(For online publication)

Appendix to: The Growth Potential of Startups over the Business Cycle

Petr Sedláček Vincent Sterk

Contents

A Empirical robustness exercises

This section supports the results presented in the main text by conducting several robustness checks. In particular, we check robustness with respect to alternative detrending methods (A.1), different ways of constructing employment levels (A.2), and measuring data over two-year rather than one-year windows (A.3). Further, we re-visit the stylized fact using a panel regression approach (A.4), we analyze data for establishments rather than firms (A.5), we analyze underlying micro data which allow us to follow cohorts of establishments beyond age five (A.6), we re-visit the stylized facts within sectors (A.7) and we consider the importance of very small firms (A.8). Finally, we investigate time variation in the contributions of the intensive and extensive margins to cohort-level employment variations over the sample period (A.9).

A.1 Detrending method

Figure 1 in the main text reports results based on log-differences from sample means. Nevertheless, the results change very little when one considers HP-filtered or linearly detrended data instead.[1] Figure 1 shows results for HP-filtered (left panel) and linearly detrended (right panel) employment levels of entrants, five year old firms (of the same cohort) and aggregate employment. The figure displays almost identical patterns as Figure 1 in the main text.

Similarly, Figure 2 plots the autocorrelations of cohort-level and aggregate employment at various ages when using a linear trend and the HP filter to detrend the data. The HP filtered data gives somewhat stronger results, but qualitatively the results remain the same even with linearly detrended data.

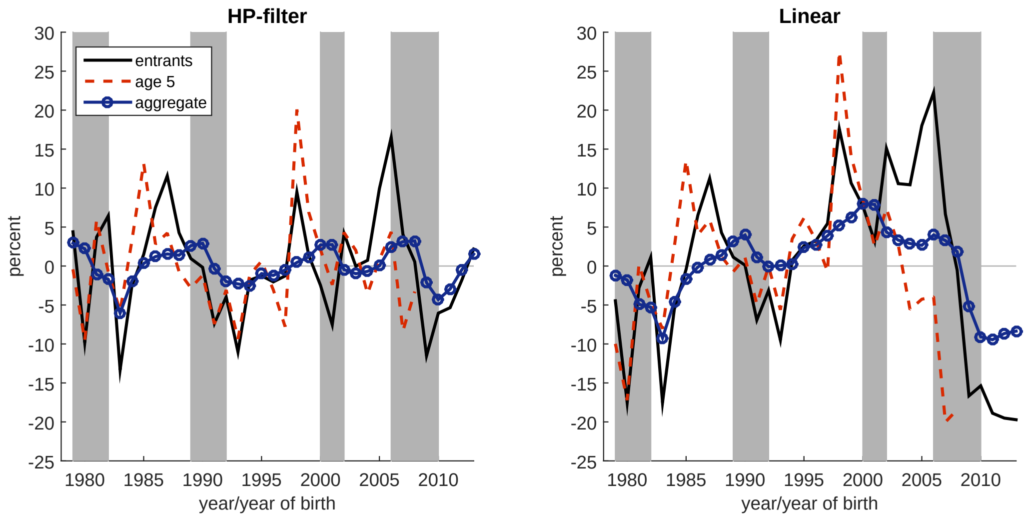
Finally, Figure 3 plots the contributions of average size to the variance of cohort-level employment at various ages using linearly detrended and HP-filtered data. The results barely change when considering linearly detrended data instead of HP-filtered data.

A.2 Alternative ways of constructing cohort-level employment

The BDS contains several variables one can use to construct measures of employment levels. The main text used cumulative net job creation (NJC) as a definition of employment levels. Alternatively, one can use employment levels directly provided by the BDS (the "EMP" variable). One can also make use of the "DENOM" variable which is an average between employment in the current and in the previous year. Therefore, as a first robustness check, we redo our analysis using the EMP variable. As a second alternative,

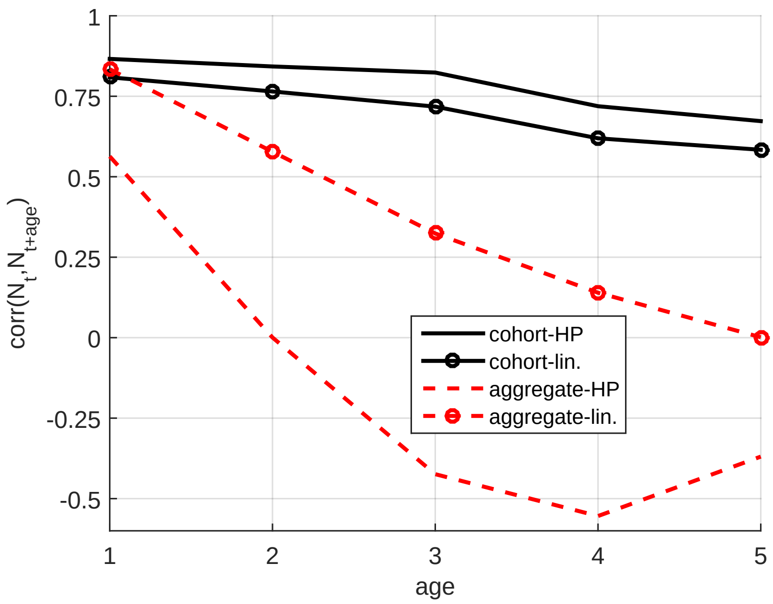
[1] Similarly, using HP-filtered data with a smoothing parameter 6.23 (instead of 100 as in the main text) as suggested by Ravn and Uhlig (2002) changes little.

Figure 1: Cohort-level employment at age 0 and 5 by year of birth and aggregate employment by year: alternative detrending methods



Notes: Cohort-level and aggregate employment plotted in percentage deviations from an HP-filter (left panel) and linear (right panel) trend, respectively. Shaded areas are the NBER recessions. Source: BDS.

Figure 2: Employment autocorrelations: alternative detrending methods

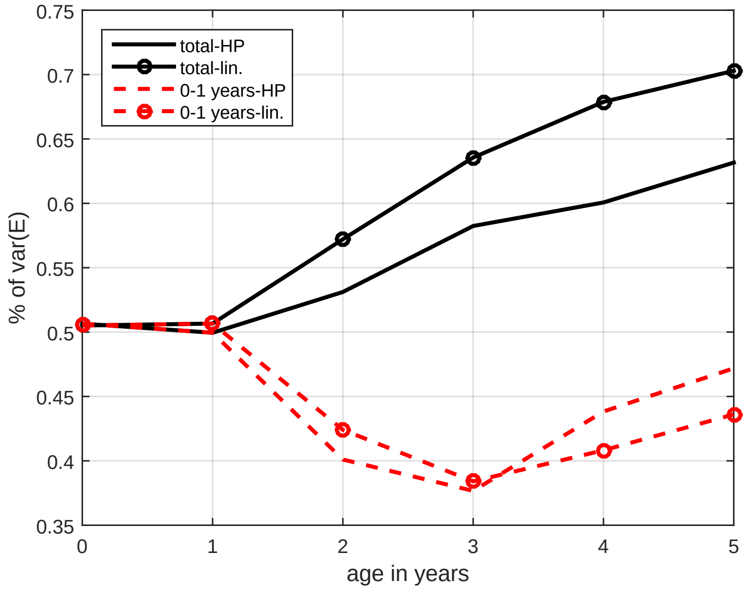


Notes: Correlation coefficients of cohort-level and aggregate employment in year $t = 0$ and in year $t+age$, with $age = 1, 2, 3, 4, 5$. "Linear" refers to linearly detrended data, "HP filter" refers to HP-filtered data. Source: BDS.

we construct employment as $\frac{1}{2}(NJC + 2DENOM)$. This, however, turns out to be fully consistent with the "EMP" variable in the latest vintage of the BDS data.

For various reasons, employment based on cumulating net job creation does not yield exactly the same numbers as the "EMP" variable in the BDS. For instance, the BDS documentation states that net job creation data is cleaned from observed entrants that

Figure 3: Contribution of average size to employment variation: alternative detrending methods



Notes: The figure plots contributions of average firm size to the variation in cohort-level employment of five year old firms expressed as percent of the total variation. "Total" denotes the share of employment variation explained by average size overall. "0-1 years" refers to the contribution of average size in the first two years of existence to total employment variation. "Linear" refers to linearly detrended data, "HP filter" refers to HP-filtered data.
Source: BDS and authors' calculations.

Table 1: Correlations of entrant size with various business cycle indicators

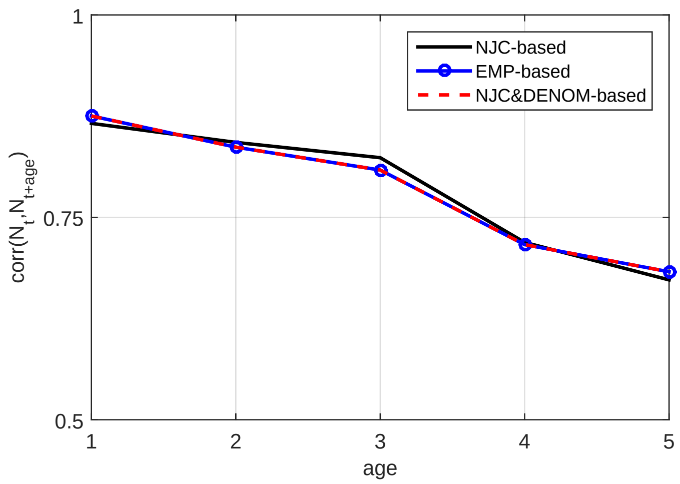
	e-rate$_L$	GDP$_L$	e-rate$_{HP}$	GDP$_{HP}$	e-rate	GDP$_{gr}$
NJC-based	0.69	0.79	0.37	0.41	0.66	0.48
EMP-based	0.67	0.77	0.36	0.40	0.64	0.46
NJC & DENOM-based	0.67	0.77	0.36	0.40	0.64	0.46

Notes: The table reports correlation coefficients between various business cycle indicators and entrant sizes. "e-rate" referes to 1 minus the unemployment rate, the subscript "L", "HP" and "gr" refer, respectively, to linearly detrended data, HP-filtered data and growth rates. "NJC-based" refers to the construction of employment as in the main text, "EMP-based" refers to directly using the EMP variable in the BDS and "NJC & DENOM-based" refers to employment being defined as $0.5(NJC+2DENOM)$.
Source: authors' calculations.

are not believed to be true startups, while the employment data is not. In particular, "...it may be determined that an establishment's entry/exit as shown by the data is not credible. These establishments are excluded from the change calculations in a given year" (http://www.census.gov/ces/dataproducts/bds). Therefore, we check whether our results are not driven by a particular way of constructing employment levels.

Table 1 reports the correlation coefficients of business cycle indicators with entrant average size computed using the three alternative ways of constructing employment. All

Figure 4: Employment autocorrelations: alternative construction of employment



Notes: Correlation coefficients of cohort-level employment in year $t = 0$ and in year $t + age$, with $age = 1, 2, 3, 4, 5$. "NJC-based" refers to the construction of employment as in the main text, "EMP-based" refers to directly using the EMP variable in the BDS and "NJC & DENOM-based" refers to employment being defined as $0.5(NJC + 2DENOM)$. Source: BDS.

three methods deliver very similar results. As mentioned above, the employment levels based on cumulative net job creation and on net job creation and the DENOM variable yield identical results.

Figure 4 plots the autocorrelations of cohort-level employment based on the three construction methods. Again, the high persistence of cohort-level employment is independent of the construction method used. Note that aggregate employment is not affected by the construction method of cohort-level employment.

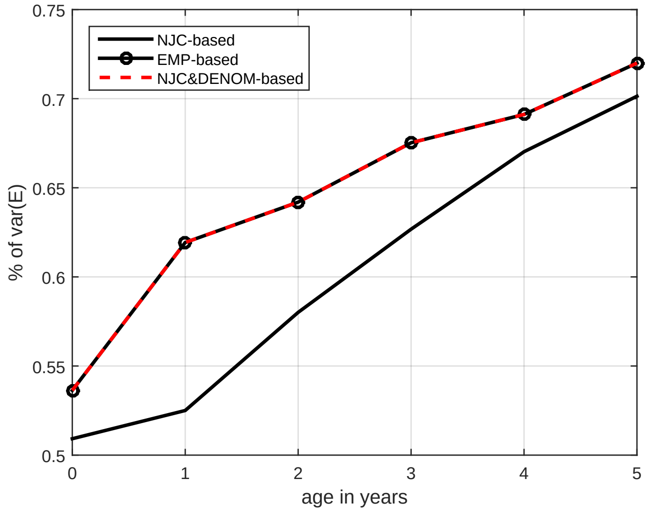
Finally, Figure 5 plots the contribution of average size to the variation in total employment at various ages, again based on different construction methods for employment. In all cases, the importance of average size in driving the variation in total employment increases with the age of the cohort and explains the majority of volatility of employment among five year old firms. The alternative methods of constructing employment levels deliver even stronger results with average size accounting for over 75% of employment variation at age five.

A.3 Results using data averaged over a two-year window

One concern may be that the annual timing of the BDS potentially introduces some undesired variation.[2] This may be valid in particular for entrants. To give an example, the entrant data include firms that start up just before the March deadline for reporting, even if they turn out to exit very shortly afterwards (i.e. not actually living through their first year). Although one could expect that, if anything, noise may weaken the correlations

[2]Noisiness of the data due to a low number of firm observations is unlikely to be an issue. The minimum number of observations in a given age-year cell is 196,397 for five year old firms.

Figure 5: Contribution of average size to employment variation: alternative construction of employment



Notes: The figure plots contributions of average firm size to the variation in cohort-level employment of five year old firms expressed as percent of the total variation. "NJC-based" refers to the construction of employment as in the main text, "EMP-based" refers to directly using the EMP variable in the BDS and "NJC & DENOM-based" refers to employment being defined as $0.5(NJC + 2DENOM)$.
Source: BDS and authors' calculations.

Table 2: Correlations of entrant size with various business cycle indicators: averaged data

e-rate$_L$	GDP$_L$	e-rate$_{HP}$	GDP$_{HP}$	e-rate	GDP$_{gr}$
0.66	0.78	0.24	0.32	0.62	0.54

Notes: The table reports correlation coefficients between various business cycle indicators and entrant sizes. "e-rate" referes to 1 minus the unemployment rate, the subscript "L", "HP" and "gr" refer, respectively, to linearly detrended data, HP-filtered data and growth rates. All variables are averaged over two years prior to detrending. Source: authors' calculations.

we report, we check for robustness by repeating our empirical analysis on a sample of 2-year averages. Specifically, all variables are average across two years backwards (that is a 1980 observation is the average between a 1979 and a 1980 observation etc.). This should also alleviate concerns that our results are driven by a certain year/cohort.

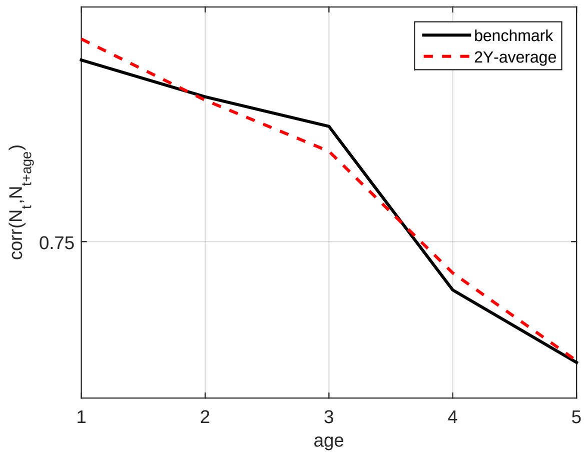
Table 2 shows the correlations between entrant size and various business cycle indicators. Averaging over two years produces only very small changes in the autocorrelations.

Figure 6 and 7 show the autocorrelations of cohort-level employment and the contributions of average size to employment variation both for the benchmark specification and for the two-year averaged data. The results for the two methods are almost identical.

A.4 Panel regressions

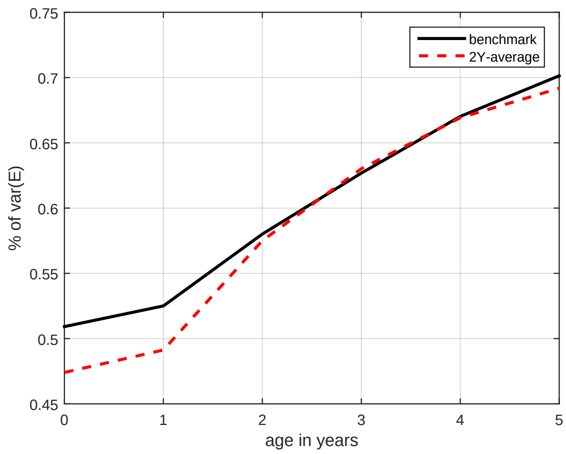
Instead of inspecting the autocorrelations of detrended cohort-level data as in Figure 2, one can estimate the following panel regression:

Figure 6: Employment autocorrelations: averaged data



Notes: Correlation coefficients of cohort-level employment in year $t = 0$ and in year $t + age$, with $age = 1, 2, 3, 4, 5$. "Benchmark" denotes the benchmark time series used in the main text, "2Y-average" refers to the case when the time-series is average over two years prior to detrending. Source: BDS and authors' calculations.

