Supplementary Appendix

A Data Appendix

A.1 Sample summary and descriptive statistics

Table A.1 lists the countries in the sample, their survey waves, and the number of firm-level observations recorded.

Table A.1: Harmonized WBES sample merged with GDP p.c.

Country	Survey Waves	Num. Obs.
Austria	2021	600
Bahamas, The	2010	150
Belgium	2020	614
Croatia	2007 2013 2019 2023	1871
Cyprus	2019	240
Denmark	2020	995
Estonia	2009 2013 2019 2023	1257
Finland	2020	759
France	2021	1566
Germany	2021	1694
Greece	2018 2023	1198
Hungary	2009 2013 2019 2023	2237
Ireland	2020	606
Israel	2013	483
Italy	2019	760
Kazakhstan	2009 2013 2019	2590
Latvia	2009 2013 2019	966
Lithuania	2009 2013 2019	904
Luxembourg	2020	170
Malaysia	2015 2019	2221
Malta	2019	242
Netherlands	2020	808
Poland	2009 2013 2019	2366
Portugal	2019 2023	2069
Romania	2009 2013 2019 2023	2842
Russian Federation	2009 2012 2019	6547
Saudi Arabia	2022	1573
Slovak Republic	2009 2013 2019	972
Slovenia	2009 2013 2019	955
Spain	2021	1051
Sweden	2014 2020	1191

SOURCE: WBES, WDI, and authors' calculation.

Table A.2 reports summary statistics for the variables used in Section 2.

Table A.2: Summary statistics

Statistics	Mean	Median	SD	P25	P75
Firm size	57.949	48.069	32.042	36.075	71.117
Firm size cumulative growth, %	113.50	114.07	31.51	91.284	135.04
Firm size annualized growth, %	7.5342	7.3405	2.2200	5.9854	8.7939
Firm age (years)	21.418	19.631	7.4695	15.901	28.770
Firm performing R&D	0.2230	0.1975	0.1461	0.0917	0.3130
Firm performing process innovation	0.3118	0.2830	0.1797	0.1755	0.4354
Firm performing product innovation	0.4183	0.4082	0.2022	0.2357	0.5657
Sales (log)	14.018	13.976	0.7352	13.572	14.570
Material expenditure (log)	12.726	12.684	0.9570	12.091	13.504
Capital expenditure (log)	12.665	12.614	0.8514	11.969	13.437
Foreign	0.1120	0.1203	0.0608	0.0610	0.1501
Shareholding	0.7213	0.8302	0.2873	0.5375	0.9511

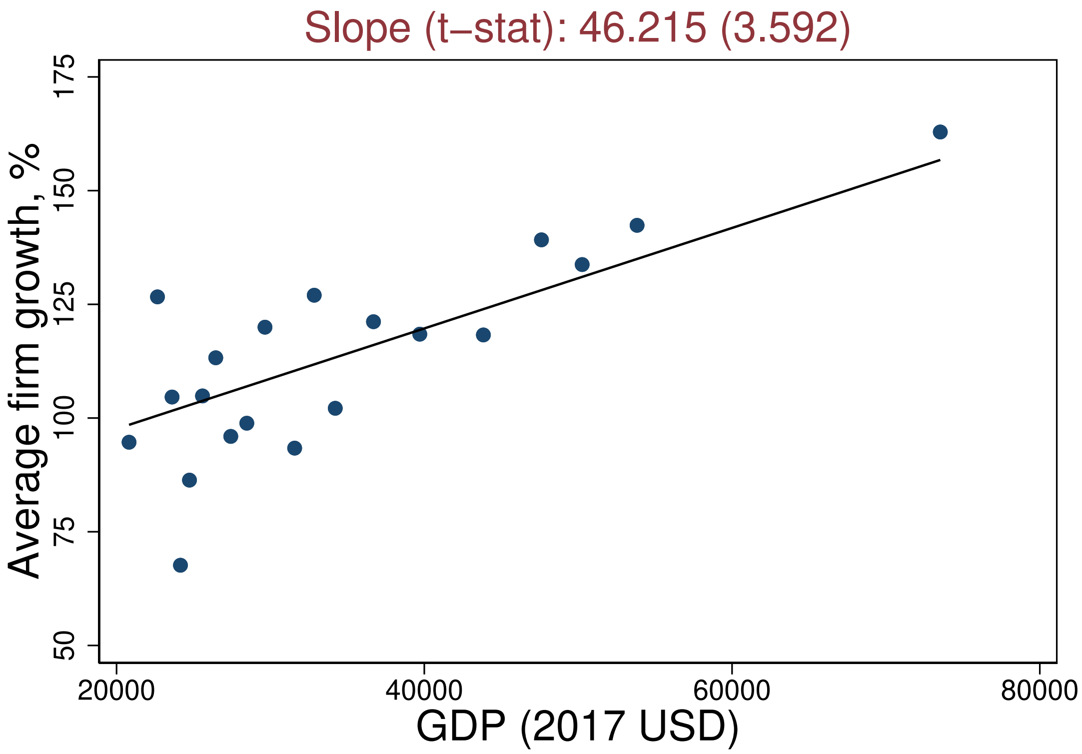
NOTES: Firm size to the current number of employees. Firm size cumulative growth is computed as the log difference between the current number of employees and the number of employees recorded in the first year of operations. Firm size annualized growth is computed dividing firm size cumulative growth by the recorded firm age. Sales refer to the establishment's total annual revenues. Material expenditure refers to the cost of raw materials and intermediate goods used in production in the last fiscal year. Capital expenditure refers to the cost for the establishment to re-purchase all of its machinery. Sales, material, and capital expenditure are deflated using the US GDP deflator and expressed in 2009 USD. Foreign is a dummy variable taking value 1 if the firm is owned by private foreign individuals, companies or organizations by any percent, 0 otherwise. Shareholding is a dummy variable taking value 1 if the firm is shareholding company (including both those with traded and non-traded shares), 0 otherwise. SOURCE: WBES and authors' calculation.

