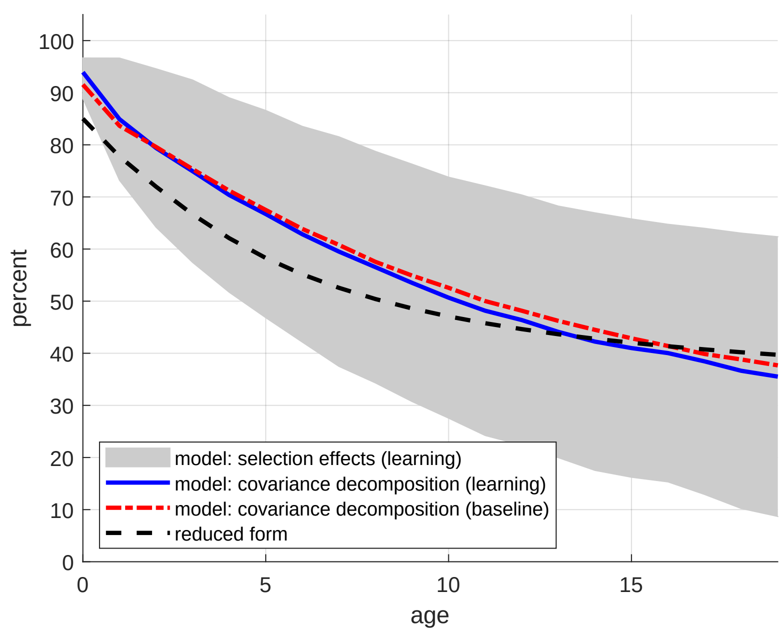
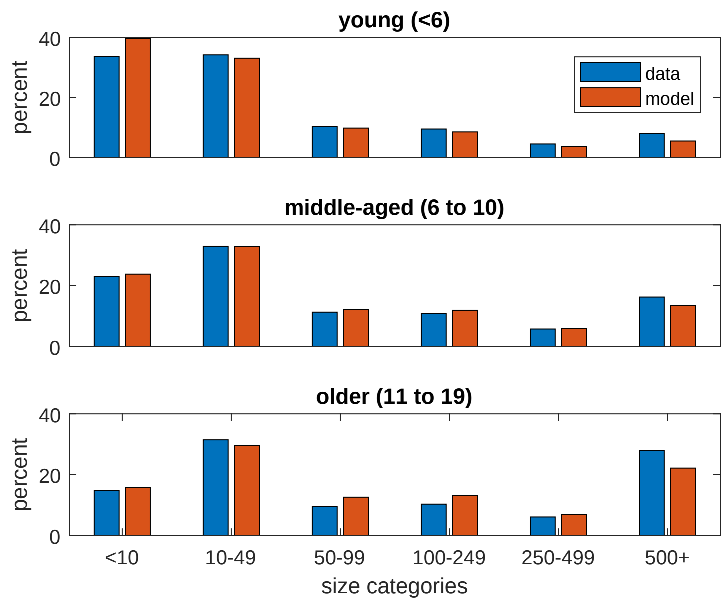


Figure E.9: Size distribution: alternative calibration



Notes: Employment shares by firm age and size (employment). Values are expressed as percentages of total employment in firms between 0 to 19 year old firms, both in the data and the model. Data are obtained from the Business Dynamics Statistics, an aggregated and publicly available version of the LBD over the corresponding time period.

Figure E.10: Remaining targets: alternative calibration

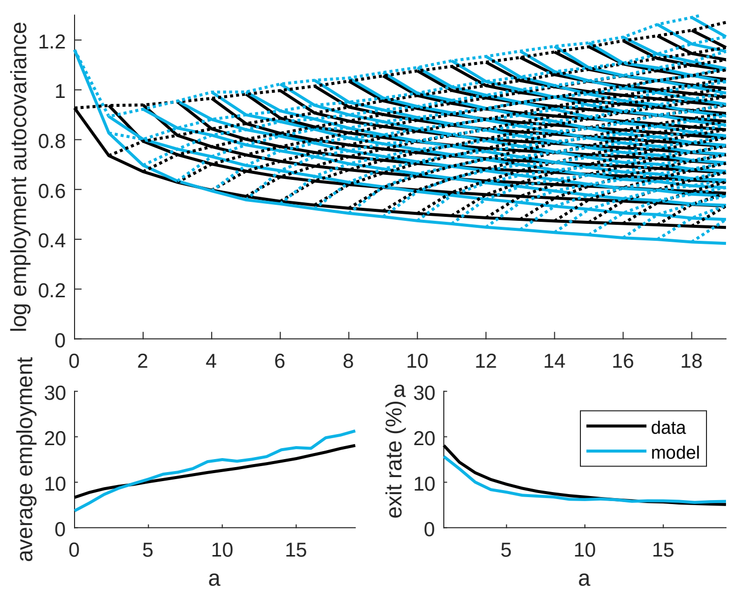
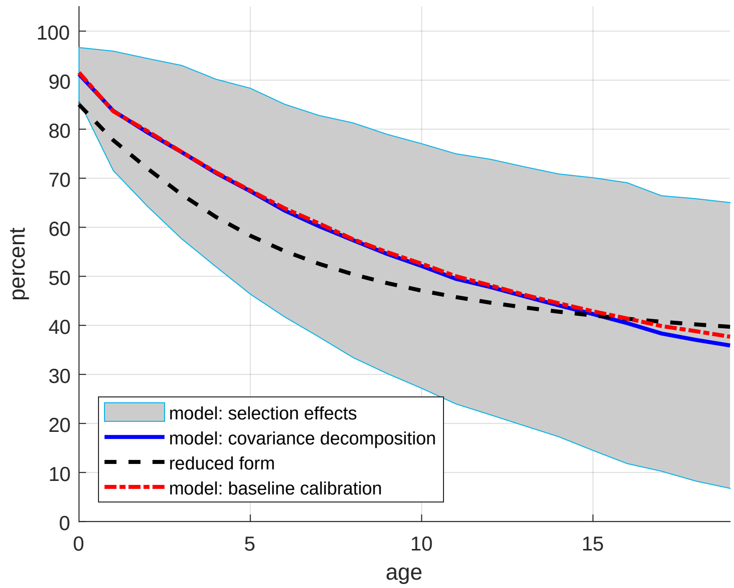


Figure E.11: Variance decomposition: alternative calibration



well as the baseline. Importantly for our main message, the figure shows that the decomposition when also targeting the size distribution is almost identical to the baseline.