Figure 7: Contribution of average size to employment variation: averaged data



Notes: The figure plots contributions of average firm size to the variation in cohort-level employment of five year old firms expressed as percent of the total variation. "Benchmark" denotes the benchmark time series used in the main text, "2Y-average" refers to the case when the time-series is average over two years prior to detrending. Source: BDS and authors' calculations.

$$\ln N_{a,t} = \alpha_0 + \alpha_1 \ln N_{0,t-a} + \alpha_2 \ln N_{0,t} + \alpha_3 \ln N_{0,t-a} a + \alpha_4 a + \alpha_5 a^2 + u_{a,t}, \qquad (1)$$

where a indicates age and $u_{a,t}$ is the residual term. While $N_{0,t-a}$ is entrant job creation of the given cohort at birth, $N_{0,t}$ is the employment level of *current* entrants and as such measures the current aggregate conditions. The elasticity of cohort-level employment at a given age with respect to that at birth is then given by $\alpha_1 + \alpha_3 a$. We estimate the above panel regression using data from 1979 until 2013 for firms aged 1 to 5 years.

The left panel of Table 3 reports the elasticities of cohort-level employment at ages

Table 3: Elasticity of cohort-level employment and average size with respect to entrant size

	employment		average size	
$a = 1$	0.831	(0.124)	0.901	(0.208)
$a = 2$	0.812	(0.125)	0.882	(0.208)
$a = 3$	0.794	(0.140)	0.864	(0.231)
$a = 4$	0.776	(0.166)	0.845	(0.270)
$a = 5$	0.758	(0.199)	0.826	(0.320)

Notes: The table reports elasticities of cohort-level employment and average size at ages 1 to 5 with respect to entrant average size together with the appropriate standard errors in brackets.

1 to 5 together with their respective standard errors. It shows that this alternative way of investigating the persistence in cohort-level employment delivers very similar results to those in Figure 2. One can run the panel regressions also for average size, rather than employment. The results are reported in the right panel of Table 3 and reveal even higher elasticities than for cohort-level employment. This is consistent with the variance decomposition in Figure 3 which shows that the majority of the variation in cohort-level employment is driven by changes in average firm size.

A.5 Establishments

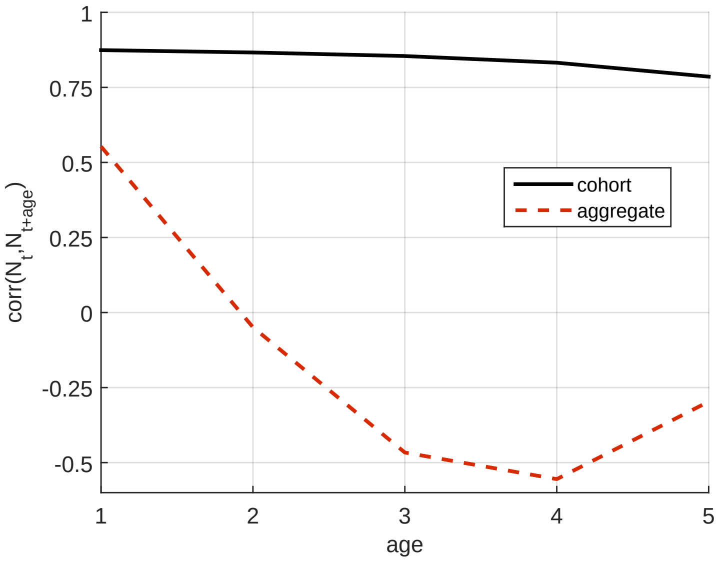
The main text documents new stylized facts for firms. Using the BDS data one can also inspect establishment-level information. An establishment is defined as a single physical location where business is conducted or where services or industrial operations are performed. A firm, on the other hand, is a business organization consisting of one or more establishments that were specified under common ownership or control. Therefore, the firm and the establishment are the same for single-establishment firms, but existing firms can create new establishments. The following paragraphs show that at the establishment-level our empirical findings remain to hold.

As for firms, the variation in the number of jobs created by new establishments is robustly pro-cyclical and large. The correlation coefficient of establishment entrant job creation with the employment rate (real GDP) is 0.63 (0.69) using linear detrending. The correlations when considering HP-filetered data or data in levels remain large and positive.[3] Moreover, the volatility of jobs created by new establishments (in logs) is large, amounting to 5.4 times that of the volatility of (log) real GDP.

Figure 8 shows the correlation coefficient of employment in year t with that in year $t+a$ of the same cohort. The figure shows a very high persistence for cohort-level employment, which strongly contrasts that of aggregate employment. Notice, that the correlation of

[3]For HP-filtered data the correlation coefficients are 0.35 (0.38) when considering the employment rate (real GDP) as business cycle indicators. The correlation with the level of the employment rate (growth rate in real GDP) is 0.66 (0.15).

Figure 8: Employment autocorrelations: establishments



Notes: Correlation coefficients of employment at establishments in year $t = 0$ and in year $t + a$, with $a = 1, 2, 3, 4, 5$, at both the level of a cohort born in period $t = 0$ and at the aggregate level. Source: BDS.

employment of entrants and five year old establishments of the same cohort is even higher than that computed using firm-level data (0.79 for establishments compared to 0.70 for firms).

Finally, when decomposing employment variation of five year old establishments into the intensive (average establishment size) and extensive (number of establishments) margins, one finds that the majority is driven by the intensive margin (57%). This contribution is somewhat smaller than the one found for firms, which is to be expected, given that firm growth may involve opening new establishments.

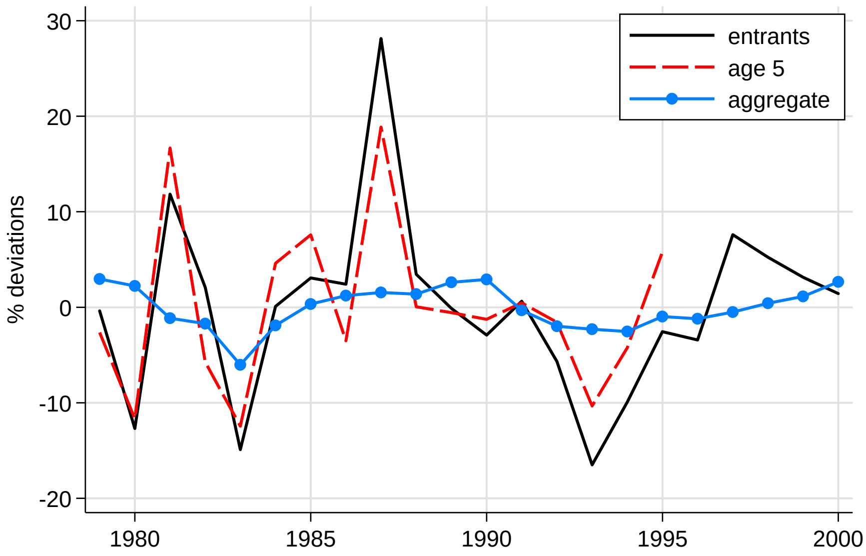
A.6 Micro-data

The main text analyzes aggregated, publicly available firm-level data. Appendix A.5 showed that the stylized facts remain to hold also for establishments. In this Appendix, we check the robustness of our main results by verifying that the stylized facts found in the publicly available BDS data (which are firm-level) also hold in the so called Longitudinal Business Database (LBD), which contains the micro data from which the BDS is constructed . To this end, we redo our analysis using the "Synthetic LBD" (SynLBD), to which we were able to obtain access.[4]

The SynLBD includes 21 million establishment records covering the period between 1976 and 2000. This shorter time frame together with the focus on establishments rather

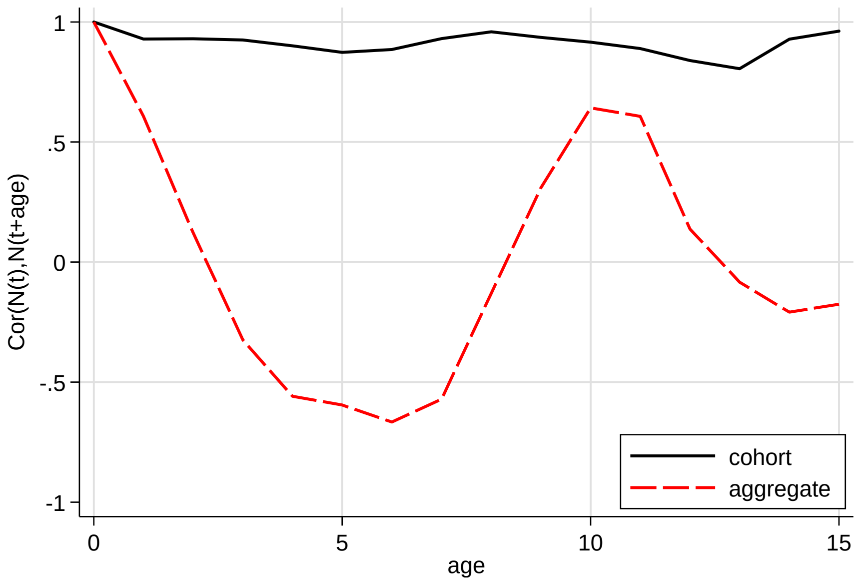
[4]Detailed information on the SynLBD and its methodology can be found at `http://www.census.gov/ces/dataproducts/synlbd/`. We thank Javier Miranda for help using the synthetic LBD data.

Figure 9: Cohort-level employment at age 0 and 5 by year of birth and aggregate employment by year: SynLBD data



Notes: Cohort-level and aggregate employment (level) are plotted in percentage deviations from a HP-trend. Source: BLS, SynLBD.

Figure 10: Employment autocorrelations: SynLBD data

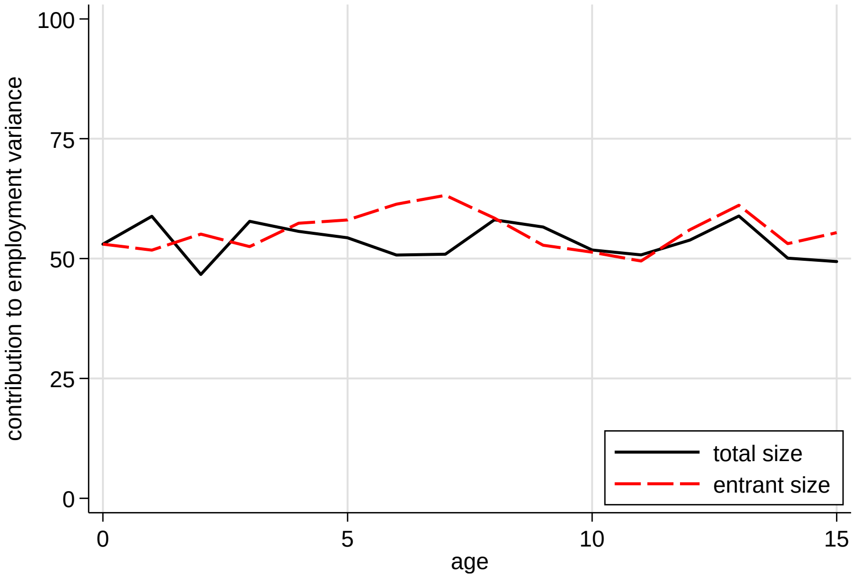


Notes: Correlation coefficients of employment in year t = 0 and in year t + age, with age = 1,2,...15 at both the level of a cohort born in period t = 0 and at the aggregate level. Source: BLS, SynLBD.

than firms are the reasons why the main text uses the aggregated BDS data. The main advantage of the SynLBD data used in this section is that it is possible to track establishments for a longer horizon than the five years reported in the BDS.

Figures 9 to 11 are the LBD equivalents of Figures 1 to 3 in the main text. Figure 9 shows the total employment level of entering establishments. Like in the BDS data, entrant employment correlates positively with aggregate employment (blue line with circles). The same figure also plots the employment level of the cohorts five years after

Figure 11: Contribution of average size to employment variation: SynLBD data



Notes: The figure plots contributions of average firm size and entrant size to the variation in cohort-level employment as a percentage of its variation at different ages. Source: SynLBD.

birth (red dashed line), with the year of birth on the horizontal axis. Like in the BDS data, there is a strong positive co-movement between employment in the year of birth and employment within the same cohort five years after birth.

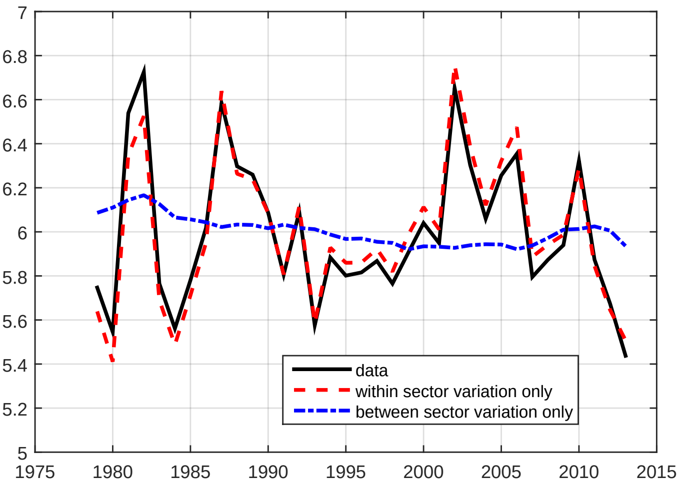
The persistence of cohort-level employment is depicted in Figure 10. The black solid line plots the correlation between cohort-level employment by age with the employment level in the year of entry. The figure makes clear that cohort-level employment is extremely persistent, even up to 15 years after entry. For comparison, the figure also plots the autocorrelation in aggregate employment, which is generally much less persistent.

Finally, Figure 11 conducts the variance decomposition described in the main text. The black solid line shows that, regardless of age, about half of the variations in employment across cohorts is driven by the "intensive margin" (average firm size). The remainder is driven by the extensive margin (number of firms). Next, we further decompose the contribution of the intensive margin by age. The red dashed line shows the contribution of entrant size, which fluctuates around 50 percent. Thus, fluctuations in the intensive margin appear almost fully driven by the year of entry. Overall, the results based on the SynLBD data are very close to those in the main text based on the firm-level BDS data.

A.7 Sectoral evidence

This section investigates to what extent our empirical findings may be driven by cyclical sectoral composition changes of entering firms. For this purpose we first use the BDS

Figure 12: Average entrant size: data and sectoral counterfactuals



Notes: "within sector variation only" average size is constructed by fixing the sectoral shares of entrants to their sample average. "Between sector variation only" average size is constructed by fixing the average size of entrants within sectors to their respective sample averages. Source: BDS, authors' calculations.

sectoral breakdown, which includes information on nine 1-digit sectors.[5] Second, we use an additional dataset, the Quarterly Workforce Indicators (QWI). While this dataset has some limitations in regards to our needs, it's benefit is a much finer sectoral break-down. In both cases, we show that our stylized facts hold within (even narrowly) defined sectors.

A.7.1 Broad sectoral evidence

To gain insight into the importance of sectoral shifts for aggregates, we compute a two counterfactual time series of average entrant size. First, we construct a counterfactual entrant size under the assumption that the distribution of the number of entrants over the nine sectors remains fixed over time, setting the fractions equal to their sample averages. This series captures variation that is due within sector variations in average size only. Second, we compute a counterfactual series that captures only between-sector shifts, by setting the average entrant size within each sector equal to the sample average, but let fractions of entrants in the nine sectors to vary over time as in the data. Figure 12 displays the two counterfactual time series, as well as the actual series for average size within newborn cohorts. It is immediately clear that within-sector variations account for almost all of the variation in average size; between-sector shifts appear to play an extremely limited role.

Next, we repeat our empirical analysis within each of the nine sectors in the BDS separately. The results are reported in Table 4 and show that our earlier findings also broadly hold within sectors. This gives further support that the economy-wide results

[5]The data is broken down into the following sectors: (i) Agriculture, Forestry, and Fishing, (ii) Mining, (iii) Construction, (iv) Manufacturing, (v) Transportation, Communication, and Public Utilities (vi) Wholesale Trade, (vii) Retail Trade, (viii) Finance, Insurance, and Real Estate, (iv) Services.