A.2 Firm Growth across Countries

Figure A.1 reports the average cumulative firm size growth across countries with different GDP per capita. Specifically, the cumulative firm size growth is about 90 percent on average in countries with a GDP per capita of 25,000 USD, and it increases to around 120 percent in countries with a GDP per capita of 60,000 USD.

Figure A.2 complements Figure A.1, and it reports the average cumulative firm size growth for firms that are 40 years old across countries. The average firm growth is higher in richer countries, even when conditioned on firm age, and it

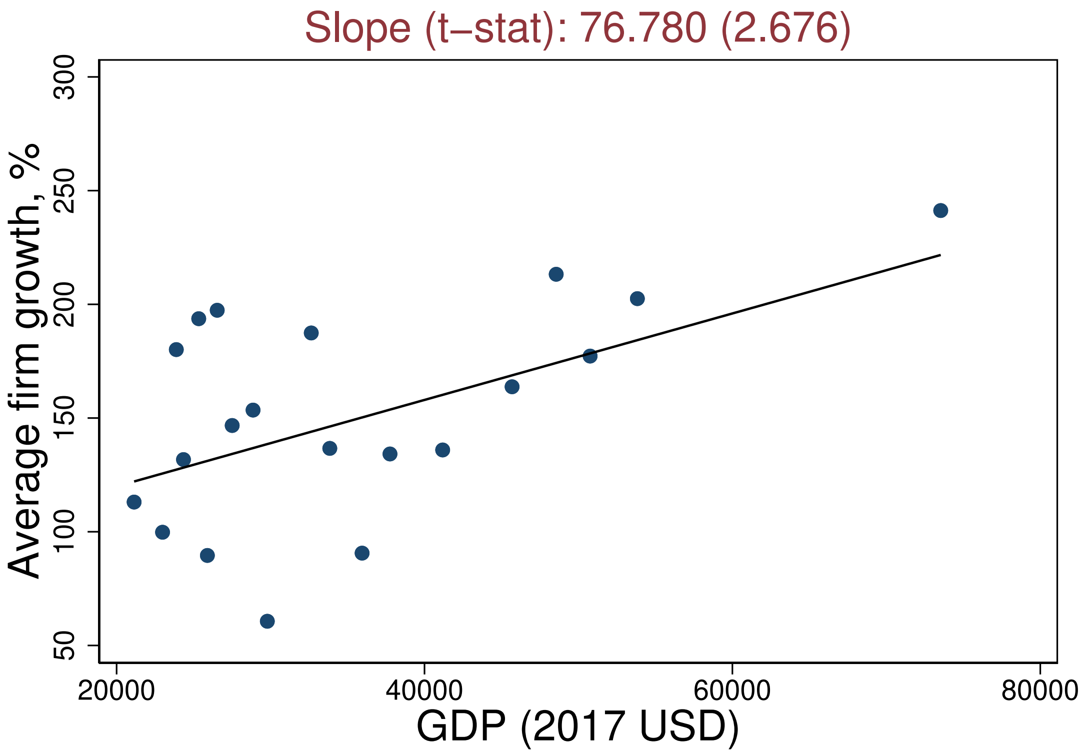
Figure A.1: Average cumulative firm growth across countries



NOTES: This figure is a binscatter plot of the average cumulative firm size growth since birth across countries over GDP per capita. Fitted lines are obtained from an auxiliary regression on GDP per capita. In the caption, we report the estimated slope of the regression and the t-stat (in parentheses). GDP per capita is expressed in 2017 USD constant prices. SOURCE: WBES, WDI, and authors' calculation.

increases from around 75 percent in countries with a GDP per capita of 20,000 USD to more than 150 percent in countries with a GDP per capita above 60,000 USD.

Figure A.2: Average cumulative firm size growth, 40 years



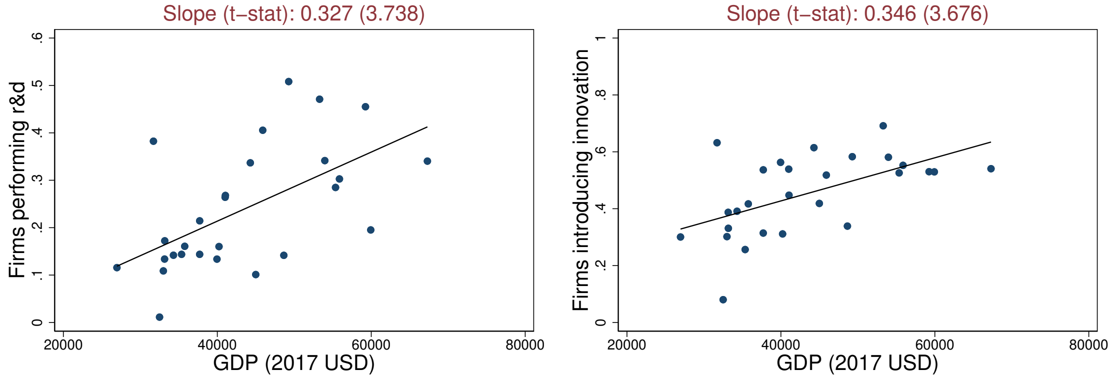
NOTES: This figure is a binscatter of the country-average cumulative firm size growth after 40 years of operations over GDP per capita. The fitted line is obtained from an auxiliary regression on GDP per capita. GDP per capita is expressed in 2017 USD constant prices.
SOURCE: WBES, WDI, and authors' calculation.