E.3 Flexible labor supply

The baseline model in the main text assumes fixed labor supply. In this appendix we relax that assumption. Specifically, we now assume that preferences include not only consumption (as in our baseline), but also labor supply (employment):

$$U_0 = \sum_{t=0}^{\infty} \beta^t \left[C_t - \upsilon \frac{N_t^{1+\psi}}{1+\psi} \right],$$

where υ scales the disutility of labor and ψ is the inverse of the Frisch elasticity of labor supply. This setup delivers the familiar optimal labor supply decision of households which equates the real wage (recall that preferences are linear in consumption) and the marginal disutility of labor:

$$\frac{W_t}{P_t} = \upsilon N_t^{\psi}.$$

In the version with flexible labor supply, this condition replaces our baseline assump-

Table E.7: Late vs Early sample aggregate changes (in %)

	fixed LS	flexible LS
output	−4.5	−6.6
employment	0	−4.4
real wage	−4.0	−4.4

tion that $N = 1$. Moreover, we calibrate the flexible labor supply version by setting υ such that $N = 1$ (This is a standard calibration strategy). Since no other part of the model is affected by the introduction of flexible labour supply, it follows that the aggregate outcomes and distributions in the two models coincide precisely. Therefore, all the main results in Section 3 remain unchanged, since they are obtained from decompositions or partial-equilibrium exercises.

Some of the other results may be affected, however. We now explore how flexible labor supply affects the results of the split sample analysis of Section 4.[10] In particular, using the flexible labor supply assumption above we re-estimate the model on the early and late samples of the data. In doing so, we set υ such that $N = 1$ in the early sample and we assume $\psi = 1$ (and thus also a Frisch elasticity of 1), which is in the middle of the range of estimates in the literature.

As mentioned earlier, because we are targeting average size life-cycle profiles (together with exit rates and the autocovariance matrix), only aggregate results are affected. In particular, Table E.7 shows the changes in aggregate values, going from the early to the late sample (in percent of early sample values), for our baseline (fixed labor supply) and the alternative (flexible labor supply). The results show that with flexible labor supply, the estimated decline in aggregate output is stronger. This is because the real wage falls going from early to the late sample which, in turn, discourages households from supplying labor.

[10]For the micro frictions exercises in Section 4, the issue is less pressing, since these are illustrative examples in which we compare two versions of the model (the baseline and one with a restricted shock process), but both under the same assumption on labor supply.

E.4 Treatment of "gazelles"

The main text defines gazelles as those startups with an ex-ante projected growth rate of at least 20 percent annually, over the first five years, and an expected employment level of at least 10 workers at some point during their lifetimes. Under this definition, we find 5.4% of startups are gazelles. In this appendix, we analyze different definitions of gazelles and to what extent gazelles defined on ex-ante projected growth profiles actually achieve their potential.

E.4.1 Different definitions of gazelles

There are other ways of defining gazelles which can be found in the literature. For instance, Guzman and Stern (forthcoming) define "high-growth" firms as those which achieve a "growth outcome". The specific growth outcome in their paper is an initial public offering (IPO) or an acquisition "at a meaningful positive valuation within 6 years of registration". This particular definition results in high-growth firms accounting for 0.1% of all businesses in their sample.

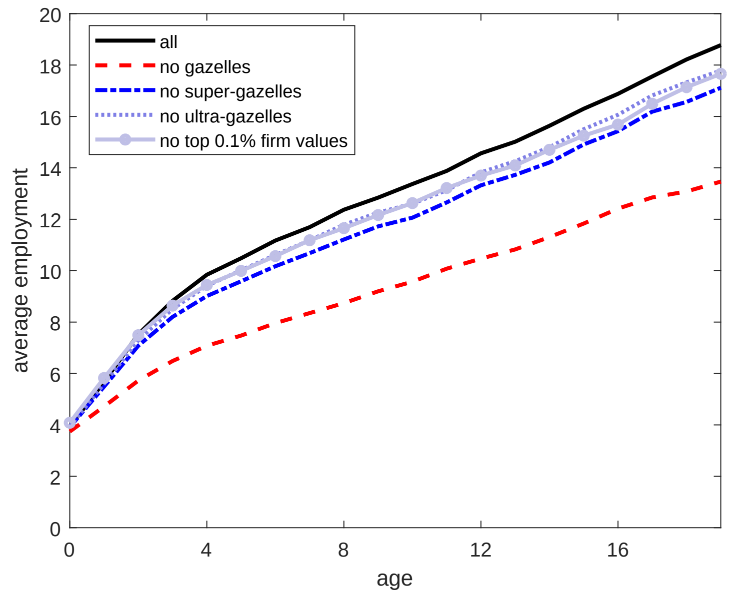
Despite the fact that the two approaches are not directly comparable, we investigate whether more narrowly defined gazelles in our framework still retain the power to impact aggregate outcomes. To do so, in addition to the gazelles defined in our main text, we define three other sets of gazelles: i) super-gazelles, ii) ultra-gazelles and iii) value-gazelles. Super-gazelles and ultra-gazelles are defined in the same way as our baseline gazelles using employment growth rates. However, the average growth rates are set such that the resulting group of firms accounts for 1% and 0.5% of startups in the case of super- and ultra-gazelles, respectively.[11] Value-gazelles are defined in a way that attempts to mimic the definition of Guzman and Stern. Specifically, these are the firms which fall into the top 0.1% of firm values at the age of five.

Figure E.12 replicates our results of Figure 9 in the main text, but adds super-, ultra- and value-gazelles. Interestingly, ultra- and value-gazelles have very similar growth profiles, despite not being the same firms (the group of ultra-gazelles is six time larger with 85% having firm values below the lowest firm value in the group of value-firms). The figure also highlights that even gazelles defined more restrictively can still have an aggregate impact. Specifically, value-gazelles, which are close to those defined in Guzman and Stern and account for only 0.1 percent of all startups

[11] The associated average growth rates are 41% and 47%, respectively.

still have a noticeable impact on average firm size.

Figure E.12: The importance of high-potential startups



Note: Average size by age, unbalanced panel, in the baseline and three counterfactuals.

E.4.2 Ex-ante vs ex-post gazelles

While we define gazelles based on their ex-ante potential, the rest of the literature does so based on ex-post realizations. In what follows, we compare the two approaches. Specifically, we classify firms as (regular-, super- or ultra-) gazelles using the definitions described above both based on their ex-ante predicted growth profiles, but also based on their ex-post realized growth profiles. We can then identify to what extent ex-ante identified gazelles live up to their potential and also realize their high growth and, conversely, how often do "regular" firms defy ex-ante expectations and realize high rates of growth ex-post. Table E.8 below summarizes the results.

The results show that about one third of ex-ante identified gazelles fail to realize their potential and grow at rates below 20 percent annually in their first five years. This failure rate increases as the definition of gazelles becomes more stringent. Conversely, there is a small fraction of firms which are not identified as gazelles ex-ante, but which nevertheless manage to grow at very high rates. For the case of gazelles, this share is almost 13 percent, but it drops to only about 3.5 percent in the case

Table E.8: Living up to potential and defying odds

	living up to potential	defying ex-ante odds
gazelles	67.9%	12.7%
super-gazelles	62.0%	4.7%
ultra-gazelles	58.5%	3.4%

Note: "Living up to potential" provides the share of ex-ante identified gazelles which manage to realize their potential and grow at the pre-defined rate also ex-post. "Defying ex-ante odds" provides the share of firms which are not classified as gazelles ex-ante (and are also overcome the lower limit of 10 workers at some point over their life-cycle) which end up growing at the pre-defined rates ex-post. The pre-defined growth rates for gazelles, super- and ultra-gazelles are, 20, 41 and 47 percent, respectively.

of ultra-gazelles. Therefore, Table E.8 shows that ex-ante characteristics identified through our framework are very powerful in predicting actual growth outcomes.