Table 4: Summary of stylized facts within sectors

		AGR	MIN	CON	MAN	TCU	WHO	RET	FIRE	SRV
Cyclicality										
firms	e-rate	0.53	−0.39	0.81	0.68	0.27	0.57	0.29	0.66	0.51
	GDP	0.62	−0.42	0.84	0.78	0.36	0.65	0.46	0.78	0.63
establishments	e-rate	0.64	−0.41	0.80	0.69	0.32	0.60	0.65	0.66	0.55
	GDP	0.72	−0.44	0.85	0.77	0.41	0.65	0.70	0.74	0.61
Persistence										
firms	cohort	0.05	0.79	0.60	0.43	0.75	0.58	0.32	0.11	0.58
establishments	cohort	0.41	0.84	0.56	0.52	0.84	0.69	0.58	0.61	0.76
	total	−0.07	0.37	−0.14	0.25	0.19	0.15	−0.11	−0.29	0.20
Variance decomposition										
firms	extensive	0.39	0.53	0.83	0.33	0.04	0.44	0.39	0.21	0.12
	intensive	0.61	0.47	0.17	0.67	0.96	0.56	0.61	0.79	0.89
establishments	extensive	0.36	0.58	0.87	0.49	0.40	0.48	0.56	0.56	0.20
	intensive	0.64	0.42	0.13	0.51	0.60	0.52	0.44	0.44	0.80

Notes: "Cyclicality" reports the correlation coefficients between linearly detrended log job creation of entrants and the employment rate or real GDP in the different sectors for firms and establishments. "Persistence" reports the correlation coefficients between (linearly detrended) entrant job creation and employment in 5 year old firms or establishments within the same cohort, both for the individual cohorts and for the sector as a whole. Finally, "Variance decomposition" reports the contribution of the extensive (number of firms or establishments) and intensive (average size) margin to variation in employment of 5 year old firms or establishments (based on linearly detrended data).

Table 5: Elasticity of cohort-level employment and size with respect to that at entry

age	AGR	MIN	CON	MAN	TCU	WHO	RET	FIRE	SRV
					employment				
$a = 1$	0.96(0.17)	0.35(0.11)	0.63(0.11)	1.08(0.14)	1.43(0.09)	0.70(0.12)	0.66(0.30)	0.94(0.15)	1.75(0.12)
$a = 2$	0.94(0.17)	0.37(0.11)	0.63(0.11)	1.05(0.13)	1.40(0.09)	0.70(0.12)	0.65(0.30)	0.92(0.15)	1.71(0.12)
$a = 3$	0.92(0.16)	0.38(0.11)	0.63(0.11)	1.02(0.13)	1.37(0.09)	0.70(0.12)	0.63(0.29)	0.90(0.15)	1.67(0.11)
$a = 4$	0.90(0.16)	0.39(0.11)	0.62(0.11)	0.99(0.13)	1.34(0.09)	0.70(0.11)	0.61(0.29)	0.88(0.15)	1.63(0.11)
$a = 5$	0.87(0.16)	0.40(0.11)	0.62(0.10)	0.95(0.13)	1.31(0.09)	0.70(0.11)	0.59(0.28)	0.86(0.14)	1.59(0.11)
					average size				
$a = 1$	0.58(0.28)	0.76(0.23)	1.21(0.28)	0.91(0.24)	1.26(0.14)	0.44(0.31)	0.43(0.30)	0.53(0.22)	1.15(0.25)
$a = 2$	0.56(0.27)	0.75(0.23)	1.17(0.28)	0.90(0.24)	1.26(0.13)	0.46(0.31)	0.40(0.30)	0.55(0.21)	1.14(0.25)
$a = 3$	0.54(0.27)	0.75(0.22)	1.14(0.27)	0.88(0.23)	1.25(0.13)	0.49(0.30)	0.37(0.29)	0.57(0.21)	1.13(0.24)
$a = 4$	0.52(0.26)	0.74(0.22)	1.10(0.27)	0.87(0.23)	1.25(0.13)	0.51(0.30)	0.35(0.29)	0.59(0.21)	1.13(0.24)
$a = 5$	0.50(0.26)	0.74(0.22)	1.07(0.27)	0.85(0.23)	1.24(0.13)	0.53(0.29)	0.32(0.28)	0.61(0.20)	1.12(0.24)

Notes: The table reports elasticities of cohort-level employment and average size at ages 1 to 5 with respect to that at entry (standard errors in brackets).

are not driven by cyclical sectoral shifts.

In particular, all sectors are characterized by large persistence in cohort-level employment, which is in stark contrast to the persistence found in the sector as a whole. Most sectors are also characterized by strongly pro-cyclical job creation by entrants. The exception is mining which is strongly counter-cyclical. Mining, however, accounts for a very small fraction of firms and employment in the economy and therefore it is unlikely to influence the aggregate cyclical properties. Finally, in most sectors it is the intensive margin which drives the majority of variation of employment among five year old firms or establishments. The exception is construction where the intensive margin contributes with 17%.

Table 5 reports the elasticities of cohort-level employment (average size) with respect to entrant employment (average size) based on the panel regressions described in Appendix A.4. Again, the resulting elasticities within all sectors are similar to those in the aggregate with the exception of manufacturing for employment and whole sale and retail trade for average size which are somewhat lower.

Our findings are also related to results of Lee and Mukoyama (2013) who document that in recessions entering plants in manufacturing are on average larger than those entering in booms. Their findings are based on the Annual Survey of Manufacturers from the U.S. Census Bureau for the period 1972-1997. Their measure of the business cycle is given by the growth rate of manufacturing output. Interestingly, we confirm their finding in the BDS. When we compute the correlation of average size of newborn firms in manufacturing and the growth rate of real GDP, we find it is significantly negative. However, for other de-trending methods and business cycle indicators and when using data on establishments this correlation drops to virtually zero in the BDS data.

A.7.2 Evidence at the 4-digit industry level

The Quarterly Workforce Indicators data includes information on employment broken down by firm age and 4-digit industry at the state level (U.S. wide data at the 4-digit level is not available). While this dataset provides additional valuable information, it lacks certain features important for our main analysis and we therefore focus on the BDS in the main text. First, the data starts (at the earliest) in 1990. Second, there is no information on the number of firms. Finally, the coverage across states is relatively sparse.[6] Thus, while the QWI is relatively suitable for studying patterns at a narrow industry level, but it is relatively less suitable for studying aggregate business cycle patterns.

We first investigate our stylized fact regarding the cyclicality of entrant employment. Firm age is grouped in two-year bins in the QWI, rather than one-year bins as in the BDS. Therefore, entrant employment refers to employment in firms aged 0 and 1 years.

[6] To give an idea about the data availability, out of the possible 51×312 possible state-industry observations, 27 percent are available in 2000 (only about 2 percent are available in 1990).

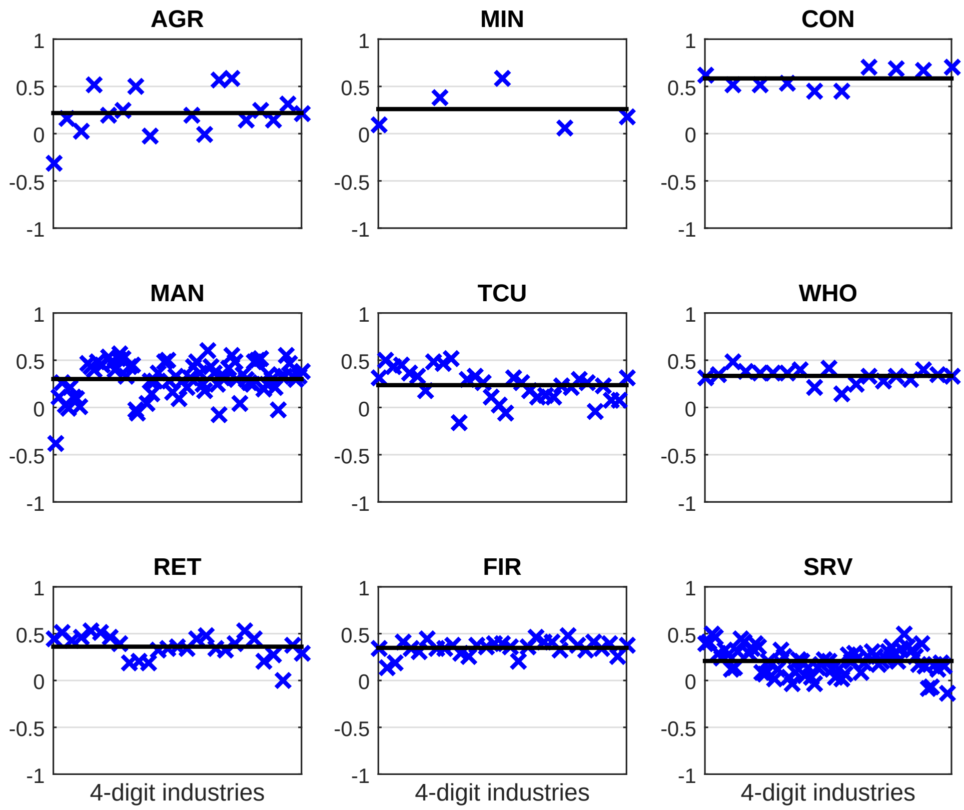
Table 6: Cyclicality of entrant employment

	AGR	MIN	CON	MAN	TCU	WHO	RET	FIR	SRV
corr(N_{0-1},Nr_{state})	0.22	0.26	0.60	0.32	0.26	0.35	0.37	0.36	0.21
corr(N_{0-1},Nr_{agg})	0.21	0.15	0.59	0.34	0.23	0.34	0.40	0.36	0.24
corr(N_{0-1},Δ GDP)	0.16	0.02	0.37	0.26	0.17	0.21	0.27	0.21	0.15
nobs	246	62	404	752	432	549	1090	781	2486

Notes: The table reports correlations between the log-deviations of employment in 0-1 year old firms from their respective means in a given industry and state with business cycle indicators. As the latter, the top row takes the employment rate at the state-level, the second row takes the aggregate employment rate and the third row considers real GDP growth. The values are based on simple averages across individual industry-state correlations within the broad sectors. The bottom row indicates the number of such individual industry-state correlations available in each of the broad sectors. The broad categories: AGR, MIN, CON, MAN, TCU, WHO, RET, FIR and SRV stand for, respectively, agriculture, mining, construction, manufacturing, telecommunications, wholesale, retail trade, finance, insurance and real estate and services.

.

Figure 13: Cyclicality of entrant employment



Notes: The figure plots the correlation between employment of 0-1 year old firms (in log deviations from the respective mean) and the state-level employment rate for each industry-state time series, averaged over states. The nine panels group the individual industry-level correlations into the broader sectors of the BDS. The black line indicates the average within the broad sectors (identical to the first row in Table 6). The blue crosses indicate the individual industry correlations.

Table 7: Persistence of cohort-level employment

	AGR	MIN	CON	MAN	TCU	WHO	RET	FIR	SRV
corr($N_{0-1,t-2}$,$N_{2-3,t}$)	0.67	0.59	0.58	0.53	0.49	0.55	0.56	0.53	0.60
corr($N_{0-1,t-4}$,$N_{4-5,t}$)	0.54	0.49	0.27	0.43	0.31	0.33	0.40	0.36	0.46
nobs	164	35	287	525	333	494	753	691	1708

Notes: The table reports correlations between the log-deviations of employment in 0-1 year old firms from the respective cohort HP-trend with that of 2-3 (top row) and 4-5 (middle row) years old firms of the same cohort in each industry and state. The values are based on simple averages across individual industry-state correlations within the broad sectors. The bottom row indicates the number of such individual industry-state correlations available in each of the broad sectors. The broad categories: AGR, MIN, CON, MAN, TCU, WHO, RET, FIR and SRV stand for, respectively, agriculture, mining, construction, manufacturing, telecommunications, wholesale, retail trade, finance, insurance and real estate and services.

.

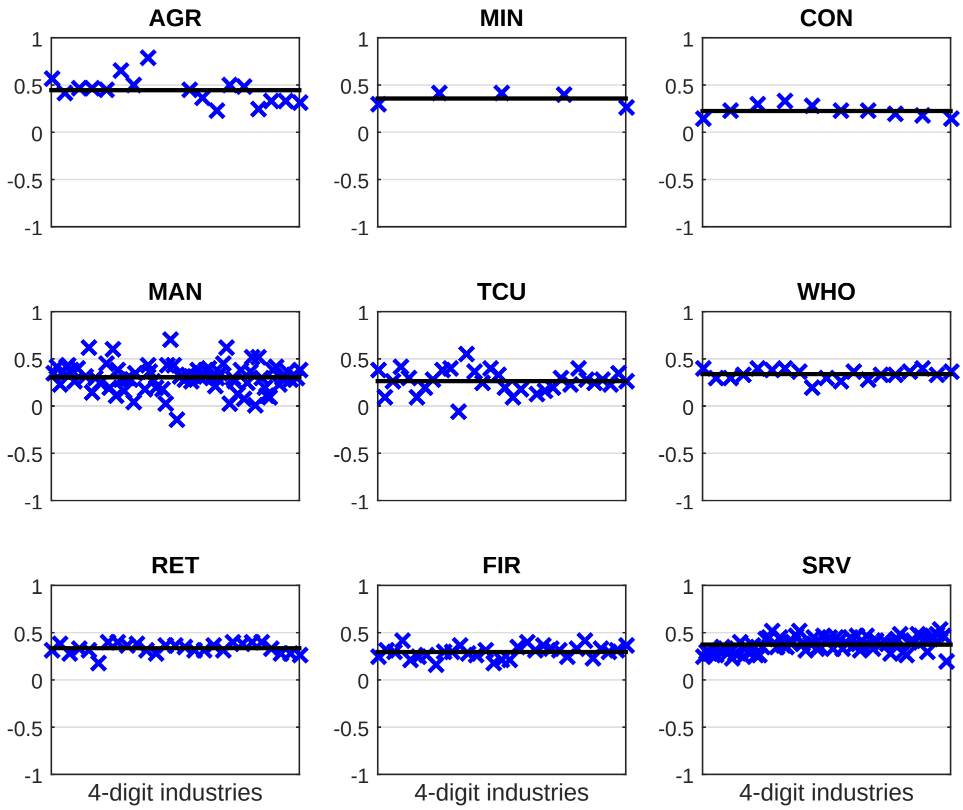
To this end, we computed, by 4-digit industry and by state, the correlation between state-level entrant employment (in log-deviations from the respective mean) and three business cycle indicators (state-level and aggregate employment rate and real GDP growth). For the sake of parsimonious presentation, we then averaged these correlations across states.[7] Table 6 shows the resulting average correlations within the broader sectors in the BDS. To get a sense of the distribution of these values across industries, Figure 13 shows scatter plots of the individual 4-digit industries (averaged over states) within the broader sectors. While there is heterogeneity across broad sectors and within them, the overall picture is that entrant employment is by-and-large pro-cyclical.

Next, we investigate our second stylized fact regarding the persistence of cohort-level employment within the narrow industries. We do so by computing the autocorrelation of cohort-level employment at age 0-1 with that of 2-3 year old firms and 4-5 year old firms, again by state and 4-digit industry. Table 7 shows these correlations, averaged over states and industries within the broader industry classes of the BDS. Figure 14 then plots the individual industry-level correlations of cohort-level employment of 0-1 year old firms with that of 4-5 year old firms, again averaged over states. As with the first stylized fact, there is heterogeneity across industries, but the persistence of cohort-level employment is a robust feature of the data.

Overall, we interpret the evidence from the QWI as giving additional support to our stylized facts, showing that they are broadly relevant in many sectors of the economy. Of course, there is heterogeneity across sectors and we exploit this heterogeneity in Appendix A.10 to provide further empirical support to our model mechanism.

[7] Taking simple averages or employment-weighted averages changes very little.