A.3 Technology Adoption across Countries

Here we complement the evidence described in Section 2 on technology adoption across countries. To this purpose, we use data from the Community Innovation Survey (CIS). The CIS, is conducted in every European Union (EU) Member State to collect data on firm-level R&D and innovation in businesses, i.e. on product innovation (goods or services) and process innovation (e.g. production methods; logistics; information processing and communication; accounting and administrative operations; organizational, management and marketing aspects). We use country-wide aggregated statistics provided by the Eurostat for 2022.

Figure A.3: Firm technology adoption over development

(a) R&D **(b)** Innovation



NOTES: Panel A binscatters the share of firms that conduct R&D across countries over their GDP per capita. Panel B binscatters the share of firms that conduct innovation (both product and process) across countries over their GDP per capita. Fitted lines are obtained from auxiliary regressions on GDP per capita. GDP per capita is expressed in 2017 USD constant prices. SOURCE: CIS, EUROSTAT, WDI, and authors' calculation.

Figure A.3 binscatters the share of firms performing R&D (panel A) and the share of firms conducting innovation (panel B), including product and process innovations. This figure replicates the evidence on technology adoption discussed in Section 5.1 in the main text. As we move from low to high GDP per capita countries in the EU, the share of firms performing R&D increases from 0.10 to 0.50. Similar, the share of firms conducting either product or process innovation (or both) doubles from 0.30 to 0.6.

A.4 Wage markdown across Countries

We measure labor market power at the firm level by comparing the firm's marginal revenue product of labor to the wage paid (Amodio et al., 2024). To do so, we first assume the following revenue function specification,

$$\ln y_{it} = \alpha + \xi \ln \ell_{it} + \gamma \ln k_{it} + \delta \ln m_{it} + \omega_{it} + \epsilon_{it}$$

where y_{it} is firm sales, ℓ_{it} denotes number of employees, k_{it} is capital, m_{it} materials of firm i and time t. Finally, ω_{it} captures a combination of productivity differences across firms and demand-side factors affecting the output price, while ϵ_{it} is instead an unobserved i.i.d. idiosyncratic shock to revenues with mean zero.

We estimate the parameters of the revenue production function separately for each country in the sample using a control function approach as in Levinsohn and Petrin (2003). This method relies on three main assumptions: (i) the term ω_{it} evolves according to a first-order Markov process; (ii) the term ω_{it} is the only unobservable in the firm's input demand function; and (iii) the input demand function is invertible in ω_{it}. Under these three assumptions, we can control for unobserved productivity and demand shocks non-parametrically, using materials as proxy variables. Let

$$m_{it} = F_M(k_{it}, \omega_{it})$$

be the firm-level demand function for materials. As commonly assumed, capital k_{it} is an argument of this function because it is assumed to be already determined when materials are chosen. Assuming monotonicity of F_M, one can invert the materials demand function and express productivity ω_{it} as a function of capital, k_{it}, and material, m_{it}, i.e.,

$$\omega_{it} = \Lambda_M(k_{it}, m_{it}).$$

Substituting the latter back into the production function equation, we can write:

$$\ln y_{it} = \alpha + \xi \ln \ell_{it} + \phi(k_{it}, m_{it}) + \epsilon_{it}, \tag{13}$$

where

$$\phi(k_{it}, m_{it}) = \gamma \ln k_{it} + \delta \ln m_{it} + \Lambda_M(k_{it}, m_{it}).$$

As in Levinsohn and Petrin (2003), we can obtain an estimate of the revenue elasticity of labor, $\hat{\xi}$, by estimating equation (13) with OLS, using a fully interacted second-order polynomial in k_{it} and m_{it} to approximate $\phi(k_{it}, m_{it})$.

We estimate equation (13) separately for every country in the sample. Table A.3

reports selected summary statistics for the distribution of estimated country-specific revenue elasticities of employment, $\hat{\xi}$.

Table A.3: Estimated revenue elasticity of employment

Statistics	Mean	Median	SD	P25	P75
$\hat{\xi}$	0.4472	0.4110	0.2042	0.3267	0.5557

NOTES: This table reports summary statistics for the distribution of estimated country-specific revenue elasticities of employment, $\hat{\xi}$. SOURCE: WBES and authors' calculation.

The mean estimate for the elasticity is 0.4472. The distribution of estimates is not excessively dispersed across countries: the standard deviation is 0.2042, and the inter-quartile range is 0.2283.