By comparison Guzman and Stern (2015) propose a measure of "entrepreneurial quality" using an array of observable characteristics at the time of startup (e.g. location, whether a business holds a patent, the name of the firm etc). They report that "77% of all growth firms are in the top 5% of our estimated growth probability distribution" and that "the average firm within the top 1% of estimated entrepreneurial quality has only a 14% chance of realizing a growth outcome". While a direct comparison is difficult as we adopt a different methodology, our framework implies that 87% of all high-growth firms were identified as high-growth ex-ante and that 68% of ex-ante identified high-growth firms realize their growth potential.

F Structural model: the importance of ex-ante heterogeneity

F.1 Ex-ante heterogeneity and the macroeconomic impact of micro-level frictions

This section provides further details to the analysis of adjustment cost in Section 4. It also shows results for other frictions: a change in the entry cost and in the fixed cost of operation.

F.1.1 Adjustment costs: additional details

The main text considered a model extended to include an adjustment cost to demand accumulation. All the parameter values for the baseline and the restricted version of the model are the same as in Table D.1. In addition, we introduce adjustment costs s.t. average costs paid by adjusting firms are 1 percent of their output. For the baseline and the restricted version of the model, this amounts to $\kappa_{\text{baseline}} = 0.045$ and $\kappa_{\text{restricted}} = 0.037$, respectively.

While the main text considers the baseline parametrization and then introduces adjustment costs, we have also investigated the impact of the presence of adjustment costs on the estimates of our structural parameters (not shown). In other words, using the same parametrization strategy as in the main text we have re-estimated the structural model in the case of positive adjustment costs, amounting to 1 percent of output of adjusting firms.

Intuitively, the dispersion of the permanent component of the demand fundamental, σ_θ is somewhat narrower than in the benchmark model. This is because part of the cross-sectional dispersion in firm sizes at old ages is now also driven by adjustment costs and not only firm types. Nevertheless, the variance decomposition of the cross-sectional variation in firm size into the contributions of ex-ante and ex-post factors is very similar to the baseline. Therefore, while adjustment costs introduce an additional margin of adjustment, they do not alter the main qualitative or quantitative conclusions regarding the relative importance of ex-ante heterogeneity and ex-post shocks.

F.1.2 Financial friction: additional details and results

The main text considered a model extended to include a financial friction. Here, we provide additional details and results for this version of the model. In the model with a financial friction, we assume that firms can hold a risk free asset, denoted b_i, which can also be held by the households. Market clearing therefore implies that the asset pays a net real interest rate $r = \frac{1}{\beta} - 1$ in the steady state. However, the firm is subject to a borrowing limit:

$$b_i \geq \zeta$$

where $\zeta \leq 0$. If at any point in time the firm cannot meet this limit, it is forced to exit. Upon entry, a firm receives an initial equity injection to cover the entry cost, but no additional equity injections are possible.

When a firm exits with positive assets ($b_i \geq 0$), then these assets are returned to the owners. If a firm exits with debt ($b_i < 0$) then the owners must settle the remaining debt. In this setting, the shadow value of wealth within a firm always exceeds the interest rate $r = \frac{1}{\beta} - 1$. That is, the firm does not pay dividends until it exits.

More formally, a firm is forced to exit if end-of-period asset holdings drop below ζ, i.e. if

$$b' = (1 + r)\, b + \pi - d < \zeta$$

where $d \geq 0$ is a dividend paid out to the firm owners. To simplify notation, we omit the firm index i. The timing within a period is as follows:

1. The firm decides whether or not to exit voluntarily or pay the fixed cost and continue. If the firm exits, it pays a final dividend $d = (1 + r)\, b$ to the owners of the firms (the households).[12]
2. The demand fundamental $\mathbf{s}'$ is realized
3. The firm decides on a plan for n, d, π and b' (in the current period)
4. The financial constraint is enforced:
   (a) The firm is forced to exit if the plan does not satisfy $b' \geq \zeta$. In that case, the firm owners must settle a remaining debt and pay $-b + f$.
   (b) If the firm is allowed to continue, the firm implements its plan and continues on to the next period.

Given that $r = \frac{1}{\beta} - 1$, one can verify that (i) a continuing firm always sets $d = 0$, since the shadow value of firms inside the firm exceeds the value outside the firm, (ii) a firm chooses n to statically maximize $\pi(\mathbf{s})$ in precisely the same way as in the model without financial frictions. Intuitively, maximizing profits increases the wealth of the firm, which in turn makes it less likely that the constraint will bind.

[12]Note that this dividend may be negative. In that case the households settle the remaining debt of the firm, i.e. pay $-(1 + r)\, b$.

It follows that, given b, there is a threshold level on the demand level such that the firm ends up being forced to exit (due to the financial constraint) if and only if:

$$\varphi(\mathbf{s}') < \overline{\varphi}(b).$$

In what follows, we let $I(\varphi(\mathbf{s}'), b)$ be an indicator function which equals one if the firm satisfies the financial constraint and zero otherwise.

Given these preliminaries, now consider the stage-1 value of a firm *excluding its financial wealth*, $V(\mathbf{s}, b)$.[13] This value is given by:

$$V(\mathbf{s}, b) = \max\left\{\mathbb{E}\left[\pi(\mathbf{s}') + \beta(1-\delta)\int V(\mathbf{s}, b')I(\varphi(\mathbf{s}'), b)F(d\mathbf{s}'|\mathbf{s}))\right], 0\right\}$$

where

$$\pi(\mathbf{s}') = \varphi(\mathbf{s}')\left(\frac{\eta}{\eta-1}\right)^{-\eta} w^{-\eta}C$$
$$b' = (1+r)\,b + \pi(\mathbf{s}')$$

Above, the expression for $\pi(\mathbf{s}')$ statically maximizes profits, precisely as in the baseline. The constraint below is the evolution of firm wealth, using that the firm does not pay dividends until it exits.