Figure 14: Persistence of cohort-level employment

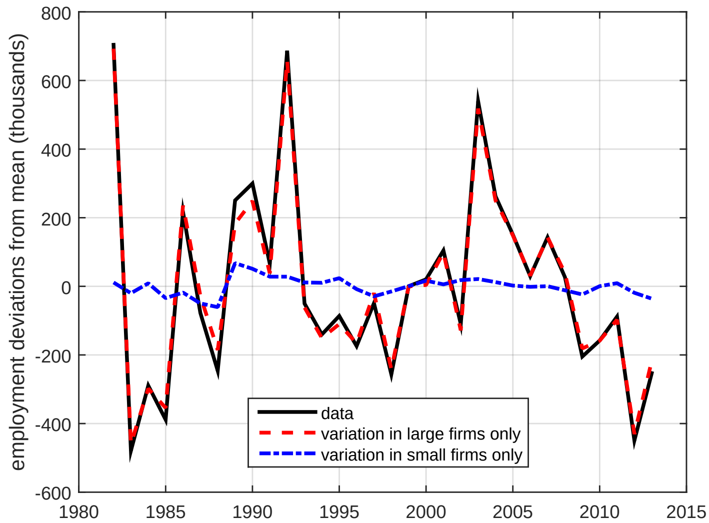


Notes: The figure plots the correlation between employment of 0-1 year old firms and that of 4-5 year old firms (in log deviations from the cohort-level HP trend) in each industry and state, averaged over states. The nine panels group the individual industry-level correlations into the broader sectors of the BDS. The black line indicates the average within the broad sectors (identical to the second row in Table 7). The blue crosses indicate the individual industry correlations.

A.8 The impact of very small firms

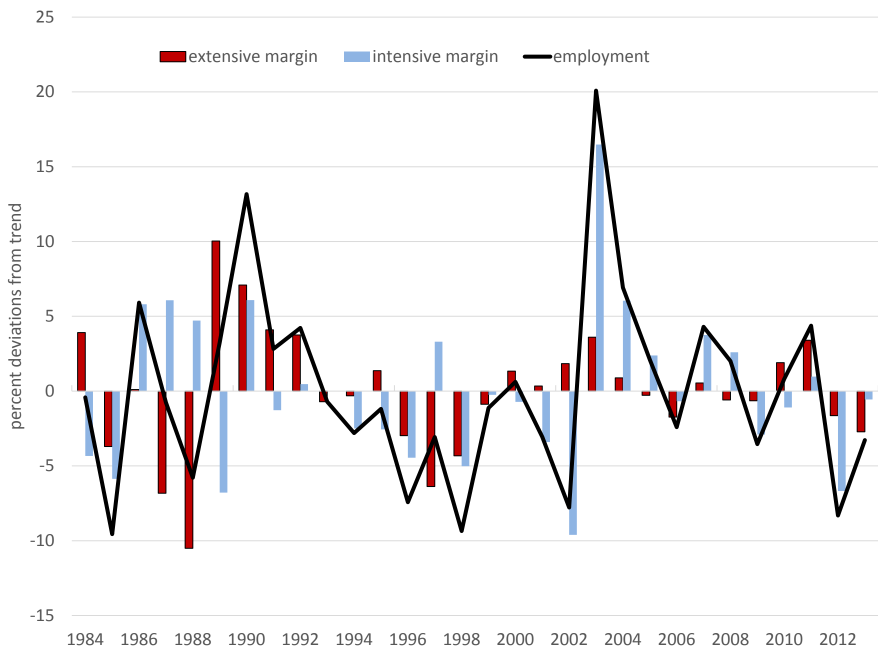
As mentioned in the main text, one possible explanation for our results is that they are driven by fluctuations in the entry of very small firms. To investigate the importance of small firms for the variation in cohort-level employment, Figure 15 plots the deviations from mean employment of five year old firms together with the relative contributions of small and large firms (where small is defined as less than 10 employees). The figure shows that the vast majority of employment variation is driven by large firms. This observation does of course not refute the existence of necessity entrepreneurship, but it appears unlikely that cyclical variations in this entrepreneurship motive are driving our stylized facts.

Figure 15: Employment of five year old firms: data and necessity entrepreneur counterfactuals



Notes: The figure plots employment of five year old firms in deviations from its mean ("data") together with the relative contributions of "small firms" and "large firms" to this variation. Small firms are defined as having fewer than 10 employees.

Figure 16: Contribution of extensive and intensive margin to variation in cohort-level employment of five year old firms



Notes: Contributions of average firm size and number of firms to cohort-level employment of five year old firms (in percent deviations from HP-filter trend).

A.9 The contribution of the intensive and the extensive margins over time

Our third stylized fact describes the *average* contribution of the intensive and the extensive margin to variation in cohort-level employment and it does so conditional on the age of the cohort. This Appendix investigates how these contributions evolved over time (we thank an anonymous referee for this suggestion).

In particular, Figure 16 presents the contribution of the extensive and intensive margins for cohort-level employment variation of five year old firms over time. The figure suggests that, if anything, the intensive margin has gained importance during recent decades.

A.10 Empirical support for the demand channel

In addition to the aggregate empirical support for our model mechanism presented in the main text, this section of the Appendix gives further empirical support for the demand channel using detailed industry-level data.

Recall that the model predicts that changes in the composition of firms are (mainly) driven by demand shocks and the response to them by firms which are heterogeneous in their marketing elasticities. The latter is important as it determines the degree to which firms are sensitive to demand shocks. With low marketing elasticities of demand, the effect of demand shocks is muted and vice versa.

Therefore, our model would predict that subsectors with higher marketing elasticities (and thus higher advertising expenditure shares) would be more sensitive to demand shocks and in turn display stronger cohort-effects. To investigate this conjecture in the data, we link the QWI 4-digit industry data (described in Appendix A.7) with information from input-output tables of the Bureau of Economic Analysis (BEA). This enables us to observe a relationship between employment patterns by firm age and advertising expenditures in the given 4-digit industries. Using this data, we show that, consistent with the model, cohort effects are stronger in industries with higher expenditure shares on advertising and other forms of marketing.

To construct a measure for the strength of cohort effects we run, for each 4-digit sector, a linear regression of employment at age 4-5 on employment at age 0-1, lagged by four years (i.e. for the same cohort of firms). We express the data in log deviation from the mean (by state and 4-digit industry), and in our baseline include year fixed effects. The estimated coefficient on lagged employment parsimoniously captures the magnitude of cohort effects. As an alternative, we compute the autocorrelation at age 0-1 and age 4-5. Although results turn out to be similar, we prefer the aforementioned regression approach for the purpose of comparing sectors.[8]

[8] The autocorrelation captures both magnitude of cohort-effects and post-entry shocks. Suppose we compare two sectors, one of which is relatively sensitive to the demand shocks. According to the mechanism in our model, that sector will have relatively strong cohort effects, pushing up the autocorrelation. At the same time, however, the impact of post-entry demand shocks is relatively large in that sector, pushing down the autocorrelation. The net effect could in principle be ambiguous. By contrast, the regression-based measure captures only the former of these effects providing a more direct measure on the strength of cohort effects across sectors.

Table 8: Correlations between 4-digit measures of cohort effects and marketing shares

	expenditure share		output share	
	narrow	broad	narrow	broad
Baseline measure of cohort effects	0.265***	0.275***	0.215**	0.205**
Baseline cohort-effect measure, no year fixed effects	0.266***	0.277***	0.217**	0.209**
Alternative cohort-effect measure	0.205**	0.202**	0.188**	0.180**

Notes: "narrow" refers to marketing expenditures on "Advertising, public relations, and related services"and "broad" refers to expenditures on "Advertising, public relations, and related services"together with "Marketing research and all other miscellaneous professional, scientific, and technical services"of a given industry. Both values are taken from the input-output tables (at purchasers prices) of the BEA. These expenditures are expressed either as a share of total expenditures ("expenditure share") or as a share of industry output ("output share"). "Baseline measure" to regression-based measure of cohort-effects. "Alternative measure" refer to the autocorrelation (without removing year fixed effects). Two and three stars indicate significance at the 1, and 5 percent level, respectively.

.

We then correlate these measures for the strength of cohort effects with two measures of marketing expenditures taken from the input-output tables of the BEA. Specifically, we consider a "narrow" marketing measure which consist of "Advertising, public relations, and related services"(541800), as well as a "broad" measure which further adds "Marketing research and all other miscellaneous professional, scientific, and technical services"(541800 + 5419A0). These values we express as shares of either total expenditures or output in the given industry.

Table 8 presents the correlations of our cohort-effect measures and the two marketing shares. The relations are positive and statistically significant. This result is robust to removing the time fixed effect and to considering the alternative, autocorrelation-based measure. Thus, subsectors with relatively high marketing expenditure shares tend to have stronger cohort effects. A similar exercise for establishments, conducted using micro data from the Longitudinal Business Database, can be found in Moreira (2015), who comes to the same conclusions.

The evidence from the QWI supports is in line with the mechanism in the model, which predicts that demand shocks have less of an impact when marketing elasticities of demand, and hence advertising expenditure share, are lower. We corroborated this point by considering an alternative calibration in which the marketing elasticity of demand parameter, μ_i, is increased uniformly for all firm types, increasing the aggregate marketing expenditure share. We then computed the above measures for the strength of cohort effects in the model and, as expected, found that cohort effects are stronger in the calibration with a higher marketing expenditure share.

B Model details

This appendix presents supplemental model derivations (B.1), a formal definition of the equilibrium (B.2), and a formal discussion of the endogenous composition effects based on a simplified model (B.3).

B.1 Model derivations

The household's optimization problem. After substituting out C_t, the household's optimization problem can be expressed as:

$$\begin{aligned}
&\max_{N_t,\{c_{j,t}\}_{j\in\Omega_t}} \ln\left(\left(\int_{j\in\Omega_t}[\kappa_j(s_{j,t})]^{\frac{1}{\eta}}\, c_{j,t}^{\frac{\eta-1}{\eta}}\,dj\right)^{\frac{\eta}{\eta-1}}\right)-\nu Z_t N_t,\\
&s.t.\\
&\int_{j\in\Omega_t} p_{j,t}c_{j,t}dj = P_tW_tN_t+\Pi_t.
\end{aligned}$$

The first-order condition with respect to N_t is given by:

$$\nu Z_t = \lambda_t P_t W_t,$$

where λ_t is the Lagrange multiplier on the budget constraint. The first-order condition for $c_{j,t}$ reads:

$$C_t^{\frac{1-\eta}{\eta}}\left(\kappa_j(s_{j,t})\right)^{\frac{1}{\eta}} c_{j,t}^{-\frac{1}{\eta}} = p_{j,t}\lambda_t,$$

where we have used the definition of C_t. Rewriting this condition gives an expression for $c_{j,t}$:

$$c_{j,t} = p_{j,t}^{-\eta}\lambda_t^{-\eta}C_t^{1-\eta}\kappa_j(s_{j,t}).$$

Next, substitute out $c_{j,t}$ in the budget constraint to obtain:

$$\lambda_t^{-\eta}C_t^{1-\eta}\int_{j\in\Omega_t} p_{j,t}^{1-\eta}\kappa_j(s_{j,t})dj = P_tW_tN_t+\Pi_t,$$

or, equivalently,

$$\lambda_t^{-\eta} = \frac{P_tW_tN_t+\Pi_t}{C_t^{1-\eta}P_t^{1-\eta}},$$

where we used the aggregate price index, i.e. $P_t \equiv \left(\int_{j\in\Omega_t}\kappa_j(s_{j,t})p_{j,t}^{1-\eta}dj\right)^{\frac{1}{1-\eta}}$, which is defined such that $P_tC_t = \int_{j\in\Omega_t}p_{j,t}c_{j,t}dj$. Plugging this expression for $\lambda_t^{-\eta}$ back into the first-order condition for $c_{j,t}$ gives:

$$c_{j,t} = \left(\frac{p_{j,t}}{P_t}\right)^{-\eta}\kappa_j(s_{j,t})\left(W_tN_t+\Pi_t/P_t\right).$$

Using the budget constraint it now follows that the household's consumption demand for good j equals:

$$c_{j,t} = \left(\frac{p_{j,t}}{P_t}\right)^{-\eta} \kappa_j(s_{j,t})C_t.$$

Note further that:

$$\lambda_t^{-\eta} = \frac{P_t C_t}{C_t^{1-\eta} P_t^{1-\eta}},$$

from which it follows that $\lambda_t = \frac{1}{P_t C_t}$. Substituting out gives λ_t in the first-order condition for N_t gives:

$$\nu Z_t = \frac{W_t}{C_t}.$$

To verify that P_t is indeed the price index associated with the household's consumption decisions, note that:

$$\begin{aligned}
C_t &= \left(\int_{j\in\Omega_t} [\kappa_j(s_{j,t})]^{\frac{1}{\eta}} c_{j,t}^{\frac{\eta-1}{\eta}} dj\right)^{\frac{\eta}{\eta-1}}, \\
&= (W_t N_t + \Pi_t/P_t)\, P_t^{-\eta}\left(\int_{j\in\Omega_t} \kappa_j(s_{j,t}) p_{j,t}^{1-\eta} dj\right)^{\frac{\eta}{\eta-1}}, \\
&= (W_t N_t + \Pi_t/P_t)\, P_t^{-\eta} P_t^{\eta}, \\
&= \frac{\int_{j\in\Omega_t} p_{j,t} c_{j,t} dj}{P_t}.
\end{aligned}$$

Demand constraint. We now show how the households' first-order condition for $c_{j,t}$ leads to the firms' demand constraint. First, note that the amount of variety j used per unit of the aggregate consumption good is given by $\frac{c_{j,t}}{C_t} = \kappa_j(s_{j,t})\left(\frac{p_{j,t}}{P_t}\right)^{-\eta}$. Each entry attempt requires X_t units of the aggregate consumption bundle, and hence an amount $X_t\kappa_j(s_{j,t})\left(\frac{p_{j,t}}{P_t}\right)^{-\eta}$ of goods variety j. The total amount of variety j used for entry purposes, denoted $x_{j,t}$, is therefore given by $x_{j,t} = \sum_{i=1}^{I} e_{i,t} X_t \kappa_j(s_{j,t})\left(\frac{p_{j,t}}{P_t}\right)^{-\eta}$, where $e_{i,t}$ is the total number of entry attempts. Total demand for good j is thus given by:

$$\begin{aligned}
y_{j,t} &= c_{j,t} + x_{j,t}, \\
&= \kappa_j(s_{j,t})\left(\frac{p_{j,t}}{P_t}\right)^{-\eta} Y_t,
\end{aligned}$$

where the second equality follows from the resource constraint.

The firms' optimization problem. After substituting out $n_{j,t}^G = \kappa_j(s_{j,t}) \left(\frac{p_{j,t}}{P_t}\right)^{-\eta} Y_t/A_t$ and $n_{j,t}^M = \zeta(g_{j,t})$, the firms' optimization problem can be expressed as:

$$V_{j,a}(s_{j,t-1}, \mathcal{F}_t) = \max_{g_{j,t}, p_{j,t}, s_{j,t}} \left[\begin{array}{c} \kappa_j(s_{j,t}) \left(\frac{p_{j,t}}{P_t}\right)^{1-\eta} Y_t - \left(\kappa_j(s_{j,t}) \left(\frac{p_{j,t}}{P_t}\right)^{-\eta} Y_t/A_t + \zeta(g_{j,t})\right) W_t \\ + (1-\rho_a)\, \mathbb{E}_t \beta \frac{C_t}{C_{t+1}} V_{j,a+1}(s_{j,t}, \mathcal{F}_{t+1}) \end{array} \right]$$

$s.t.$

$s_{j,t} = s_{j,t-1} + Q_t g_{j,t}.$

The first-order condition for $p_{j,t}$ can be written as:

$$(1-\eta)\, p_{j,t}^{-\eta} P_t^{\eta-1} + \eta p_{j,t}^{-\eta-1} P_t^{\eta} W_t/A_t = 0,$$

or

$$p_{j,t} = \frac{\eta}{\eta-1} P_t W_t/A_t.$$

Note that all firms set the same price. Using the above equation to substitute out $p_{j,t}$, the firms' problem can be simplified to:

$$V_{j,a}(s_{j,t-1}, \mathcal{F}_t) = \max_{s_{j,t}, g_{j,t}} \left[\begin{array}{c} \kappa_j(s_{j,t}) \left(\frac{\eta}{\eta-1} W_t/A_t\right)^{1-\eta} Y_t - \left(\kappa_j(s_{j,t}) \left(\frac{\eta}{\eta-1} W_t/A_t\right)_t^{-\eta} Y_t/A_t + \zeta(g_{j,t})\right) W_t \\ + (1-\rho_a)\, \mathbb{E}_t \beta \frac{C_t}{C_{t+1}} V_{j,a+1}(s_{j,t}, \mathcal{F}_{t+1}) \end{array} \right];$$