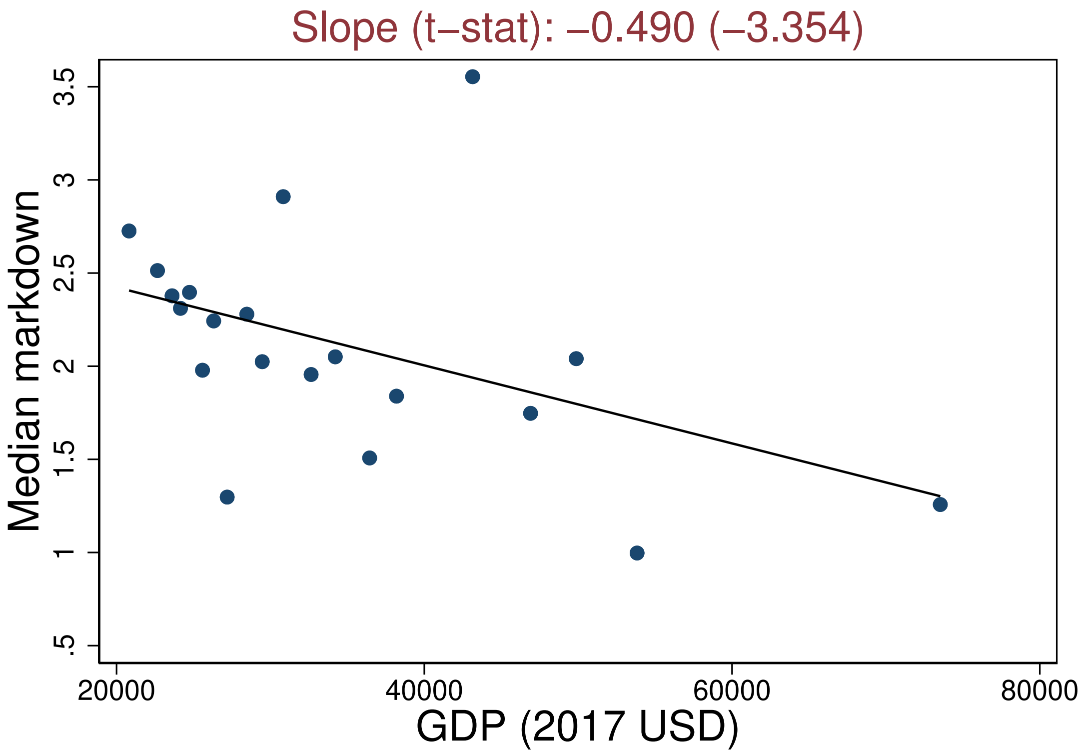
Using our estimates for the revenue elasticity of labor, $\hat{\xi}$, we derive the wage markdown as a ratio between the marginal revenue product of labor and the wage paid by firm i at time t,

$$\mu_{it} = \frac{\text{MRPL}_{it}}{w_{it}} \quad \text{where} \quad \text{MRPL}_{it} = \frac{\partial y_{it}}{\partial \ell_{it}} = \hat{\xi}\frac{y_{it}}{\ell_{it}},$$

Notice that if the product market were not perfectly competitive, firms could raise their output price above marginal costs. On the other hand, price markup will not confound the estimate for wage markdown obtained from the equation above because it is based on the revenue elasticity of labor (Amodio et al., 2024).

Finally, we complement the analysis of labor market power by estimating firm-level markdowns controlling for year fixed effects and firm-level observables, such as location fixed effects, 2-digit ISIC v.3.1 industry fixed effects, firm-age fixed effects, a dummy for foreign ownership, and a dummy for shareholding companies. Figure A.4 binscatters the estimated median markdown against GDP per capita of countries in the sample. Controlling for firm-level observables does not alter the cross-country patterns uncovered in Section 2.

Figure A.4: Median markdown across countries - Controlling for firm-level observables.



NOTES: This figure shows a binscatter of the median firm-level markdown across countries over GDP per capita. Firm-level markdown are estimated controlling for year fixed effects and firm-level observables, such as location fixed effects, 2-digit ISIC v.3.1 industry fixed effects, age fixed effects, a dummy for foreign ownership, and a dummy for shareholding companies. Fitted lines are obtained from an auxiliary regression on log GDP per capita. In the caption of each panel, we report the estimated slope of the regression and the t-stat (in parentheses). GDP per capita is expressed in 2017 USD constant prices. SOURCE: WBES, WDI, and authors' calculation. SOURCE: WBES, and authors' calculation.

B Model Appendix

B.1 Derivation of labor supply function

Equation (3) in the main text shows the that labor supply function to a firm j is equal to

$$L_j = L \int_{\mathcal{Z}\times\mathcal{A}} p_{ij}\phi(z_i, a_i) dz_i da_i$$

where p_{ij} is the probability that a worker i chooses to work for a firm j, while $\phi(z_i, a_i)$ is the equilibrium distribution of workers i across productivity and amenities.

Recall that p_{ij} is equal to:

$$p_{ij} = \frac{\exp\left(\frac{U(z_i,a_i,z_j,a_j)}{\sigma_\nu}\right)}{\int_L^1 \exp\left(\frac{U(z_i,a_i,z_k,a_k)}{\sigma_\nu}\right) dk}$$

By a change of variable and expanding the value functions, we can re-write the previous expression as:

$$p_{ij} = \frac{\exp\left(\frac{\epsilon^L \ln(w_j)+a_j+\beta(1-\delta_w)\mathbb{E}_{z_i'} \max\{V(z_i',a_i),\tilde{U}(z_i',a_i)\}}{\sigma_\nu}\right)}{E\int_{\mathcal{Z}\times\mathcal{A}} \exp\left(\frac{\epsilon^L \ln(w_k)+a_k+\beta(1-\delta_w)\mathbb{E}_{z_i'} \max\{V(z_i',a_i),\tilde{U}(z_i',a_i)\}}{\sigma_\nu}\right)\psi(z_k, a_k) dz_k da_k}$$

Substituting the last expression for p_{ij} into the labor supply function, L_j, we can write:

$$L_j = L \int_{\mathcal{Z}\times\mathcal{A}} \left(\frac{\exp\left(\frac{\epsilon^L \ln(w_j)+a_j+\beta(1-\delta_w)\mathbb{E}_{z_i'} \max\{V(z_i',a_i),\tilde{U}(z_i',a_i)\}}{\sigma_\nu}\right)}{E\int_{\mathcal{Z}\times\mathcal{A}} \exp\left(\frac{\epsilon^L \ln(w_k)+a_k+\beta(1-\delta_w)\mathbb{E}_{z_i'} \max\{V(z_i',a_i),\tilde{U}(z_i',a_i)\}}{\sigma_\nu}\right) d\psi(z_k,a_k)} \right) d\phi(z_i,a_i)$$