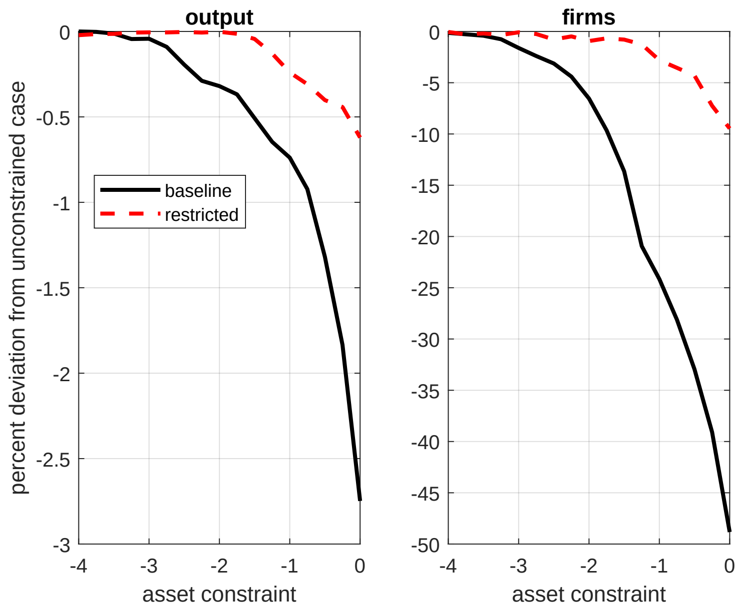
Financial frictions: additional results The main text reported results for the case when $\zeta = 0$. Figure F.13 below shows the long-run effects on output and the number of firms of a range of different values for ζ. The impact is expressed in terms of deviations from the respective frictionless version of the model.

A number of results stand out. First, for a range of values of ϕ, the introduction of the financial friction only very small effects. Second, for a given value of ζ the losses created by the financial frictions tend to be larger in the baseline model than in the version with an AR(1) process. At the maximally tight constraint ($\zeta = 0$) of this exercise, the output loss is almost three times as large as in the restricted version.

A key reason for this is that in the frictionless version of the baseline model, there are firms which choose not to exit despite making losses for extended periods of time.

[13] The firm value including financial wealth would be given by $\widetilde{V}(\mathbf{s}, b) = V(\mathbf{s}, b) + (1+r)b$. Because of properties (i) and (ii) discussed above, any solution which maximizes $V(\mathbf{s}, b)$ also maximizes $\widetilde{V}(\mathbf{s}, b)$.

Figure F.13: Effect of the financial friction



Note: the long-run effects of the introduction of the financial on aggregate output and the number of firms, relative to the model without financial frictions.

They do so because of high, long-run, potential. With the financial friction in place, however, such firms are no longer allowed to survive. The restricted version of the model features much fewer of such firms, because in that economy all firms are "the same in the long run". Indeed, Figure F.13 shows that in the baseline, the effect on the number of firms is much larger than in the restricted version.

F.1.3 Costs of entry and of operation

In this subsection, we investigate the aggregate consequences of two other micro-level frictions: a change in the entry cost and in the fixed cost of operation. As with the firm-level adjustment costs discussed in the main text, our baseline economy behaves in a quantitatively different manner in response to these two frictions compared to the restricted version of the model.

The common source of these differences is the *distribution of firm values* which differs between the two economies. Importantly, this happens despite the two economies having essentially an identical firm size distribution. In contrast to firm sizes, which often are a calibration target but which do not matter for firm decisions per se, firm values are crucial for forward-looking firm-level decisions such as entry and exit. In turn, the distribution of (expected) firm values then crucially impacts the equilibrium wage (via the free entry condition) and through it the rest of the aggregate economy.

F.1.4 Aggregate impact of micro-level frictions

Before providing a quantitative evaluation of the sensitivity of the baseline economy and the restricted version of the model to the two micro-frictions, let us first discuss the importance of the distribution of firms for these results. As mentioned above, changes in average (expected) firm values influence the equilibrium wage through the free entry condition. In turn this impacts average firm sizes and given the assumption of fixed labor supply also the number of firms. Therefore, the elasticity with which (average) firm values respond to changes in a given parameter (or variable) x, $\epsilon_{\overline{V},x} = \frac{\partial \overline{V}}{\partial x}\frac{x}{\overline{V}}$, is key for our results. This elasticity can be decomposed as follows

$$\epsilon_{\overline{V},x} = \overline{\epsilon_{V_i,x}} + \text{Cov}\left[\epsilon_{V_i,x}, V_i\right]. \tag{F.1}$$

where $\overline{\epsilon_{V_i,x}}$ is the average sensitivity of firm values with respect to x and $\text{Cov}\left[\epsilon_{V_i,x}, V_i\right]$ is the extent to which these individual elasticities covary with firm values.

Importantly, both terms in (F.1) will in general depend on the *distribution of firm values*. While we have documented that the restricted version of the model and our baseline have very similar distributions of firm sizes, this does not imply the same for the distribution of firm values. Indeed, the baseline economy has much more dispersed distribution owing to the presence of (permanent) ex-ante heterogeneity.

In other words, even though the firm size distribution is often a calibration target, it does not matter for firm decisions per se. In contrast, the distribution of (unobserved) firm values, which crucially depend on the nature of the underlying firm-level driving forces, is what influences forward-looking firm decisions (such as exit and entry) and in turn aggregate equilibrium outcomes.

Table F.9 documents the quantitative impact of the two considered changes in micro-frictions. In both cases, the behavior of the restricted version of the model is considerably different from that of the baseline economy. Interestingly, the two micro-frictions have qualitatively different implications for the two economies. While the baseline economy is less sensitive to an entry cost increase, it is somewhat more sensitive to a change in fixed costs of operation. The macro effects of these changes are non-trivially determined by the interaction between the distribution of firm values

Table F.9: Aggregate impact of entry and operation costs (percent change)

	output	wage	size	exit	firms
			Entry cost		
baseline economy	−5.6	−6.0	+11.5	−2.7	−10.0
restricted economy	−11.4	−6.3	+15.2	−1.4	−13.0
			Operation cost		
baseline economy	−2.1	−2.1	+50.8	+4.1	−33.7
restricted economy	−1.9	−1.6	+30.6	+0.9	−23.6

Notes: Long-run impact of increasing entry costs (top panels) and fixed costs of operation (bottom panels) in the baseline economy and in the restricted version where $\rho_u = \sigma_\theta = \sigma_{\tilde{u}} = 0$ and $\rho_v = \rho_w$. In both economies costs are increased by 50 percent. Reported values are relative to the respective values prior to the increase. Output refers to aggregate production, wage is the real wage rate, size is average firm size, exit is the average exit rate and firms refers to the number of incumbent firms.

and the general equilibrium effects which arise from firms' responses.

F.2 Changes in the Nature of Firm Growth: baseline versus restricted model

In this Appendix, we revisit the split-sample analysis of Section 4 in the restricted model, which does not properly account for ex-ante heterogeneity. As is highlighted in the main text for the case of the macroeconomic impact of micro-level distortions, also in the case of the split-sample analysis the baseline and the restricted versions of the model deliver starkly different conclusions. In particular, while aggregate output falls from the early to the late period in the baseline model, it *increases* in the restricted model.