$s.t.$

$s_{j,t} = s_{j,t-1} + Q_t g_{j,t},$

or

$$V_{j,a}(s_{j,t-1}, \mathcal{F}_t) = \max_{s_{j,t}} \left[\begin{array}{c} \kappa_j(s_{j,t}) Y_t (W_t/A_t)^{1-\eta} \left(\frac{\eta}{\eta-1}\right)^{-\eta} \frac{1}{\eta-1} - \zeta\left(\frac{s_{j,t}-s_{j,t-1}}{Q_t}\right) W_t \\ + (1-\rho_a)\, \mathbb{E}_t \beta \frac{C_t}{C_{t+1}} V_{j,a+1}(s_{j,t}, \mathcal{F}_{t+1}) \end{array} \right],$$

where we used that

$$\begin{aligned} & \kappa_j(s_{j,t}) \left(\frac{\eta}{\eta-1} W_t/A_t\right)^{1-\eta} Y_t - \kappa_j(s_{j,t}) \left(\frac{\eta}{\eta-1} W_t/A_t\right)^{-\eta} Y_t W_t/A_t, \\ = \; & \kappa_j(s_{j,t}) \frac{\eta}{\eta-1} W_t/A_t \left(\frac{\eta}{\eta-1} W_t/A_t\right)^{-\eta} Y_t - \kappa_j(s_{j,t}) \left(\frac{\eta}{\eta-1} W_t/A_t\right)^{-\eta} Y_t W_t/A_t, \\ = \; & \kappa_j(s_{j,t}) W_t/A_t \left(\frac{\eta}{\eta-1} W_t/A_t\right)^{-\eta} Y_t \left(\frac{\eta}{\eta-1} - 1\right), \\ = \; & \kappa_j(s_{j,t}) Y_t (W_t/A_t)^{1-\eta} \left(\frac{\eta}{\eta-1}\right)^{-\eta} \frac{1}{\eta-1}. \end{aligned}$$

The first-order condition of the above problem is:

$$\zeta'\left(\frac{s_{j,t}-s_{j,t-1}}{Q_t}\right)W_t/Q_t=\kappa_j'(s_{j,t})Y_t\left(W_t/A_t\right)^{1-\eta}\left(\frac{\eta}{\eta-1}\right)^{-\eta}\frac{1}{\eta-1}$$
$$+\left(1-\rho_a\right)\mathbb{E}_t\beta\frac{C_t}{C_{t+1}}\zeta'\left(\frac{s_{j,t+1}-s_{j,t}}{Q_{t+1}}\right)W_{t+1}/Q_{t+1}.$$

The above equation simplifies to Equation (4) in the main text, making use of the fact that

$$\begin{aligned}
\epsilon_{j,t}^{\kappa,s}\frac{n_{j,t}^G}{s_{j,t}}\frac{W_t}{\eta-1} &= \frac{\kappa_j'\left(s_{j,t}\right)s_{j,t}}{\kappa_j\left(s_{j,t}\right)}\frac{y_{j,t}}{A_t s_{j,t}}\frac{W_t}{\eta-1}\\
&= \kappa_j'\left(s_{j,t}\right)\left(\frac{p_{j,t}}{P_t}\right)^{-\eta}Y_t\frac{W_t}{A_t}\frac{1}{\eta-1}\\
&= \kappa_j'\left(s_{j,t}\right)Y_t\left(\frac{W_t}{A_t}\right)^{1-\eta}\left(\frac{\eta}{\eta-1}\right)^{-\eta}\frac{1}{\eta-1}
\end{aligned}$$

where the first equality uses the production function, $y_{j,t}=A_t n_{j,t}^G$.

Aggregate resource constraint. Next, we derive the aggregate resource constraint. First note that aggregate net profits, in real terms, are given by total sales minus the total wage bill minus total entry costs:

$$\Pi_t/P_t=\int_{j\in\Omega_t}\left(\frac{p_{j,t}}{P_t}y_{j,t}-W_t\left(n_{j,t}^G+n_{j,t}^M\right)\right)dj-\sum_{i=1}^{I}e_{i,t}X_t.$$

Recall that the labor market clearing condition is given as:

$$N_t=\int_{j\in\Omega_t}\left(n_{j,t}^G+n_{j,t}^M\right)dj.$$

Plugging these two expressions into the household's budget constraint gives the aggregate resource constraint:

$$
\begin{aligned}
C_t &= W_t N_t + \Pi_t / P_t, \\
&= \int_{j\in\Omega_t} W_t \left(n^G_{j,t} + n^M_{j,t}\right) dj + \int_{j\in\Omega_t} \left(\frac{p_{j,t}}{P_t} y_{j,t} - W_t \left(n^G_{j,t} + n^M_{j,t}\right)\right) dj - \sum_{i=1}^{I} e_{i,t} X_t, \\
&= \int_{j\in\Omega_t} \frac{p_{j,t}}{P_t} y_{j,t} dj - \sum_{i=1}^{I} e_{i,t} X_t, \\
&= \sum_i \sum_a \frac{p_{j,t}}{P_t} m_{i,a,t} y_{i,a,t} - \sum_{i=1}^{I} e_{i,t} X_t, \\
&= Y_t - \sum_{i=1}^{I} e_{i,t} X_t.
\end{aligned}
$$

Note further that:

$$
\begin{aligned}
Y_t &= \int_{j\in\Omega_t} \frac{p_{j,t}}{P_t} y_{j,t} dj, \\
&= \int_{j\in\Omega_t} \kappa_j(s_{j,t}) \left(\frac{p_{j,t}}{P_t}\right)^{1-\eta} Y_t dj, \\
&= \int_{j\in\Omega_t} \kappa_j(s_{j,t}) dj \left(\frac{\eta}{\eta-1} W_t/A_t\right)^{1-\eta} Y_t,
\end{aligned}
$$

which gives the following restriction:

$$
\left(\frac{\eta}{\eta-1} W_t/A_t\right)^{\eta-1} = \int_{j\in\Omega_t} \kappa_j(s_{j,t}),
$$

which in turn delivers the "variety effect" equation stated in the main text.

B.2 Equilibrium definition

For the sake of parsimony, we substitute out several variables before defining the equilibrium. As mentioned in the main text, we also replace firm index j by age-type indices i, a. The equations for firm values ($V_{i,a,t}$), and the first-order conditions for marketing capital ($s_{i,a,t}$) can respectively, be written as:

$$
\begin{aligned}
V_{i,a,t} = \quad & \kappa_j(s_{j,t}) \left(W_t/A_t\right)^{1-\eta} \left(\frac{\eta}{\eta-1}\right)^{-\eta} \frac{1}{\eta-1} Y_t - \zeta\left(\frac{s_{i,a,t} - s_{i,a-1,t-1}}{Q_t}\right) W_t \\
+ \quad & (1-\rho_a)\, \mathbb{E}_t \beta \frac{C_t}{C_{t+1}} V_{i,a+1,t+1},
\end{aligned}
$$

$$\zeta'\left(\frac{s_{i,a,t}-s_{i,a-1,t-1}}{Q_t}\right)\frac{1}{Q_t} = \kappa_j'(s_{j,t})Y_t\left(W_t/A_t\right)^{1-\eta}\left(\frac{\eta}{\eta-1}\right)^{-\eta}\frac{1}{\eta-1} + \left(1-\rho_a\right)\mathbb{E}_t\beta\frac{C_t}{C_{t+1}}\zeta'\left(\frac{s_{i,a+1,t+1}-s_{i,a,t}}{Q_{t+1}}\right)\frac{1}{Q_{t+1}}\frac{W_{t+1}}{W_t},$$

for $a\in\mathbb{N}_{>0}$ and $i=1,2,..,I$. The free-entry condition and accounting equation for the masses of firms are:

$$\begin{aligned}
m_{i,0,t} &= \psi_i\left(\frac{V_{i,0,t}}{X_t}\right)^{\frac{1-\phi}{\phi}}\\
m_{i,a,t} &= \left(1-\rho_{a-1}\right)m_{i,a-1,t-1} \text{ for } a\in\mathbb{N}_{>0} \text{ and } i=1,2,..,I.
\end{aligned}$$

The labor market clearing condition, the aggregate resource constraint, the aggregate output definition, and the first-order condition for labor can be expressed as:

$$\begin{aligned}
&N_t=\sum_{i=1}^{I}\sum_{a\in\mathbb{N}}m_{i,a,t}\left(\kappa_j(s_{j,t})\left(\frac{\eta}{\eta-1}W_t/A_t\right)^{-\eta}Y_t/A_t+\zeta\left(\frac{s_{i,a+1,t+1}-s_{i,a,t}}{Q_t}\right)\right),\\
&C_t+\sum_{i=1}^{I}\left(m_{i,0,t}/\psi_i^{\phi}\right)^{\frac{1}{1-\phi}}X_t=Y_t,\\
&1=\sum_{i=1}^{I}\sum_{a\in\mathbb{N}}m_{i,a,t}\kappa_j(s_{j,t})\left(\frac{\eta}{\eta-1}W_t/A_t\right)^{-\eta},\\
&\nu=W_t/C_t,
\end{aligned}$$

where the matching function has been used to substitute out the measure of startup attempts per type, $e_{i,t}=\left(m_{i,0,t}/\psi_i^{\phi}\right)^{\frac{1}{1-\phi}}$.

Definition (recursive equilibrium).
A recursive competitive equilibrium is defined by laws of motion for
- the representative household's labor supply, $N\left(\mathcal{F}_t\right)$*, and the consumption bundle* $C(\mathcal{F}_t)$*, and the aggregate output bundle* $Y(\mathcal{F}_t)$*,*
- the wage $W(\mathcal{F}_t)$*,*
- firm value functions $V_{i,a}\left(s_{i,a-1,t-1},\mathcal{F}_t\right)$ *and consumer bases* $s_{i,a}\left(s_{i,a-1,t-1},\mathcal{F}_t\right)$*, for* $i=1,2,..,I$ *and* $a\in N$*,*
- the measure of operating firms $m_{i,a}\left(\mathcal{F}_t\right)$*, for* $i=1,2,..,I$ *and* $a\in N$*,*
that solve the above system of equations for each period t*, for each type* $i=1,2,..,I$*, and each age* $a\in N$*, given the processes for the exogenous shock variables* A_t,Q_t,X_t *and* Z_t *and with the aggregate state the aggregate state being given by*
$\mathcal{F}_t=\left[A_t,Q_t,X_t,Z_t,\left\{m_{i,a-1,t-1},s_{i,a-1,t-1}\right\}_{i=1,..I,\ a\in\mathbb{N}>0}\right].$

B.3 Composition effects in a simplified model

This appendix formally shows how a negative demand shock (Q_t) decreases the relative profitability of "mass" firms in a simplified version of our model. In particular, we assume that (i) the entrants exit with certainty after one year, and (ii) the advertising cost is linear with $\zeta = 1$, i.e. $n^M = g$. We proceed by first showing that firm profits are more sensitive to demand shocks when the firm devotes a relative large fraction of expenditures to marketing investments. Next, we show that it is precisely "mass" firms which have high marketing cost expenditures.

Consider the steady-state real profits of an entrant in the first year, which can be expressed as:

$$
\begin{aligned}
\Pi_j/P &= n_j^G \left(\frac{p_{j,t}}{P_t} - W \right) - n_j^M W, \\
&= n_j^G \frac{W}{\eta - 1} - n_j^M W, \\
&= n_j^G \frac{W}{\eta - 1} - \frac{s_j}{Q} W,
\end{aligned}
$$

where the last equality uses the marketing capital accumulation equation.[9] It now follows that the elasticity of real profits with respect to Q, starting from the steady state and holding the firm's total output level constant, is given by:

$$
\frac{\partial \ln (\Pi_j/P)}{\partial \ln Q} = \frac{s_j W}{Q \Pi_j/P} = \frac{n_j^M W}{\Pi_j/P}.
$$

Thus, the steady-state elasticity is equal to the steady state level of marketing expenditures relative to profits.

Next we illustrate that it is precisely "mass" firms which optimally choose higher marketing expenditure shares. Towards this end, divide real profits by the labor costs of marketing investment to obtain:

$$
\begin{aligned}
\frac{\Pi_j/P}{n_j^M W} &= \frac{n_j^G}{n_j^M W} \frac{W}{\eta - 1} - 1, \\
&= \frac{1}{\epsilon_j} - 1,
\end{aligned}
$$

where ϵ_j is the firm's marketing elasticity of demand. The second equality uses the fact that the optimal marketing condition in this simplified model is given by $1 = \frac{\epsilon_j}{\eta - 1} \frac{n_j^G}{n_j^M}$. This result has already been known since Dorfman and Steiner (1954), who also showed that the optimal expenditure share of advertising (marketing) is proportional to the advertising (marketing) elasticity of demand. The above makes clear that the expenditure share

[9] Recall that in the steady state $A = 1$.

of marketing investment in real profits, $n_j^M W/(\Pi_j/P)$, is increasing in the marketing elasticity of demand, ϵ_j.[10] Therefore, firms with higher marketing elasticities of demand optimally invest relatively heavily into marketing which, in turn, renders them more sensitive to demand fluctuations.

C Marketing elasticities in the data

When parameterizing the model, we use information on average size by age to pin down the values of the heterogeneous marketing elasticities of demand. Section 4 describes this procedure in detail. Table 1 and Figure 5 in the main text make clear that a substantial amount heterogeneity in marketing elasticities of demand is needed to match the data well. The purpose of this Appendix is to assess this heterogeneity in the light of external evidence.

In the empirical Marketing literature, there are many studies which estimate the effects marketing investments on product demand. Specifically, many studies estimate the elasticity of demand with respect to advertising expenditures. A meta-analysis of 56 of these studies, estimating over 1100 elasticities for different products, is provided by Sethuraman, Tellis, and Briesch (2011). Their analysis includes studies conducted in various time-periods, using data for firms (or products) in various industries (durables, non-durables, food, non-food, pharmaceuticals,services etc.) and in various phases of their respective life-cycles (growing or mature).

Sethuraman, Tellis, and Briesch (2011) report two overall findings that are important in the light of our model. First, they find a very large degree of heterogeneity in advertising elasticities across products, with estimates of the short-term elasticity ranging between -0.35 to 1.8. Second, they find that the advertising elasticity of demand declines over the product life cycle.

To relate our model to these empirical estimates, we need to compute the model-implied elasticities of demand with respect to changes in the *flow* of marketing expenditures.[11] Using the firms' demand constraint, one can express the short-term (flow) marketing elasticity of demand (the percentage increase in demand for a 1% increase in marketing expenditures) as:

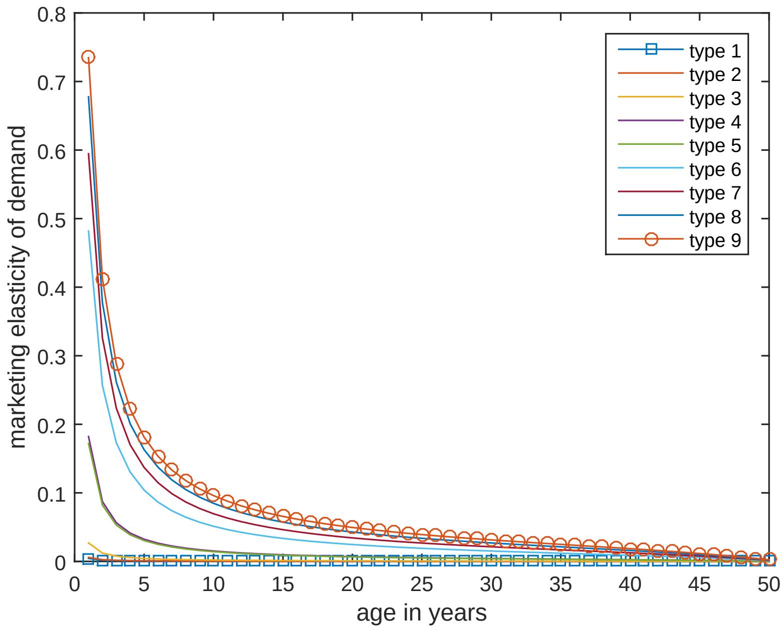
$$\epsilon_{j,t}^{y,n^M} \equiv \frac{\partial y_{j,t}}{\partial n_{j,t}^M}\frac{n_{j,t}^M}{y_{j,t}} = \frac{\mu_j}{2}\frac{g_{j,t}}{s_{j,t}}. \tag{2}$$

We compute the steady-state elasticity for each age/type category in the model. The values are between between $6.7E^{-6}$ and 0.73, falling well within the empirical range

[10] Recall that $\eta > 1$.

[11] Note that in the model μ_j is the elasticity of demand with respect to the *stock* of consumer capital, whereas the empirical studies estimate the elasticity with respect to the flow.

Figure 17: Marketing elasticities of demand by firm type and age



Notes: Marketing elasticities of demand predicted by the model and calculated according to (2).

reported in Sethuraman, Tellis, and Briesch (2011). Further, the model implies that, as in the data, demand elasticities decline over the life cycle, as shown in Figure 17, which plots the flow elasticity by type and age. We conclude that the marketing elasticities in the parameterized model are in line with the empirical evidence.