or,

$$L_j = L \int_{\mathcal{Z}\times\mathcal{A}} \left(\frac{\exp\left(\frac{\epsilon^L \ln(w_j)+a_j}{\sigma_V}\right)\exp\left(\frac{\beta(1-\delta_w)\mathbb{E}_{z_i'}\max\{V(z_i',a_i),\tilde{U}(z_i',a_i)\}}{\sigma_V}\right)}{E\int_{\mathcal{Z}\times\mathcal{A}}\exp\left(\frac{\epsilon^L \ln(w_k)+a_k+\beta(1-\delta_w)\mathbb{E}_{z_i'}\max\{V(z_i',a_i),\tilde{U}(z_i',a_i)\}}{\sigma_V}\right)d\psi(z_k,a_k)} \right) d\phi(z_i,a_i)$$

which is equivalent to:

$$L_j = L\Theta \exp\left(\epsilon^L \ln(w_j) + a_j\right)$$

where Θ is an aggregate market shifter, defined as:

$$\Theta = \frac{1}{\exp(\sigma_V)} \int_{\mathcal{Z}\times\mathcal{A}} \left(\frac{\exp\left(\frac{\beta(1-\delta_w)\mathbb{E}_{z_i'}\max\{V(z_i',a_i),\tilde{U}(z_i',a_i)\}}{\sigma_V}\right)}{E\int_{\mathcal{Z}\times\mathcal{A}}\exp\left(\frac{\epsilon^L \ln(w_k)+a_k+\beta(1-\delta_w)\mathbb{E}_{z_i'}\max\{V(z_i',a_i),\tilde{U}(z_i',a_i)\}}{\sigma_V}\right)\psi(z_k,a_k)dz_k da_k} \right) \phi(z_i,a_i)dz_i da_i$$

B.2 Equilibrium distribution

Equations (14), (15), and (16) describe the lows of motion for the measure of employers across productivity z and amenities a.

$$\begin{aligned} \Omega(z_i,a)' = {} & (1-\delta_w)p_n[(1-\rho^z(z_{i-},a))\rho^e(z_{i-},a) + (1-\rho^e(z_{i-},a))]\Omega(z_{i-},a) \\ & + (1-\delta_w)p_i\rho^z(z_{i-},a)\rho^e(z_{i-},a)\Omega(z_{i-},a) \\ & + (1-\delta_w)(1-p_n)[(1-\rho^z(z_{i+},a))\rho^e(z_{i+},a) + (1-\rho^e(z_{i+},a))]\Omega(z_{i+},a) \\ & + (1-\delta_w)(1-p_i)\rho^z(z_{i+},a)\rho^e(z_{i+},a)\Omega(z_{i+},a) \\ & + \delta_w\Psi_z(z_i)\Psi_a(a) \end{aligned} \quad (14)$$

$$
\begin{aligned}
\Omega(\overline{z}, a)' = & (1-\delta_w) p_n [(1-\rho^z(\overline{z}_-, a))\rho^e(\overline{z}_-, a) + (1-\rho^e(\overline{z}_-, a))]\Omega(\overline{z}_-, a) \\
& + (1-\delta_w) p_i \rho^z(\overline{z}_-, a)\rho^e(\overline{z}_-, a)\Omega(\overline{z}_-, a) \\
& + (1-\delta_w) p_n [(1-\rho^z(\overline{z}, a))\rho^e(\overline{z}, a) + (1-\rho^e(\overline{z}, a))]\Omega(\overline{z}, a) \\
& + (1-\delta_w) p_i \rho^z(\overline{z}, a)\rho^e(\overline{z}, a)\Omega(\overline{z}, a) \\
& + \delta_w \Psi_z(\overline{z})\Psi_a(a)
\end{aligned} \tag{15}
$$

$$
\begin{aligned}
\Omega(\underline{z}, a)' = & (1-\delta_w)(1-p_n)[(1-\rho^z(\underline{z}, a))\rho^e(\underline{z}, a) + (1-\rho^e(\underline{z}, a))]\Omega(\underline{z}, a) \\
& + (1-\delta_w)(1-p_i)\rho^z(\underline{z}, a)\rho^e(\underline{z}, a)\Omega(\underline{z}, a) \\
& + (1-\delta_w)(1-p_n)[(1-\rho^z(\underline{z}_+, a))\rho^e(\underline{z}_+, a) + (1-\rho^e(\underline{z}_+, a))]\Omega(\underline{z}_+, a) \\
& + (1-\delta_w)(1-p_i)\rho^z(\underline{z}_+, a)\rho^e(\underline{z}_+, a)\Omega(\underline{z}_+, a) \\
& + \delta_w \Psi_z(\underline{z})\Psi_a(a)
\end{aligned} \tag{16}
$$

B.3 Numerical algorithm

The algorithm to solve for equilibrium goes as follows:

1. Guess a stationary distribution of agents over productivity and amenities $\Omega(z, a)^1$.
2. Given the current distribution $\Omega(z, a)^i$:
   (a) Guess the entrepreneurship policy function $\rho^{e,j}(z, a)$.
   (b) Using $\Omega(z, a)^i$ and $\rho^{e,j}(z, a)$, compute the distributions of workers and entrepreneurs over z and a: $\phi(z, a)$ and $\psi(z, a)$, and the measures of workers, L, and entrepreneurs, E.
   (c) Given $\phi(z, a)$, $\psi(z, a)$, L and E, solve the fixed point of the value functions to obtain U, $\tilde{U}$, V, W and π.
   (d) Using V, and $\tilde{U}$, update $\rho^{e,j+1}(z, a)$.
   (e) Check for convergence of the entrepreneurship policy function, if not equal, return to step (2.b) with the new one.

3. Use Equations (14) and (15) and (16) to get $\Omega(z,a)^{i+1}$, if not sufficiently close to $\Omega(z,a)^i$ return to step 2.