The reason for this stark qualitative difference lies again in the sensitivity of the two economies to changes in structural parameters. However, this time we are not targeting a particular change in structural parameters (as we do e.g. in Section 4), but instead we effectively target a decline in average firm size. Recall that the life-cycle profile of average firm size is a target in the estimation. Matching the observed flattening of the life-cycle size profile results in a roughly 11% decrease in average firm size in both the baseline and the restricted model.

However, this common decline in average firm size is associated with different changes in the underlying distribution of the demand fundamental across the two models. The baseline model is characterized by a considerably larger drop in average demand fundamentals (about 20% compared to about 5% in the baseline model). In addition, this average size decline is accompanied by an increase in the number of firms (about 10% in both models).

The two effects above, a change in average demand and the number of firms, jointly determine the response of the real wage. This can be seen from the following (given that the nominal wage is normalized to 1 and that prices are set as a constant markup over marginal costs)

$$w = \quad P^{-1} = \left[\int_\Omega \varphi_i p_i^{1-\eta} di\right]^{\frac{\eta}{\eta-1}} = \left[\int_\Omega \varphi_i \left(\frac{\eta}{\eta-1}\right)^{1-\eta} di\right]^{\frac{\eta}{\eta-1}} = \left(\frac{\eta}{\eta-1}\right)^{-\eta} (M\overline{\varphi})^{\frac{\eta}{\eta-1}},$$

where M is the mass of firms and $\overline{\varphi} = 1/M \int_\Omega \varphi di$ is average demand. Therefore, the response of the wage depends on the relative extent of the changes in $\overline{\varphi}$ and M. In the case of the baseline model, the drop in $\overline{\varphi}$ outweighs the increase in the number of firms and therefore the wage declines. The opposite happens in the restricted model.

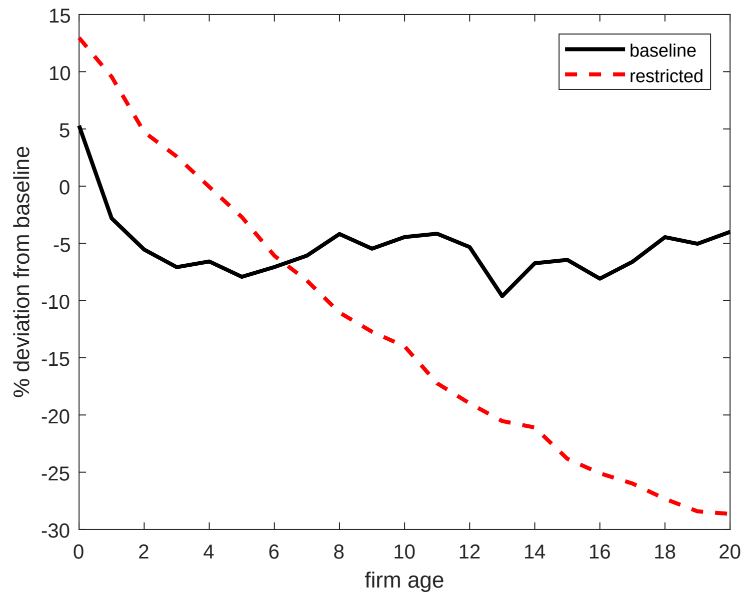
Finally, the change in the wage is directly related to aggregate output because labor supply is fixed (and therefore labor income changes one-for-one with the wage) and because all costs are expressed in labor units (and therefore consumption is equal to output).

To understand which parameters are behind the above results, let us look at the changes estimated across the two subsamples and the two model versions. In the baseline model, 7 parameters change noticeably when going from the early to the late sample. We will group these into three sets of changes:

1. c_f declines and δ rises,
2. σ_ϵ decreases, while $\sigma_{\tilde{u}}$ and σ_z increase,
3. μ_θ and σ_θ decrease.

Changes in both (1) and (2) reduce average firm size (as in the data) but *increase* aggregate output. To understand this, note that changes in (1) result in a shift away from endogenous to exogenous firm exit. This enables firms with relatively low demand to survive, reducing average firm size. At the same time, however, firms with (temporarily) low demand are able to survive and grow in the future. The latter also

Figure F.14: Response of average size profile to a 10% decrease in μ_θ



incentivizes firm entry. Similarly, changes in (2) reduce the risk of firm exit for older firms (once initial conditions have died out) and as such these effects have similar implications as changes (1).

Therefore, the key changes in understanding the decline in aggregate output are the decline in μ_θ and σ_θ. Both lower average firm size and output as they reduce the growth potential of startups.

Interestingly, the reduced form model displays very similar changes in structural parameters as the baseline. The notable difference are the changes in (3). First, $\sigma_\theta = 0$ in the restricted model and therefore a change in μ_θ impacts *all* firms in the same way. As a result, average firm size in the restricted model is very sensitive to changes in μ_θ.

We illustrate this in Figure F.14 which shows the percentage change in the average size life-cycle profile in response to a 10% decrease in μ_θ in both the baseline and the restricted model. While in the baseline model is characterized by a roughly 5% decrease in average firm size over the life-cycle. This relatively moderate decline is driven by the fact that much of the life-cycle profile in the baseline model is determined by a relatively small number of high-growth firms. In other words, the life-cycle profile is largely determined by σ_θ, rather than μ_θ.

On the other hand, the restricted model displays a much stronger decline which increases with firm age. This is because *all* firms are affected by the change in μ_θ and this effect enters multiplicatively (and in exponents) into firm size. This is also why the estimated change in μ_θ is very small across the two subsamples in the restricted model.

G Results for establishments

Below we report results for establishments. We find very similar results as for firms, with a somewhat higher degree of ex-ante heterogeneity among establishments.