D Computation and Estimation

This section discusses various issues regarding the computation of the equilibrium of the model and the estimation of the model. First, we discuss the computation of the steady-state equilibrium without aggregate uncertainty (D.1). Next, we discuss how we solve for the dynamic equilibrium with aggregate shocks (D.2). Next we discuss how we measure variables in the data (D.3). Finally, we provide further background on the estimation of the model (D.4).

To economize on notation, let us express the model compactly as:

$$\mathbb{E}_t f\left(y_{t+1}, y_t, x_{t+1}, x_t; \Upsilon, \eta\right) = 0,$$

where x_t is a vector containing the state variables (all variables in $\mathcal{F}_t$) and y_t is a vector containing the non-preditermined variables, Υ is a vector containing all parameters of the model and η is a scalar parameter pre-multiplying the covariance matrix of the shock innovations, as in Schmitt-Grohé and Uribe (2004). Importantly, the above is system of a finite number of expectational difference equations.

D.1 Solving for the steady state without aggregate uncertainty

We first solve for the equilibrium of a version of the model without aggregate uncertainty. That is, we find vectors $\overline{y}$ and $\overline{x}$ that solve $f(\overline{y}, \overline{y}, \overline{x}, \overline{x}; \Upsilon, 0) = 0$. As described in the main text, we calibrate various parameters to match long-run targets. The calibration procedure has the following steps:

1. given values for firm type parameters (μ_i and $\overline{\kappa}_i$), the price elasticity of substitution (η) and the wage rate (w), we can calculate growth paths of firm-level consumer capital (s) and thus of employment. We choose values of $\overline{\kappa}_i$ such that the model delivers average firm size of 21 to 25 year old firms as observed in the BDS size brackets. Furthermore, we choose μ_i such that the model matches average firm size by age observed in the BDS.

2. given the time-paths of consumer capital from (1.), we can also compute firm values. These together with the value of the entry cost ($\overline{X}$) pin down the startup success probabilities in each firm type through the free entry conditions.

3. given the startup probabilities from (2.) and the elasticity in the entry matching function (ϕ), we can back out the mass of business opportunities in each firm type by targeting the share of firms in each size bracket of 21 to 25 year old firms observed in the data.

4. given the mass of business opportunities from (3.), the probabilities of starting up from (2.), and the age-dependent death rates (taken from the BDS), one can calculate the mass of firms in each age-type cell.

5. given firm masses from (4.), employment choices from (1.), one can calculate all the aggregates (total employment, output, consumption etc.) and back out the disutility of supplying labor from the households first-order condition.

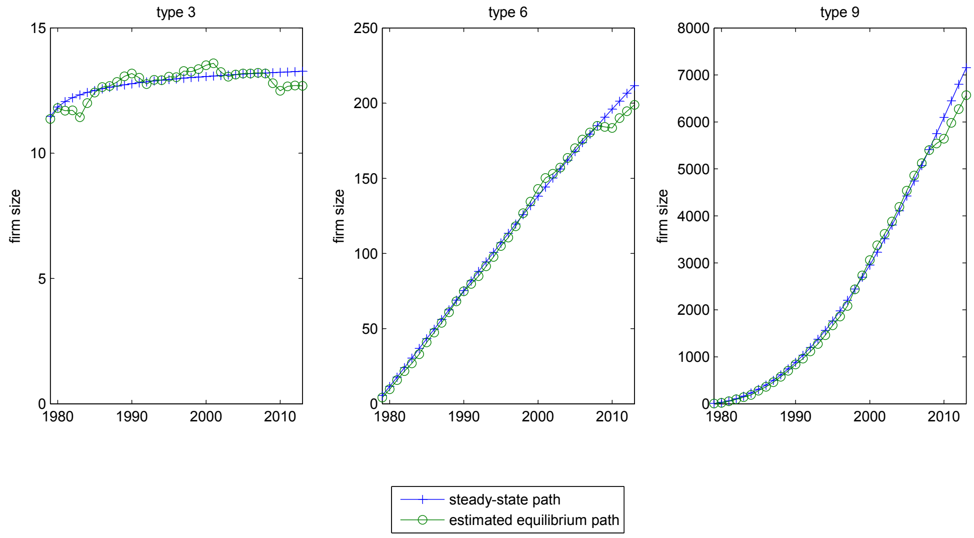
D.2 Solving for the equilibrium with aggregate uncertainty

Next, we solve for the dynamic equilibrium using first-order perturbation around the deterministic equilibrium found in the previous step. The first-order approximated solutions, denoted by hats, have the following form:

$$\begin{aligned} \widehat{x}_{t+1} &= \overline{x} + \Theta\left(\widehat{x}_t - \overline{x}\right), \\ \widehat{y}_{t+1} &= \overline{y} + \Phi\left(\widehat{x}_t - \overline{x}\right), \end{aligned}$$

where Θ and Φ are matrices containing the coefficients obtained from the approximation. The perturbation procedure is standard and carried out in one step.

Figure 18: Employment levels: types 3, 6 and 9



Notes: Green line represents the steady-state path. Blue line is the estimated equilibrium path for firms born in 1979.

An advantage of perturbation methods is that the computational speed is relatively high and many state variables can be handled. An important prerequisite for perturbations to be accurate, however, is that deviations from the steady-state are not too large. For firm dynamics models like ours this may seem problematic because differences in employment levels across firms may be very large. Our approach, however, overcomes this problem since the steady state we perturb around contains the entire *growth paths* of firms. These growth paths, captured by the constants in the above equations, are themselves non-linear functions of age and type.

Hence, the fact that most newborn firms starts off much below their eventual sizes does not involve large accuracy losses since the same is true for the steady-state sizes of newborn firms. Similarly, the fact that the equilibrium features various firm types with very different optimal sizes does not reduce accuracy since we perturb around the growth path for each individual firm type. To illustrate these points, the Figure (18) plots a simulated employment levels of firms of various types. The figure also plots the steady-state path in the absence of aggregate shocks, the center of the first-order approximation. At each point in the simulation, the employment level of the firm is close to the steady-state path used for the approximations, even though differences across type- and age-groups are very large.

D.3 Measurement

In order to estimate the model, we need to relate model variables to observables in the data. The number of firms, average firm size, and employment within a cohort of age a

in period t are given, respectively, by:

$$\begin{aligned} M_{a,t} &\equiv \sum_{i=1}^{I} m_{i,a,t}, \\ S_{a,t} &\equiv \sum_{i=1}^{I} m_{i,a,t}\left(n_{j,t}^{G}+n_{j,t}^{M}\right)/M_{a,t}, \\ N_{a,t} &= M_{a,t}S_{a,t}. \end{aligned}$$

Aggregate employment is measured as:

$$N_t = \sum_{a=1}^{K} N_{a,t}.$$

We define observed real GDP as:

$$GDP_t \equiv A_t \sum\nolimits_j n_{j,t}^{G}.$$

This is consistent with the way in which the Bureau of Economic Analysis (BEA) constructs its time series for real GDP. Since 1996, the BEA computes GDP in year t, denoted GDP_t, relative to GDP in the previous year, as:

$$\frac{GDP_t}{GDP_{t-1}} = \sqrt{I_t^L I_t^P},$$

where I_t^L is a Laspeyres quantity index defined as $I_t^L \equiv \frac{\sum_j p_{j,t-1}y_{j,t}}{\sum_j p_{j,t-1}y_{j,t-1}}$, and I_t^P is a Paasche quantity index, defined as $I_t^P \equiv \frac{\sum_j p_{j,t}y_{j,t}}{\sum_j p_{j,t}y_{j,t-1}}$. Normalizing real GDP in a certain base year to 100, a time series I_t^L and I_t^P are then used by the BEA to construct a time series for real GDP.

In our model, all firms set the same price. From the above formulas it immediately follows that measured GDP obeys

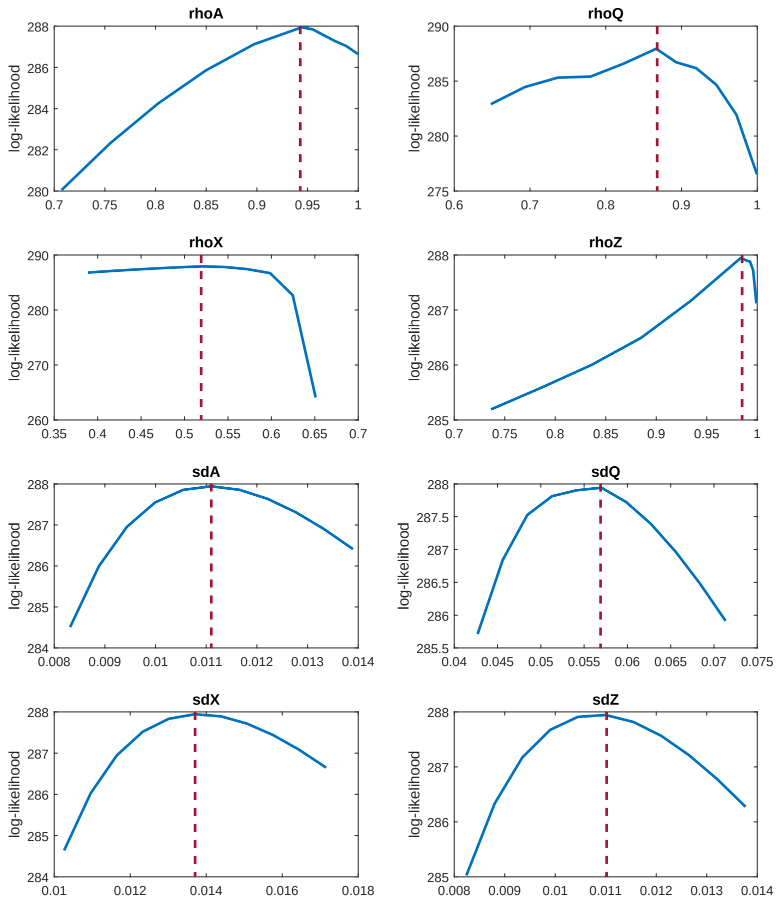
$$\frac{GDP_t}{GDP_{t-1}} = I_t^L = I_t^P = \frac{\sum_j y_{j,t}}{\sum_j y_{j,t-1}} = \frac{A_t \sum_j n_{j,t}^G}{A_t \sum_j n_{j,t-1}^G}.$$

Thus, defining observed real GDP in the model as $GDP_t = A_t \sum_j n_{j,t}^G$ is consistent with BEA methodology.

D.4 Estimation

Having solved the model for given parameter values we can compute the likelihood of the linearized model. To do so, the linearized model is set into state-space form and the parameters of the aggregate shock processes are estimated using Maximum Likelihood.

Figure 19: Partial Likelihoods



Notes: Likelihood values for deviations in parameter values

Using the Kalman filter enables one to conveniently characterize the likelihood function as well as obtain estimates of the underlying aggregate shocks, given the four observable time series. An important by-product of estimating the model is that we obtain model-predicted time series for all the variables in the model (using the Kalman smoother). This means that we obtain the entire time-varying distribution of firms (their masses, employment levels and firm values), which we can use for counterfactual analysis.

To corroborate the results of the numerical procedure used to find the likelihood, Figure 19 plots the value of the likelihood as a function of individual parameter values, each deviated in turn by fifty percent (with an upper bound of 1 in the case of the persistence parameters). The vertical lines denote our maximum likelihood estimates, which coincide with the top of the likelihood. The plot also gives a sense of how precisely the parameters are estimated. Comparing the four shocks in the model, the precision of their estimates appears similar. Moreover, the respective standard deviations appear relatively tightly estimated, compared to the persistence parameters. Note also that the likelihood of a version of the model in which the standard deviation of the composition shocks are 50 percent smaller is substantially lower than the value of the likelihood at the maximum.

E Model robustness exercises

This appendix presents exercises investigating the robustness of the model results. In particular, we discuss the importance of the assumption that exit rates are fixed over time (E.1), we provide a sensitivity analysis with respect to the degree of curvature in the convex cost of investing into the consumer base (E.2), the persistence of the exogenous shocks (E.3), the value of the maximum firm age (E.4) and finally also with respect to a positive consumer base depreciation rate (E.5).

For the sake of brevity, in all these robustness checks we summarize the model implications using two sets of results (one for cohort-level and one for aggregate implications). In particular, for each robustness check we first reproduce Figure 8 of the main text which shows the contributions of the state at birth and of all four shocks to cohort-level variation in employment and average firm size. Second, for each of the robustness checks we reproduce Figure 10 of the main text which shows the actual employment rate together with a counterfactual based only on variation in startup conditions.

E.1 Variation in firm exit rates

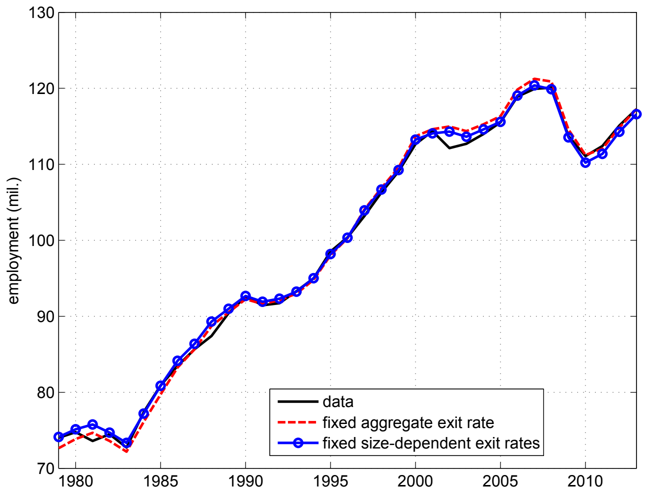
The model in the main text assumes constant, though age-dependent, firm exit rates. A concern could be that variation in exit rates is an important feature of the data responsible for a large part of variation in employment. The variance decomposition in the empirical section suggests that variation in exit rates accounts for only 3% of cohort-level employment variation on average among firms aged 1 to 5 years. Nevertheless, in this subsection we investigate the possible importance of fluctuations in exit rates further. The results suggest that incorporating variation in firm exit rates would imply only minor changes of our findings.

How important is firm exit for employment? To get a sense of how important can variation in firm exit rates be for employment we construct two counterfactual time series. The first tries to quantify how important is time-variation for the evolution of employment. The second goes a step further and acknowledges that different firm types might behave differently in terms of firm exit. Specifically, we construct a counterfactual aggregate employment time series according to

$$N_t^c = N_{t-1} + JC_t - (JD_t - JDd_t + JDd_t^c), \tag{3}$$

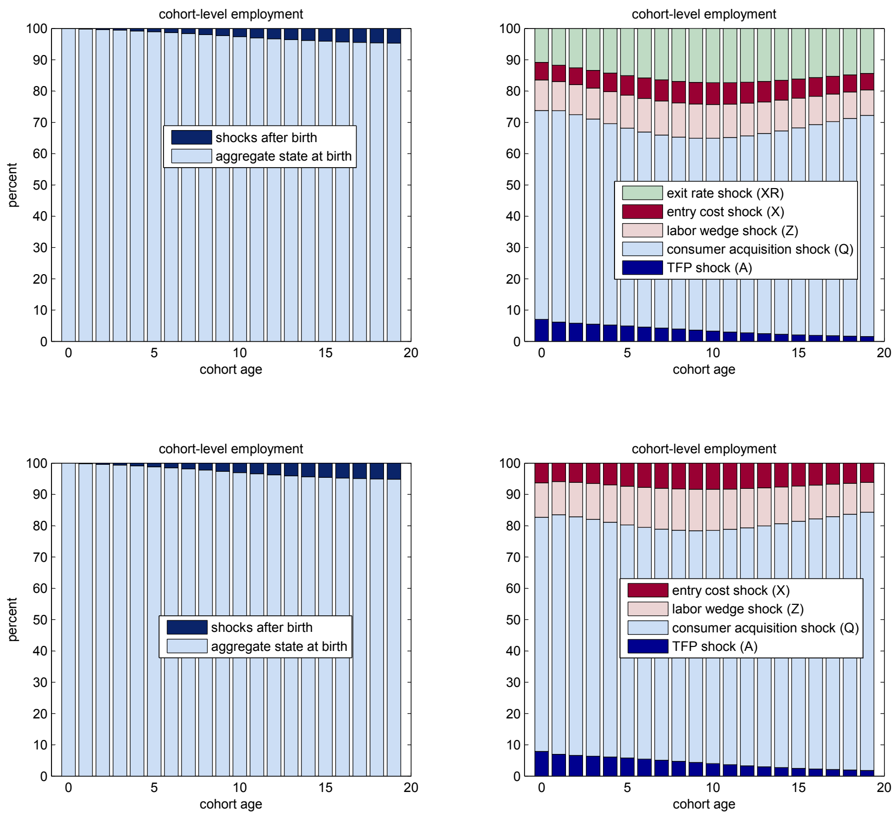
where a c in the superscript indicates a counterfactual, JC_t is gross job creation, JD_t is gross job destruction and JDd_t is gross job destruction due to firm exit in period t. The two counterfactual employment series we create differ in the way we construct JDd_t^c.