B.4 Derivation of model properties

An interior solution to the static problem of the firm presented in Section 3 satisfies the following first order condition:

$$\xi z_j L_j^{\xi-1} = w_j \left(1 + \frac{1}{\epsilon^L}\right). \tag{17}$$

That is, the wage paid by firm j, w_j, is equal to a markdown, $(1+\frac{1}{\epsilon^L})^{-1}$ over its marginal product of labor, $\xi z_j L_j^{\xi-1}$. Taking logs, we get:

$$(1-\xi)\ln L_j = \ln z_j + \ln \xi - \ln w_j - \ln\left(1 + \frac{1}{\epsilon^L}\right).$$

Using the labor supply function to substitute for (log) wages, we obtain:

$$(1-\xi)\ln L_j = \ln z_j + \ln \xi - \left(\frac{1}{\epsilon^L}\ln L_j - \frac{1}{\epsilon^L}\ln L - \frac{1}{\epsilon^L}\ln\Theta - \frac{1}{\epsilon^L}a_j\right) - \ln\left(1 + \frac{1}{\epsilon^L}\right).$$

Solving for L_j, we get:

$$\ln L_j = \frac{\epsilon^L}{1+(1-\xi)\epsilon^L}\ln z_j + \frac{1}{1+(1-\xi)\epsilon^L}a_j + C \tag{18}$$

where

$$C = \frac{1}{1+(1-\xi)\epsilon^L}\left[\epsilon^L \ln\left(\frac{\epsilon^L}{1+\epsilon^L}\right) + \epsilon^L \ln\xi + \ln L + \ln\Theta\right].$$

Notice that the log difference between the employment of firms with different productivity levels, $\underline{z}$ and $\bar{z}$, and same amenities, a, is equal to:

$$\ln L(\bar{z},a) - \ln L(\underline{z},a) = \frac{\epsilon^L}{1+(1-\xi)\epsilon^L}[\ln(\bar{z}) - \ln(\underline{z})]$$

Similarly, the log difference between the employment of firms with the same productivity, z, and different amenities, $\underline{a}$ and $\bar{a}$, we get:

$$\ln L(z, \bar{a}) - \ln L(z, \underline{a}) = \frac{1}{1 + (1 - \xi)\epsilon^L} [\bar{a} - \underline{a}]$$

Taking the exponential of both log-differences gives equations (7) and (8) in the main text. This proves proposition 1.

Consider now the average product of labor for a firm j, i.e.,

$$\text{APL}_j = \frac{z_j L_j^{\xi}}{L_j}.$$

The elasticity of APL_j with respect to productivity z_j is equal to:

$$\frac{\partial \ln \text{APL}_j}{\partial \ln z_j} = \frac{\partial(\ln z_j + (\xi - 1) \ln L_j)}{\partial \ln z_j} = 1 + (\xi - 1) \frac{\partial \ln L_j}{\partial \ln z_j}$$

Using equation (18), we know that

$$\frac{\partial \ln L_j}{\partial \ln z_j} = \frac{\epsilon^L}{1 + (1 - \xi)\epsilon^L}$$

which implies that

$$\frac{\partial \ln \text{APL}_j}{\partial \ln z_j} = 1 - \frac{(1 - \xi)\epsilon^L}{1 + (1 - \xi)\epsilon^L}$$

Similarly, the elasticity of APL_j with respect to exponential of amenities, $\exp(a_j)$, is equal to:

$$\frac{\partial \ln \text{APL}_j}{\partial a_j} = \frac{\partial(\ln z_j + (\xi - 1) \ln L_j)}{\partial a_j} = +(\xi - 1) \frac{\partial \ln L_j}{\partial \ln a_j}.$$

Using equation (18), we know that

$$\frac{\partial \ln L_j}{\partial a_j} = \frac{1}{1 + (1-\xi)\epsilon^L}$$

which implies that

$$\frac{\partial \ln \mathrm{APL}_j}{\partial a_j} = -\frac{(1-\xi)}{1 + (1-\xi)\epsilon^L}$$

This proves proposition 2 in the main text.

Finally, we can use the first order condition in equation (17) to express variable profits for firm j, π_j, as:

$$\pi_j = z_j L_j^{\xi} - w_j L_j = w_j L_j \left(\frac{1+\epsilon^L}{\xi \epsilon^L} - 1 \right)$$

This implies that the elasticity of variable profits, π_j, with respect to firm productivity, z_j, is equal to:

$$\begin{aligned}
\frac{\partial \ln \pi_j}{\partial \ln z_j} &= \frac{\partial \left[\ln w_j + \ln L_j + \ln \left(\frac{1+\epsilon^L}{\xi \epsilon^L} - 1 \right) \right]}{\partial \ln z_j} \\
&= \frac{\partial \ln w_j}{\partial \ln z_j} + \frac{\partial \ln L_j}{\partial \ln z_j} \\
&= \frac{\partial \ln w_j}{\partial \ln z_j} + \frac{\epsilon^L}{1 + (1-\xi)\epsilon^L}
\end{aligned}$$

where the last line is obtained from differentiating equation (18) with respect to $\ln z_j$.