G.1 Autocovariances for establishments

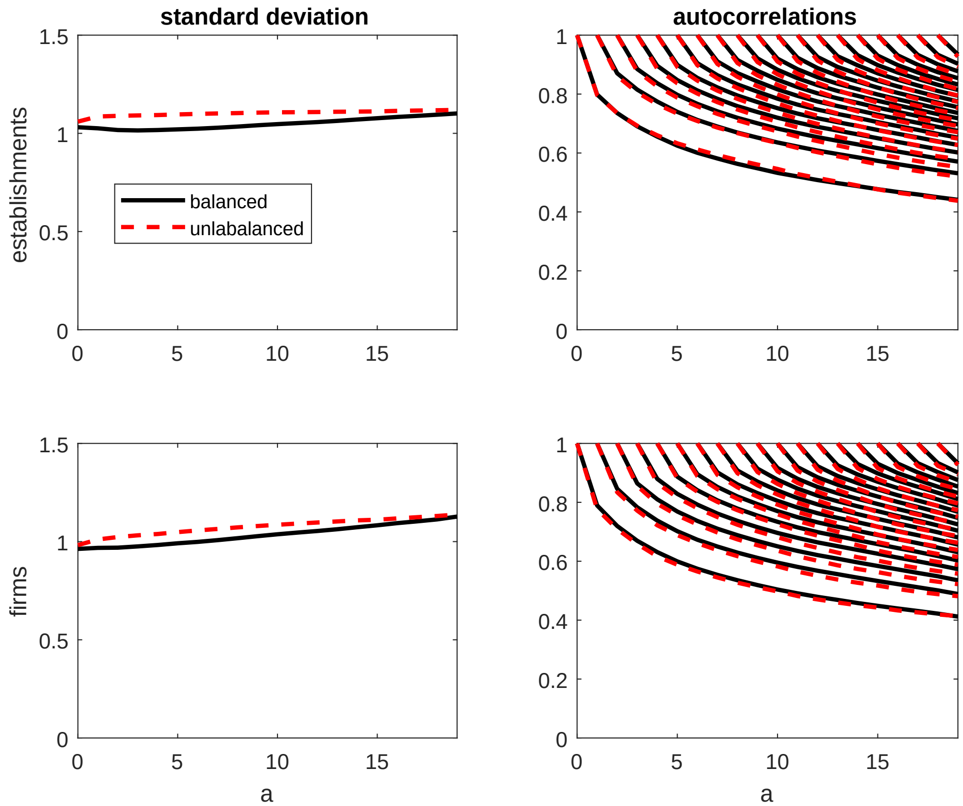
Table G.10: Autocovariance matrix of establishments for unbalanced panel

	Age $a \geq 0$																			
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
0	1.123																			
1	0.916	1.179																		
2	0.849	1.018	1.186																	
3	0.801	0.955	1.043	1.191																
4	0.765	0.908	0.984	1.057	1.195															
5	0.735	0.874	0.941	1.001	1.069	1.202														
6	0.713	0.846	0.908	0.962	1.016	1.082	1.208													
7	0.692	0.821	0.879	0.929	0.977	1.030	1.092	1.213												
8	0.674	0.800	0.854	0.900	0.945	0.992	1.041	1.100	1.217											
9	0.657	0.780	0.832	0.875	0.917	0.961	1.004	1.051	1.108	1.222										
10	0.641	0.762	0.812	0.853	0.892	0.933	0.972	1.014	1.059	1.114	1.226									
11	0.620	0.742	0.791	0.830	0.868	0.905	0.943	0.981	1.021	1.064	1.118	1.227								
12	0.605	0.726	0.773	0.811	0.847	0.883	0.917	0.953	0.990	1.027	1.070	1.122	1.229							
13	0.591	0.710	0.756	0.794	0.828	0.862	0.896	0.929	0.963	0.998	1.035	1.077	1.127	1.232						
14	0.576	0.694	0.739	0.776	0.810	0.843	0.875	0.907	0.938	0.971	1.006	1.042	1.082	1.132	1.234					
15	0.561	0.678	0.723	0.760	0.793	0.826	0.856	0.887	0.917	0.947	0.980	1.013	1.048	1.087	1.136	1.236				
16	0.549	0.665	0.709	0.745	0.778	0.811	0.841	0.870	0.899	0.928	0.958	0.989	1.022	1.057	1.094	1.142	1.242			
17	0.540	0.653	0.696	0.731	0.764	0.796	0.826	0.855	0.883	0.911	0.939	0.968	0.998	1.030	1.064	1.101	1.149	1.246		
18	0.529	0.642	0.684	0.718	0.751	0.782	0.811	0.840	0.868	0.894	0.922	0.949	0.977	1.007	1.038	1.071	1.108	1.154	1.249	
19	0.519	0.632	0.674	0.708	0.739	0.770	0.798	0.827	0.854	0.880	0.907	0.933	0.960	0.988	1.016	1.048	1.080	1.115	1.160	1.253

Table G.11: Autocovariance matrix of establishments for balanced panel

	Age $a \geq 0$																			
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
0	1.063																			
1	0.844	1.052																		
2	0.771	0.908	1.034																	
3	0.723	0.849	0.913	1.030																
4	0.687	0.808	0.865	0.925	1.034															
5	0.657	0.774	0.827	0.878	0.933	1.041														
6	0.634	0.748	0.798	0.844	0.890	0.945	1.048													
7	0.617	0.728	0.776	0.818	0.860	0.907	0.959	1.058												
8	0.601	0.710	0.757	0.797	0.836	0.879	0.923	0.974	1.070											
9	0.588	0.697	0.742	0.781	0.818	0.859	0.898	0.942	0.992	1.084										
10	0.575	0.683	0.727	0.765	0.801	0.838	0.875	0.914	0.956	1.003	1.096									
11	0.564	0.671	0.715	0.751	0.785	0.820	0.855	0.891	0.929	0.969	1.017	1.107								
12	0.554	0.660	0.703	0.739	0.772	0.806	0.839	0.873	0.908	0.943	0.983	1.030	1.118							
13	0.546	0.651	0.694	0.730	0.762	0.795	0.827	0.859	0.892	0.925	0.960	0.998	1.044	1.131						
14	0.538	0.643	0.685	0.721	0.753	0.786	0.816	0.848	0.879	0.910	0.944	0.978	1.016	1.061	1.146					
15	0.530	0.633	0.675	0.711	0.743	0.774	0.805	0.835	0.866	0.896	0.927	0.958	0.991	1.029	1.073	1.160				
16	0.522	0.625	0.667	0.701	0.733	0.764	0.794	0.824	0.854	0.883	0.912	0.942	0.972	1.006	1.042	1.087	1.174			
17	0.516	0.616	0.657	0.691	0.723	0.754	0.783	0.813	0.841	0.869	0.898	0.926	0.955	0.986	1.018	1.055	1.101	1.186		
18	0.508	0.608	0.648	0.681	0.712	0.743	0.772	0.801	0.829	0.856	0.884	0.912	0.938	0.968	0.998	1.031	1.069	1.113	1.199	
19	0.501	0.600	0.639	0.673	0.703	0.733	0.761	0.790	0.818	0.844	0.871	0.898	0.925	0.953	0.982	1.013	1.047	1.083	1.127	1.213

Figure G.15: Standard deviations and autocorrelations of log employment by age



Note: The left panels show cross-sectional standard deviations of log employment by age (a) for establishments (top left panel) and firms (bottom left panel). The right panels show cross-sectional correlations of log employment between ages a and age $h \leq a$ for establishments (top right panel) and firms (bottom right panel). "Balanced" refers to a panel of establishments (firms) which survived at least up to age 19, while "unbalanced" refers to a panel of all establishments (firms).

G.2 Statistical model: results for establishments

The main text reported only results for firms. Here we report also results for establishments. Figure G.15 presents our main piece of empirical evidence: the cross-sectional autocovariance structure of logged employment, conditional on age (a). The figure presents this information for both establishments (top panels), and for firms (bottom panels), as well as for a balanced panel, containing businesses surviving at least up to age 19, and an unbalanced panel, including all businesses in our data set. Clearly, differences in autocovariances between the balanced and unbalanced panels originate primarily from different cross-section dispersion by age, while the autocorrelations are remarkably similar across the two panels.