Figure 20: Employment levels: data and exit-based counterfactuals



Notes: The figure plots the data and two counterfactual aggregate employment levels. "fixed aggregate exit rate" is constructed by fixing the overall firm exit rate to the sample average. "Fixed size-dependent exit rates" is constructed by fixing the firm exit rates within each size bracket in the BDS and averaging over the sample.

Figure 21: Variance decomposition for a model with and without time-varying exit rates.



First, we assume that the number of jobs lost due to firm exit is fixed at its sample average, i.e. $JDd_t^{1,c} = 1/T \sum_{t=1}^{T} JDd_t$.

Second, we try to go a step further and acknowledge that firms of different types may behave differently in terms of firm exit. As a proxy for firm type, we use the BDS size brackets and construct the counterfactual number of lost jobs due to exit as

$$JDd_t^{2,c} = \sum_j N_{j,t} \frac{1}{T} \sum_{t=1}^{T} \frac{JDd_{j,t}}{N_{j,t}},$$

where j is an index representing the size brackets in the BDS data, $N_{j,t}$ is total employment in size bracket j, and $JDd_{j,t}$ is job destruction due to firm deaths in size bracket j. In other words, we first compute the sample averages of the job destruction rate due to firm exit within each type and multiply this by the (time-varying) levels of employment within each size bracket.

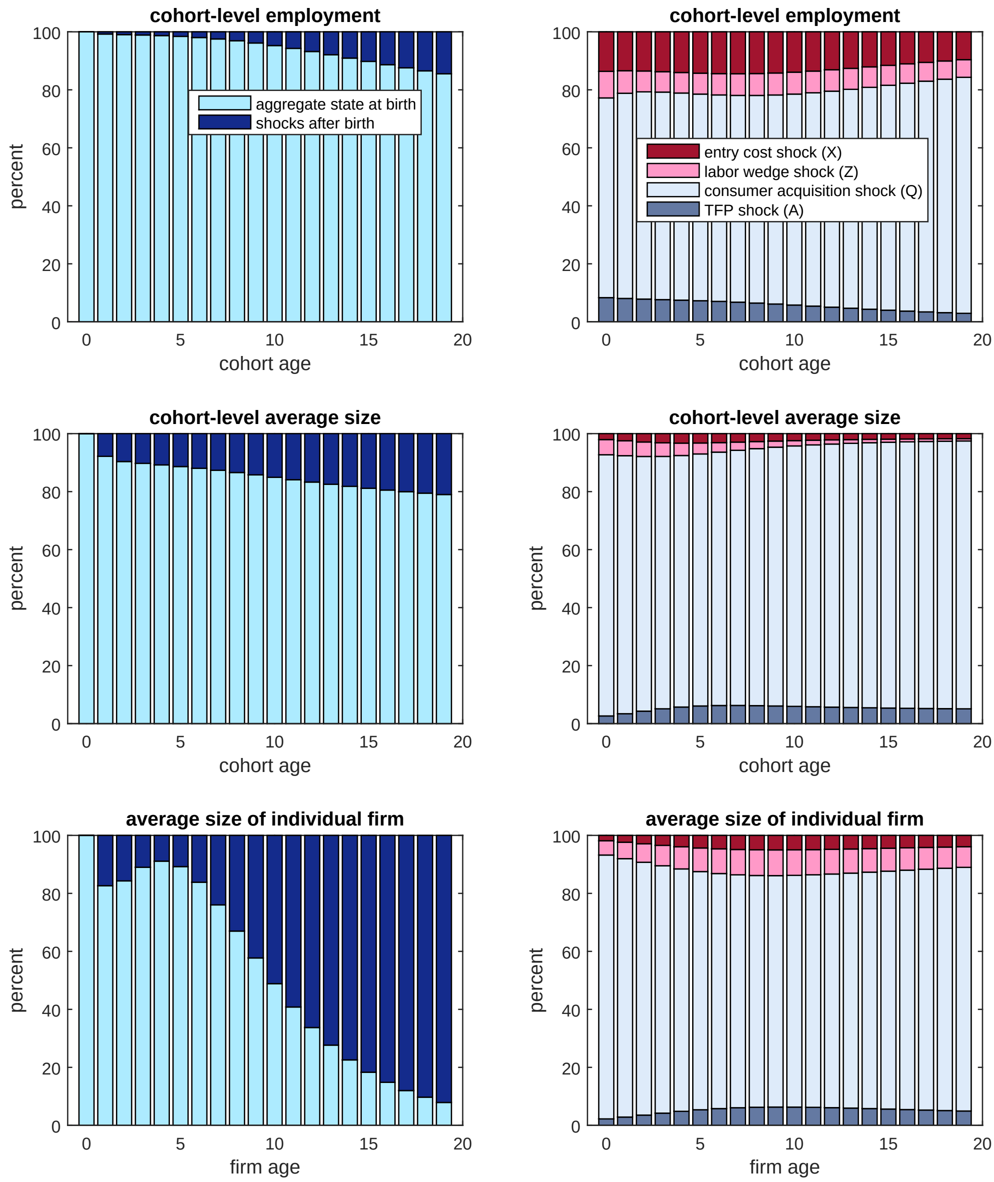
Figure 20 shows aggregate employment and the two counterfactual employment time series. The counterfactuals are very close to the data suggesting that not much information is lost by our model assumption that firm death rates (by age) are fixed over time and across types.

Incorporating time variation in firm exit in the model To investigate the importance of time variation in exit rates for our model results, we introduce a stochastic aggregate shock to the exit rate to the model. We assume the shock follows an AR(1) process with normally distributed innovations. We estimate the autocorrelation coefficient and standard deviation of the innovation from BDS data on exit rates. The estimated parameters are 0.4810 and 0.0519, respectively. After adding the shock to the model, we re-compute the variance decomposition for cohort-level employment. Figure 21 plots the result, together with the variance decomposition for our benchmark model. The results show that adding stochastic exit rate does not alter our conclusion on the importance on the state at birth. For younger firms, the initial state actually becomes more important relative to the benchmark.

E.2 Curvature in marketing costs

This Appendix conducts robustness exercises with respect to the degree of curvature in marketing adjustment costs (denoted here by ζ_1). The benchmark model assumes that these costs are convex in consumer capital investment $n^M = \frac{1}{\zeta_1} g^{\zeta_1}$, with $\zeta_1 = 2$. To check the robustness of the benchmark results, we consider two alternative parametrizations of the curvature of marketing adjustment costs. First, we assume that the curvature is closer to linear adjustment costs ($\zeta_1 = 1.5$) and second, we increase the curvature even further ($\zeta_1 = 3$). We recalibrate all other parameters such that the model matches all the targets as in the benchmark parametrization. Moreover, we re-estimate the shock processes implied by the data and these two alternative calibrations. Figures 22 and 25 depict the cohort-level implications of fluctuations in the composition of firms under the above two alternative calibrations. In both cases the state at birth accounts for the

Figure 22: Model variance decompositions: low marketing cost curvature

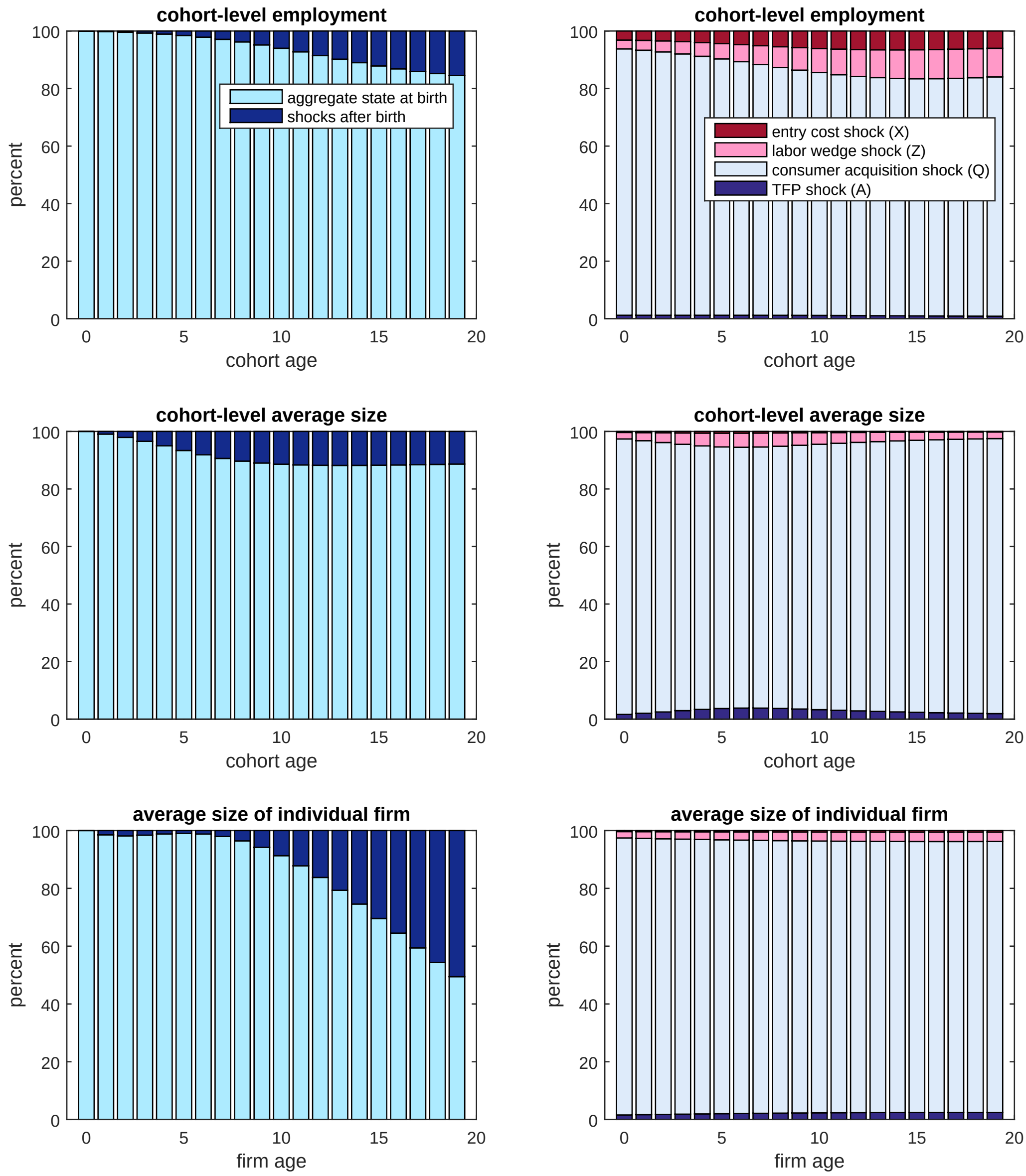


Notes: Contributions of the aggregate state at birth and post-entry shocks (left panels) and the contributions of the four aggregate shocks (right panels) to variation in cohort-level employment (top row), cohort-level average size (middle row) and individual-firm average size (bottom row). Results are based on a marketing adjustment costs curvature of $\zeta_1 = 1.5$.

majority of cohort-level fluctuations and the demand shock is the dominant driver of variation at the cohort-level as well as at the level of an individual firm.

Therefore, while all three calibrations deliver similar results in terms of the importance of composition effects for cohort-level and aggregate variables, our benchmark calibration performs somewhat better in other (untargeted) moments. In particular, the pattern of

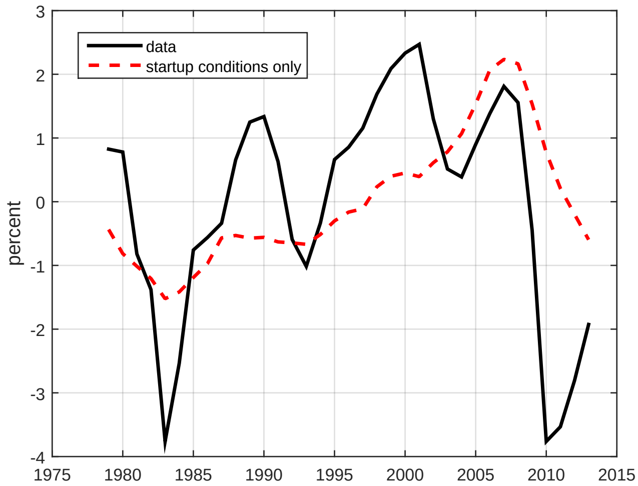
Figure 23: Model variance decompositions: high marketing cost curvature



Notes: Contributions of the aggregate state at birth and post-entry shocks (left panels) and the contributions of the four aggregate shocks (right panels) to variation in cohort-level employment (top row), cohort-level average size (middle row) and individual-firm average size (bottom row). Results are based on a marketing adjustment costs curvature of $\zeta_1 = 3$.

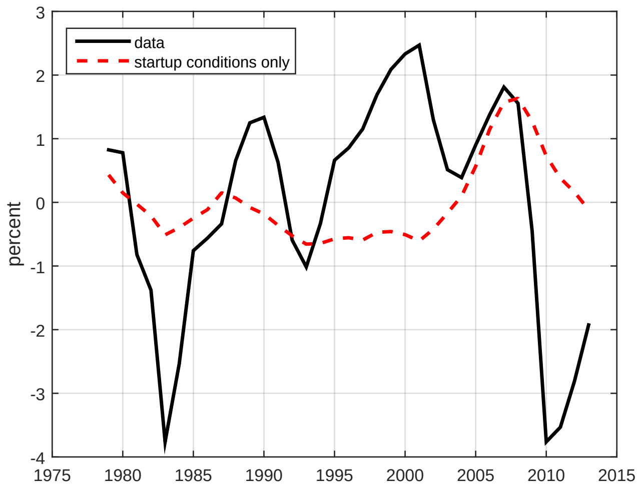
advertising-to-GDP, which crucially depends on the form of marketing adjustment costs, follows its empirical counterpart closer in the benchmark model compared to both alternative specifications. The correlation coefficient of advertising-to-GDP in the benchmark model and data is 0.50. For the two alternative specifications, this correlation becomes 0.29 and 0.32 for the case with high and low curvature, respectively.

Figure 24: Contribution of startup conditions to aggregate employment: low marketing cost curvature



Notes: Employment rate data and a model-based counterfactual employment rate based on the fixing the age/type firm sizes to their respective steady state values.Results are based on a marketing adjustment costs curvature of $\zeta_1 = 1.5$.

Figure 25: Contribution of startup conditions to aggregate employment: high marketing cost curvature

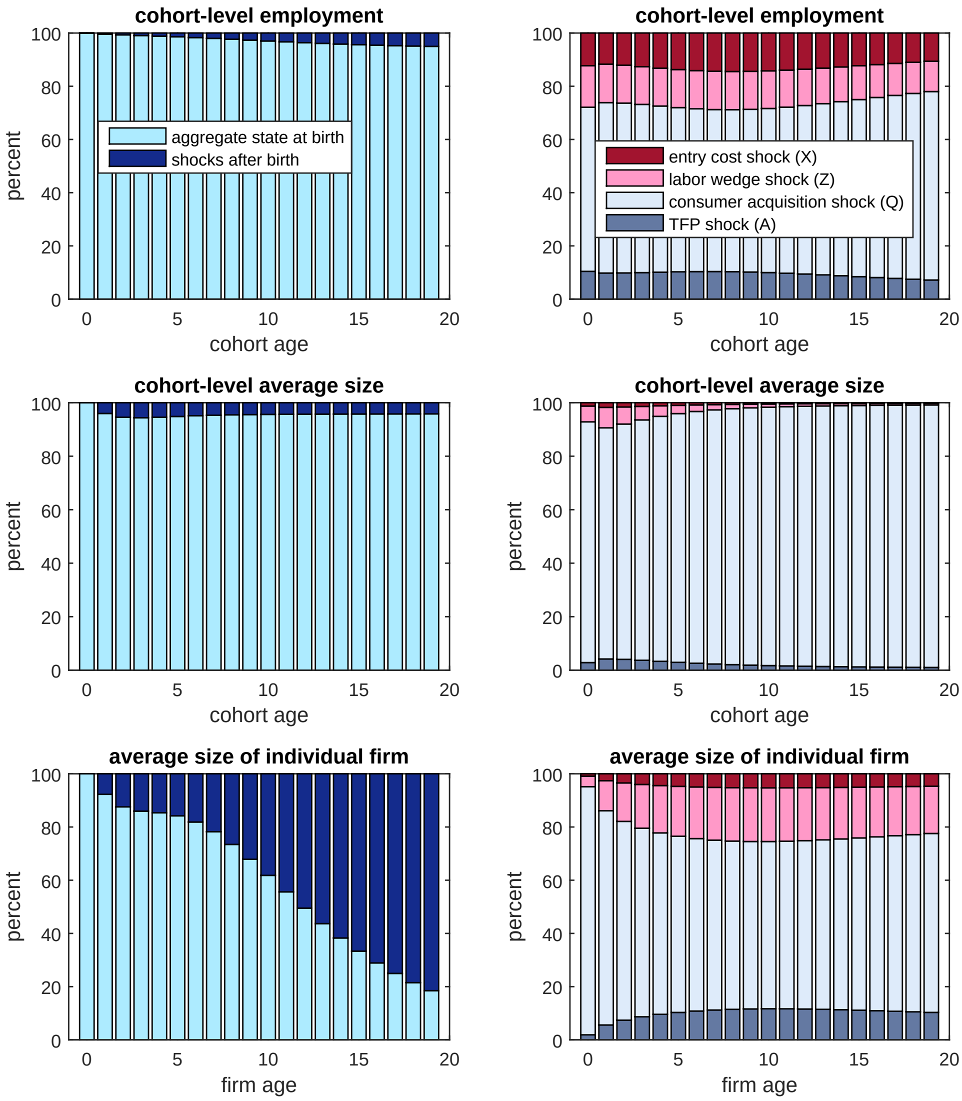


Notes: Employment rate data and a model-based counterfactual employment rate based on the fixing the age/type firm sizes to their respective steady state values.Results are based on a marketing adjustment costs curvature of $\zeta_1 = 3$.