To obtain $\partial \ln w_j / \partial \ln z_j$, we need to solve for wages, $\ln w_j$. To do, we can plug $\ln L_j$ from equation (18) into the labor supply function in equation (3) of the main text, and obtain:

$$
\begin{aligned}
\epsilon^L \ln w_j &= \ln L_j - a_j - \ln L - \ln \Theta \\
\epsilon^L \ln w_j &= \left(\frac{\epsilon^L}{1+(1-\xi)\epsilon^L} \ln z_j + \frac{1}{1+(1-\xi)\epsilon^L} a_j + C \right) - a_j - \ln L - \ln \Theta \\
\epsilon^L \ln w_j &= \frac{\epsilon^L}{1+(1-\xi)\epsilon^L} \ln z_j - \frac{(1-\xi)\epsilon^L}{1+(1-\xi)\epsilon^L} a_j + C - \ln L - \ln \Theta \\
\epsilon^L \ln w_j &= \frac{\epsilon^L}{1+(1-\xi)\epsilon^L} \ln z_j - \frac{(1-\xi)\epsilon^L}{1+(1-\xi)\epsilon^L} a_j + C - \ln L - \ln \Theta
\end{aligned}
$$

which implies:

$$
\begin{aligned}
\ln w_j = &\frac{1}{1+(1-\xi)\epsilon^L} \ln z_j - \frac{(1-\xi)}{1+(1-\xi)\epsilon^L} a_j \\
&+ \frac{1}{1+(1-\xi)\epsilon^L} \left[\ln \xi + \ln \frac{\epsilon^L}{1+\epsilon^L} - (1-\xi) \ln L - (1-\xi) \ln \Theta \right].
\end{aligned}
$$

Differentiating $\ln w_j$ with respect to $\ln z_j$, we get:

$$
\frac{\partial \ln w_j}{\partial \ln z_j} = \frac{1}{1+(1-\xi)\epsilon^L}
$$

Therefore:

$$
\frac{\partial \ln \pi_j}{\partial \ln z_j} = \frac{1}{1+(1-\xi)\epsilon^L} + \frac{\epsilon^L}{1+(1-\xi)\epsilon^L} = \frac{1+\epsilon^L}{1+(1-\xi)\epsilon^L}
$$

which is equal to equation (11) of the main text. Similarly, the elasticity of variable profits, π_j with respect to the firm amenity, $\exp(a_j)$, is equal to:

$$
\begin{aligned}
\frac{\partial \ln \pi_j}{\partial a_j} &= \frac{\partial \left[\ln w_j + \ln L_j + \ln \left(\frac{1+\epsilon^L}{\xi \epsilon^L} - 1 \right) \right]}{\partial a_j} = \frac{\partial \ln w_j}{\partial a_j} + \frac{\partial \ln L_j}{\partial a_j} \\
&= -\frac{(1-\xi)}{1+(1-\xi)\epsilon^L} + \frac{1}{1+(1-\xi)\epsilon^L} = \frac{\xi}{1+(1-\xi)\epsilon^L}
\end{aligned}
$$

which is equal to equation (12) of the main text. This proves proposition 3.

C Quantitative Appendix

C.1 Estimation

Table C.1 reports the list of parameters that are calibrated externally to the model, their values their values and targets/sources.

Table C.1: Parameters Set Without Solving the Model

Parameters	Description	Value	Targets/Source
A	Aggregate productivity shifter	1	normalization
σ_v	Type-I GEV shock scale	1	normalization
β	Discount factor	0.961	annual interest rate=0.04
δ_w	Retirement rate	0.025	40 years in the labor market
ξ	Revenue elasticity of labor	0.333	production fct. estimation
ϵ^L	Elasticity of labor supply	3.318	median markdown=1.301

NOTES: The table shows the parameters calibrated externally, their values, and the target or source used.

Table C.2 reports the list of targeted moments and their moment counterparts in the benchmark estimation. The sum of squared deviations between empirical and simulated moments is equal to 1.7 percent.

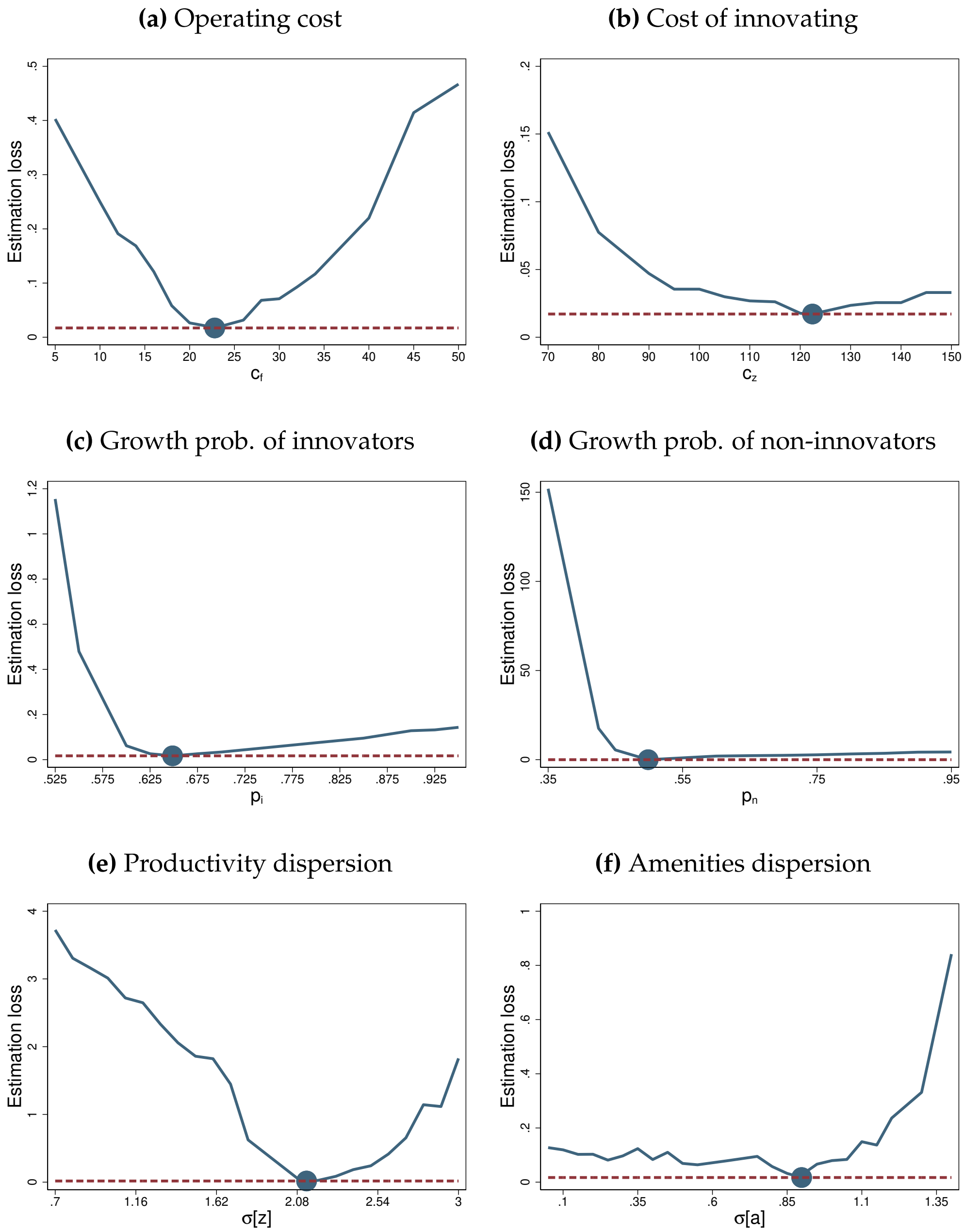
Table C.2: Targeted Moments: Data vs. Model