The corresponding parameter estimates are shown in Table G.12. For comparison,

Table G.12: Parameter estimates from reduced-form model

	ρ_u	ρ_v	ρ_w	σ_θ	$\sigma_{\widetilde{u}}$	$\sigma_{\widetilde{v}}$	σ_ε	σ_z
Estabs	0.206	0.842	0.949	0.603	2.046	0.738	0.255	0.262
	(0.002)	(0.001)	(0.001)	(0.001)	(0.017)	(0.002)	(0.001)	(0.001)
Firms	0.218	0.832	0.963	0.555	1.743	0.695	0.255	0.272
	(0.002)	(0.001)	(0.001)	(0.002)	(0.015)	(0.002)	(0.001)	(0.001)

Note: Equally-weighted minimum distance estimates of Equation (2) for both establishments and firms using the balanced panel. See Appendix Table B.2 panel A for estimates using unbalanced panel.

the table also shows the estimates for firms. Figure G.16 shows that the model fit is very good for both establishments and firms.

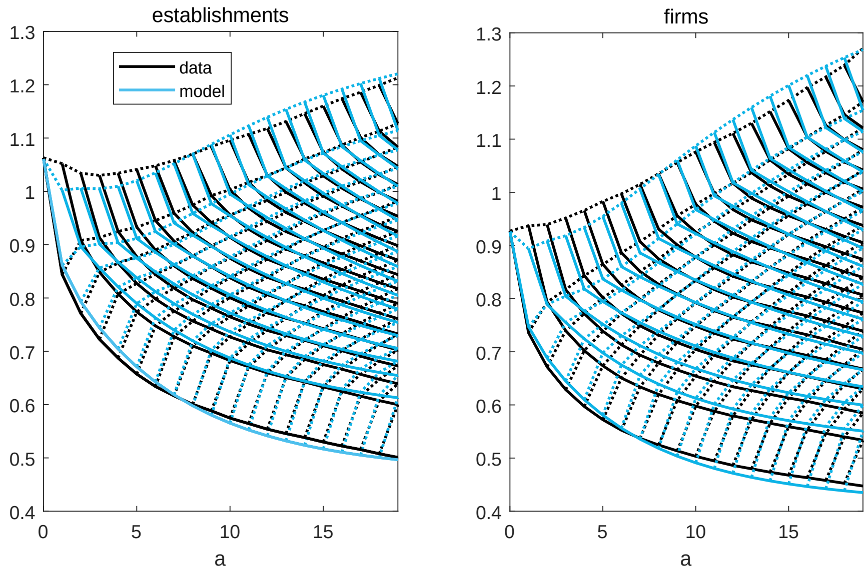
Finally, Figure 4 plots the fraction of the total variance that is accounted for by the ex-ante component. Thick lines denote the age groups used in the estimation, i.e. age zero to nineteen, whereas thin lines represent an extrapolation for businesses at age 20 or above using the point estimates. The figure shows that for businesses in the year of startup (age zero) the ex-ante component accounts for about 85 percent of the cross-sectional variance in size. The remainder is due to ex-post shocks that materialized in the first year. Considering older age groups, the contribution of ex-ante heterogeneity declines, but remains high. At age twenty, ex-ante factors account for 47 percent of the size variance among establishments, and around 40 among firms. In the data, more than seventy percent of the businesses are twenty years old or younger. Our results show that, among these businesses, ex-ante factors are a key determinant of size. Increasing age towards infinity, the contribution of ex-ante heterogeneity stabilizes at around 45 percent for establishments and 35 percent for firms. Therefore, even among very old business as ex-ante factors contribute to a large chunk of the dispersion in size.

G.3 Structural model: results for establishments

This Appendix provides results for the structural model using establishment-level data. Table G.13 shows the parameter estimates and Figure G.18 depicts the model fit.

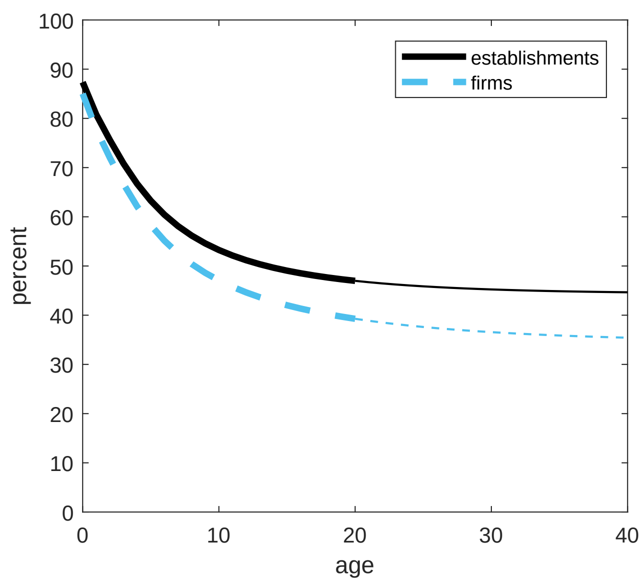
Figures G.19 and G.20 then establish that, also for establishments, ex-ante factors are a dominant force when it comes to the cross-sectional variation in employment and the establishment selection by age, respectively.

Figure G.16: Autocovariance matrices: statistical models versus data



Note: Autocovariance of log employment between age $a = h + j$ and age $h \leq a$ in the data, and in the baseline model. Results are shown for firms and establishments using the balanced panel.

Figure G.17: Contribution of ex-ante heterogeneity to cross-sectional employment dispersion



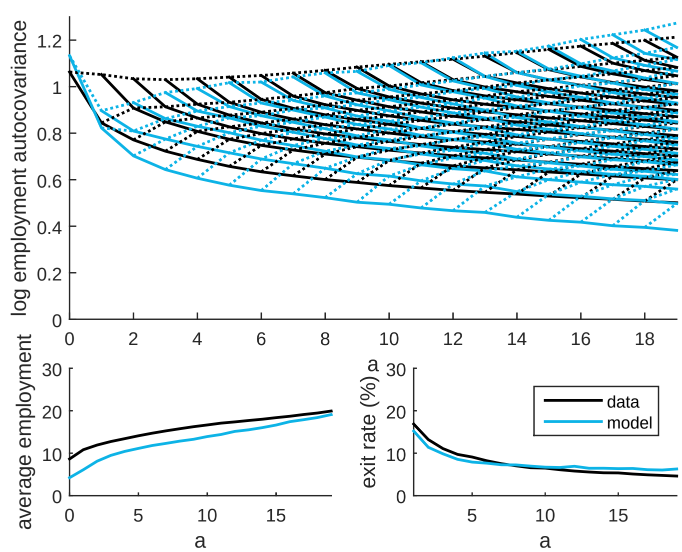
Note: Contribution of the ex-ante component, $\ln n_{i,a}^{EXA}$, to the cross-sectional variance of log employment, by age. Thin lines denote age groups not directly used in the estimation. The decomposition is based on Equation (2) with $j = 0$.