E.3 Persistence of shocks

This Appendix checks the robustness of the results with respect to the point estimates of the persistence parameters. In particular, we consider a case in which all persistence parameters have a value of at most 0.8. Figures 26 and 28 show, respectively, the model

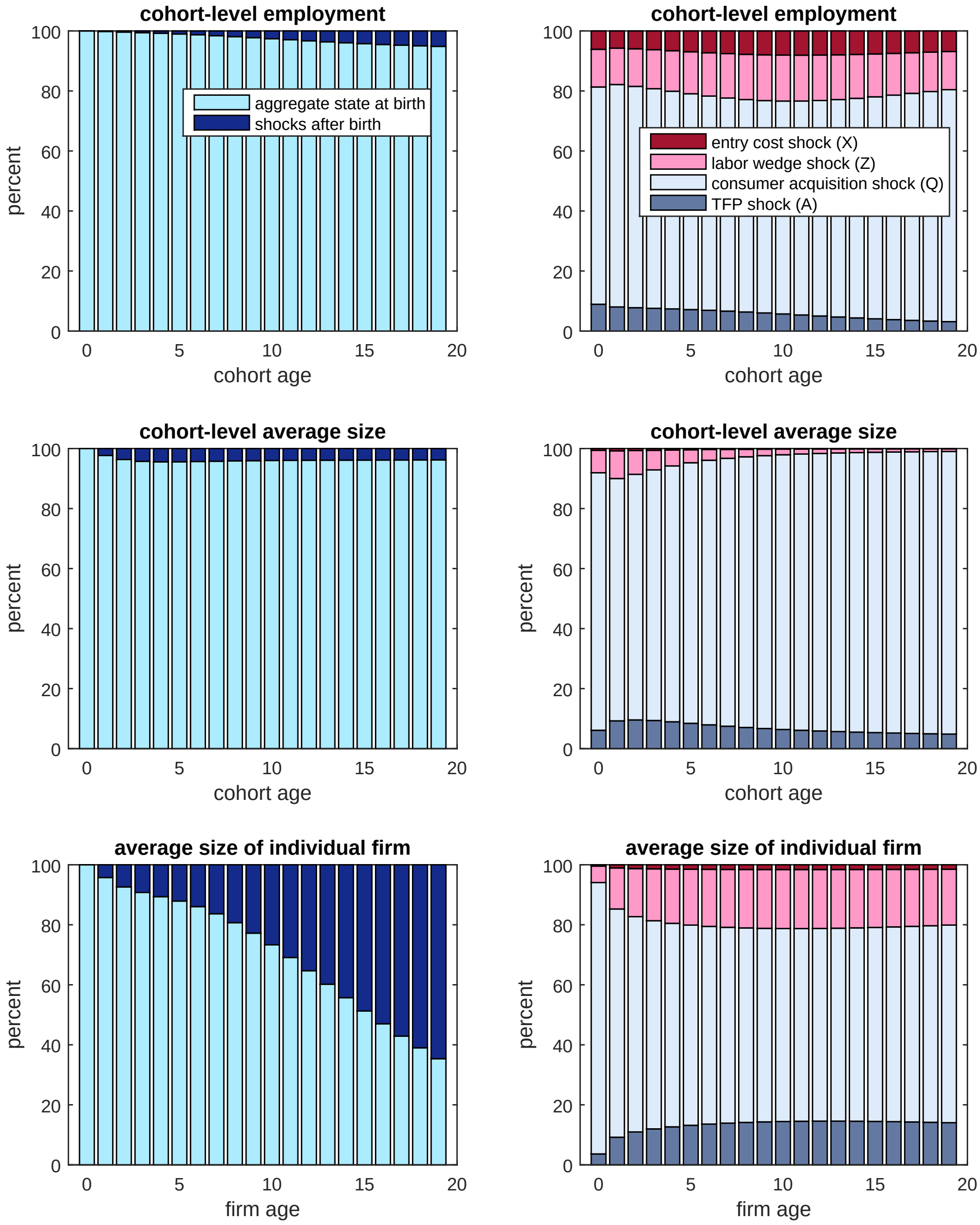
Figure 26: Model variance decompositions: lower persistence of shocks



Notes: Contributions of the aggregate state at birth and post-entry shocks (left panels) and the contributions of the four aggregate shocks (right panels) to variation in cohort-level employment (top row), cohort-level average size (middle row) and individual-firm average size (bottom row). Results are based on the case when all exogenous shocks have a persistence of at most 0.8.

variance decomposition and aggregate employment counterfactual under this alternative specification. The results are very close to those of the benchmark parametrization. The main difference can be seen in the faster decline of the influence of startup conditions in the case of an individual firm (bottom left panel).

Figure 27: Model variance decompositions: $K = 75$

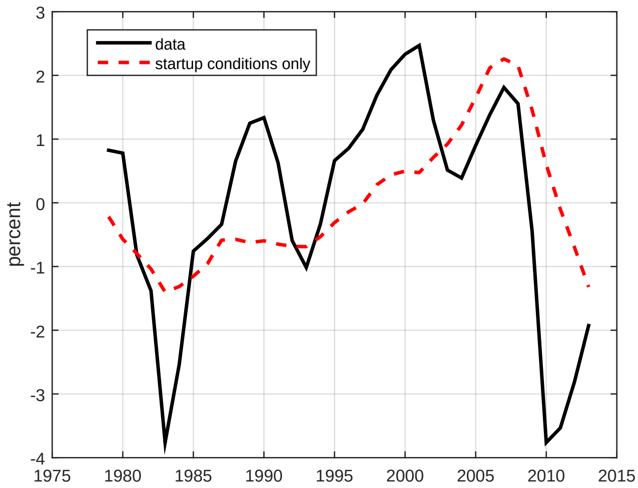


Notes: Contributions of the aggregate state at birth and post-entry shocks (left panels) and the contributions of the four aggregate shocks (right panels) to variation in cohort-level employment (top row), cohort-level average size (middle row) and individual-firm average size (bottom row). Results are based on the case when maximum firm age in the computations is increased to $K = 75$.

E.4 Maximum firm age

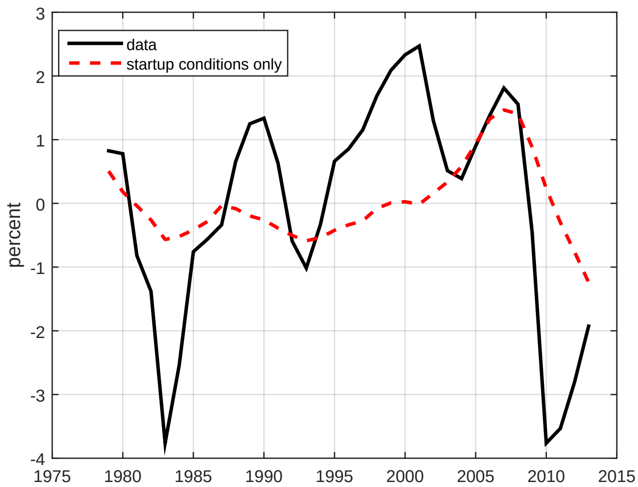
This Appendix shows that the results are robust to the maximum firm age K imposed in the solution of the model. In particular, we increase the maximum firm age from 50 to 75. Figure 27 and 29 show the model variance decomposition and the aggregate employment counterfactual based on variation in only startup conditions, respectively. In both cases the results are very similar to those in the benchmark model.

Figure 28: Contribution of startup conditions to aggregate employment: lower persistence of shocks



Notes: Employment rate data and a model-based counterfactual employment rate based on the fixing the age/type firm sizes to their respective steady state values. Results are based on the case when all exogenous shocks have a persistence of at most 0.8.

Figure 29: Contribution of startup conditions to aggregate employment: $K = 75$

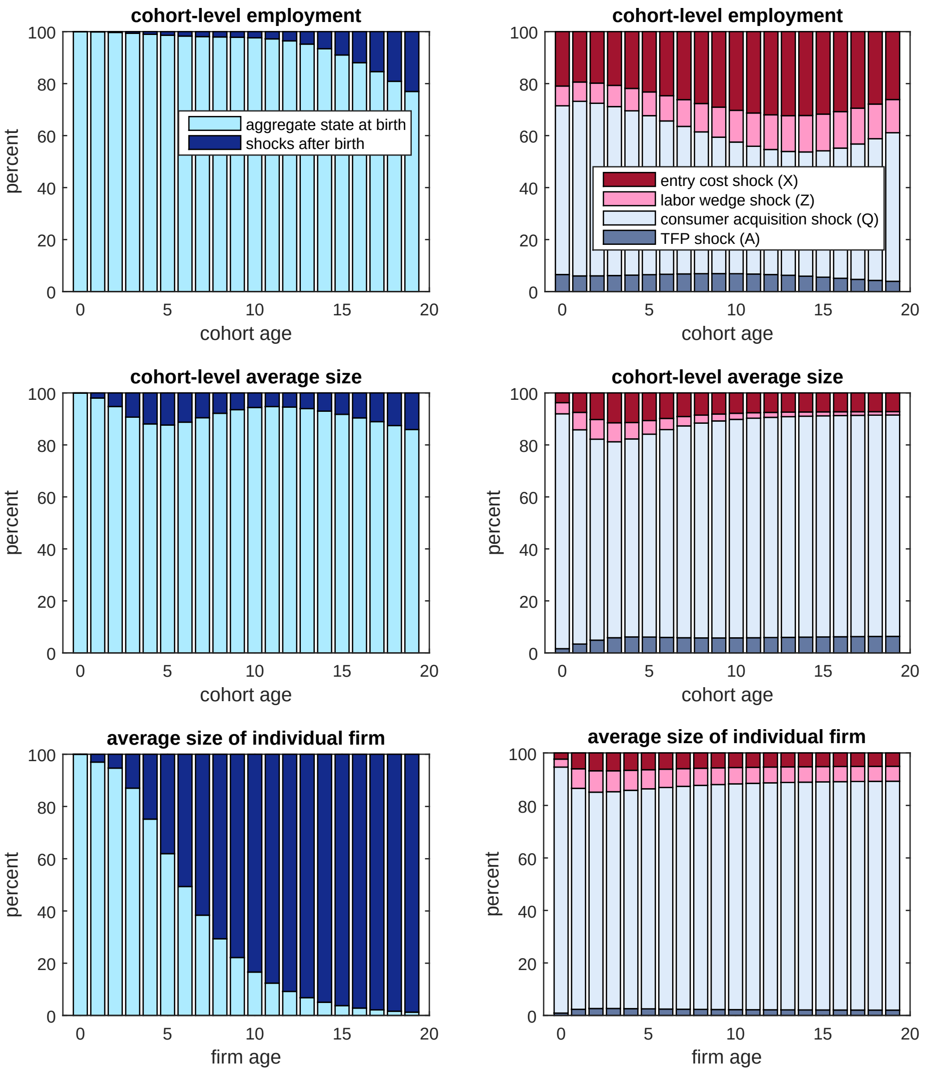


Notes: Employment rate data and a model-based counterfactual employment rate based on the fixing the age/type firm sizes to their respective steady state values. Results are based on the case when maximum firm age in the computations is increased to $K = 75$.

E.5 Positive consumer base depreciation

The benchmark model assumes zero depreciation of the consumer base. This feature is motivated by our micro-foundations which are based on assuming that firm's can relax their demand constraints by informing consumers about their products. We therefore find it realistic that consumers do not "forget" about products. Nevertheless, this appendix provides a robustenss exercise showing that the main model conclusions remain to hold even for positive values of consumer base depreciation. In particular, we assume that the

Figure 30: Model variance decompositions: positive consumer base depreciation



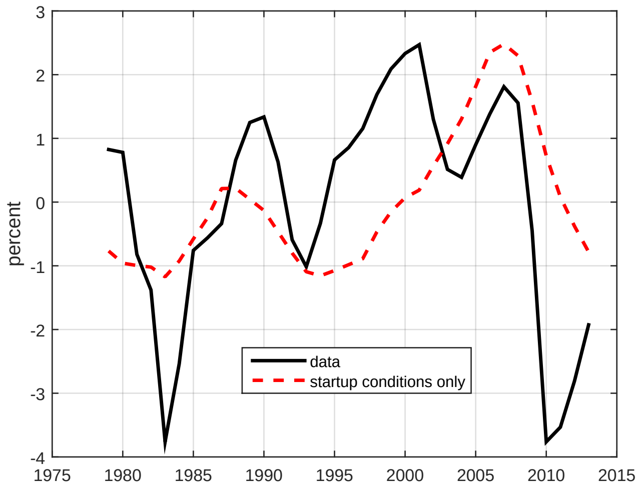
Notes: Contributions of the aggregate state at birth and post-entry shocks (left panels) and the contributions of the four aggregate shocks (right panels) to variation in cohort-level employment (top row), cohort-level average size (middle row) and individual-firm average size (bottom row). Results are based on the case when the depreciation rate of the cusumer base is set to 15%.

rate at which the consumer base depreciates is 15% as in Gourio and Rudanko (2014).[12]

Unlike with the other robustness checks, in this case it is necessary to recalibrate the firm-specific demand elasticities (and re-estimate the aggregate shocks). The reason is that a positive depreciation rate changes the shape of the growth-profiles such that our initial target (average size by age) was no longer met. Figures 30 and 31 show that even with positive depreciation of the consumer base, the state at birth and the demand shock in particular are extremely important for cohort-level variation in employment and average firm size. Moreover, as in the benchmark model, variation in startup conditions generates slow-moving changes in aggregate employment.

[12]Gourio and Rudanko (2014) acknowledge that there is large heterogeneity in depreciation rates of customer capital. Indeed, Foster, Haltiwanger, and Syverson (2016) estimate a depreciation rate close to 60% for very specific industries.

Figure 31: Contribution of startup conditions to aggregate employment: positive consumer base depreciation



Notes: Employment rate data and a model-based counterfactual employment rate based on the fixing the age/type firm sizes to their respective steady state values. Results are based on the case when the depreciation rate of the consumer base is set to 15%.

References

Dorfman, R., and P. Steiner (1954): "Optimal Advertising and Optimal Quality," *The American Economic Review*, 44(5), 826–836.

Foster, L., J. Haltiwanger, and C. Syverson (2016): "The Slow Growth of New Plants: Learning about Demand?," *Economica*, 83(329), 91–129.

Gourio, F., and L. Rudanko (2014): "Customer Capital," *Review of Economic Studies*, 81(3), 1102–1136.

Lee, Y., and T. Mukoyama (2013): "Entry, Exit, and Plant-level Dynamics over the Business Cycle," mimeo.

Moreira, S. (2015): "Firm Dynamics, Persistent Effects of Entry Conditions, and Business Cycles," mimeo.

Ravn, M., and H. Uhlig (2002): "On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations," *Review of Economics and Statistics*, 84(2), 371–380.

Sethuraman, R., G. Tellis, and R. Briesch (2011): "How Well Does Advertising Work? Generalizations from Meta-Analysis of Brand Advertising Elasticities," *Journal of Marketing Research*, 48, 457–471.