Table G.13: Parameter values (establishments)

	parameter	value
		set a priori
β	discount factor	0.96
η	elasticity of substitution	6.00
f^e	entry cost	0.448
		estimated
f	fixed cost of operation	547
δ	exogenous exit rate	0.044
μ_θ	permanent component θ, mean	−1.758
σ_θ	permanent component θ, st. dev.	1.309
$\sigma_{\widetilde{u}}$	initial condition u_{-1}, st. dev.	1.541
$\sigma_{\widetilde{v}}$	initial condition v_{-1}, st. dev.	1.206
σ_ε	transitory shock ε, st. dev.	0.303
σ_z	noise shock z, st. dev.	0.211
ρ_u	permanent component, persistence	0.393
ρ_v	transitory component, persistence	0.987

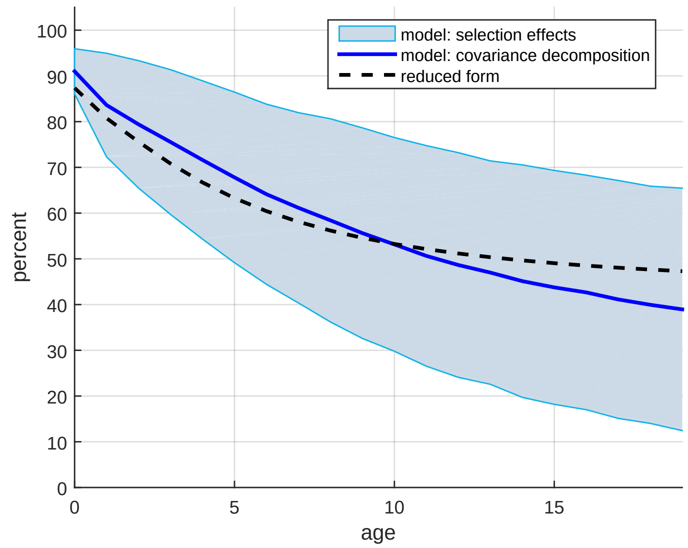
Notes: parameter values. Top three parameters are calibrated as discussed in the main text. The remaining parameters are set such that the model matches the empirical autocovariance of employment and the age profiles of average size and exit rates from age 0 to 19.

Figure G.18: Targeted moments: data and structural model (establishments)



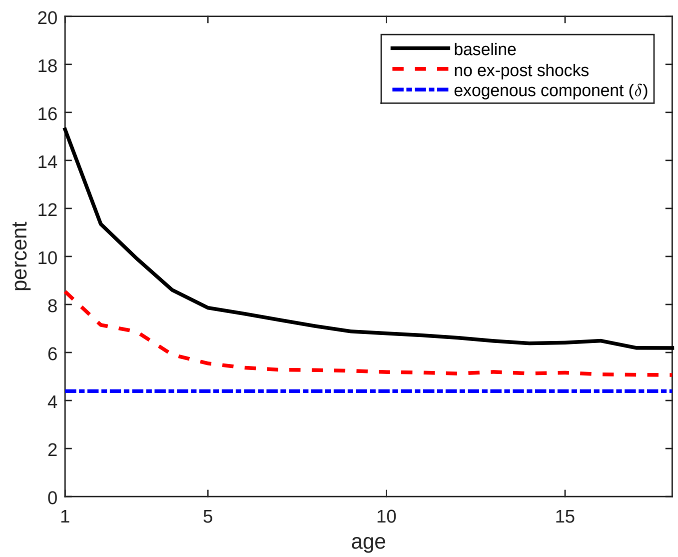
Notes: Top panel: Autocovariances of log employment between age $a = h + j$ and age $h \leq a$ in the data and the model, for a balanced panel of firms surviving up to at least age $a = 19$. Bottom left panel: Average employment by age a (unbalanced panel). Bottom right panel: exit rate by age a.

Figure G.19: Contribution of ex-ante heterogeneity to cross-sectional employment dispersion (establishments)



Note: Contributions to total cross-sectional variance by age. "Reduced-form" refers to the estimates from Figure 4, "model: covariance decomposition" is the decomposition based on the second line in Equation 4. The shaded areas ("model: selection band") is constructed based on the first equality in Equation 4 by attributing, in turn, the term $2Cov(\ln\varphi_i^{EXA}, \ln\varphi_i^{EXP})$ fully to the ex-ante component and to the ex-post component.

Figure G.20: Exit rates (establishments)



Note: exit rates by age in the baseline model, exit rates in and a counterfactual economy in with no ex-post demand shocks (but with exogenous exit), and the exogenous exit rate δ.

References

Blundell, R., L. Pistaferri, and I. Preston (2008): "Consumption Inequality and Partial Insurance," *American Economic Review*, 98(5), 1887–1921.

Chamberlain, G. (1984): "Panel Data," in *Handbook of Econometrics, Volume II*, ed. by Z. Griliches, chap. 22, pp. 1274–1318. Elsevier Science, Amsterdam, North-Holland.

Decker, R., J. Haltiwanger, R. Jarmin, and J. Miranda (2016): "Where Has All the Skewness Gone? The Decline in High-Growth (Young) Firms in the U.S.," *European Economic Review*, 86, 4–23.

Fort, T. C., S. D. Klimek, et al. (2016): "The Effects of Industry Classification Changes on US Employment Composition," Discussion paper, Tuck School at Dartmouth.

Guvenen, F. (2009): "An Empirical Investigation of Labor Income Processes," *Review of Economic Dynamics*, 12(1), 58–79.

Guzman, J., and S. Stern (2015): "Where is Silicon Valley?," *Science*, 347(6222), 606–609.

——— (forthcoming): "The State of American Entrepreneurship: New Estimates of the Quantity and Quality of Entrepreneurship for 32 US States, 1988-2014," *In American Economic Journal: Economic Policy.*

Heckler, D. E. (2005): "High-technology employment: a NAICS-based update," *Monthly Lab. Rev.*, 128, 57.

Jarmin, R. S., and J. Miranda (2002): "The longitudinal business database," Discussion paper, Census Bureau.

Lee, Y., and T. Mukoyama (2015): "Productivity and Employment Dynamics of US Manufacturing Plants," *Economics Letters*, 136, 190–193.

Pugsley, B. W., and H. Wheeler (2018): "Longitudinal Linking of Enterprises in the LBD and SSL," Discussion paper, CES TN-2018-02.

Rouwenhorst, K. G. (1995): "Asset Pricing Implications of Equilibrium Business Cycle Models," in *Frontiers of Business Cycle Research*, ed. by T. Cooley, pp. 294–330. Princeton University Press.

Sedláček, P., and V. Sterk (2017): "The Growth Potential of Startups Over the Business Cycle," *American Economic Review*, 107(10), 3182–3210.