	Data (1)	Model (2)
Average firm size	59.02	58.09
Log firm size dispersion	0.988	1.091
Average cumulative growth rate (%)	152.6	141.4
Average firm age	30.13	30.84
Log wage dispersion	0.418	0.421
Firms investing in R&D	0.407	0.404

NOTES: This table reports the estimation fit. Column (1) the value of empirical targets. Column (2) reports model-based simulated moments.

Figure C.1 reports how the estimation loss varies over the parametric space, separately for each parameters. The blue dots refer to the point estimates. The red dashed line refers to the estimation loss evaluated at the final estimates. The estimation loss is minimized over each dimension of the parametric space.

Figure C.1: Estimation Loss

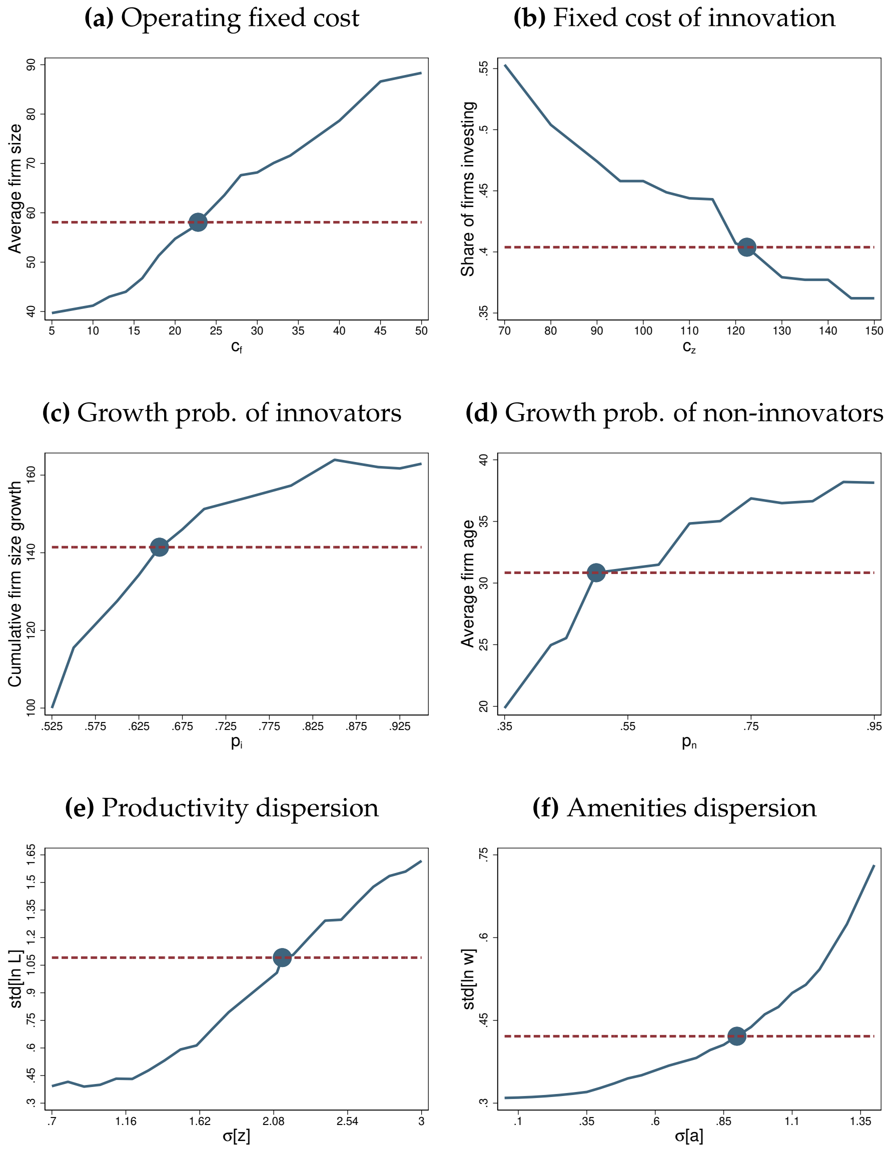


NOTES: This figure reports how the estimation loss (blue line) varies over the parametric space, separately for each parameters. The blue dots refer to the point estimates. The red dashed line refers to the estimation loss evaluated at the final estimates.

C.2 Identification

Figure C.2 reports how targeted moments change over the parametric space. Each panel shows that changing parameter values trigger a monotonic shift in the their respective target, suggesting identification is achieved.

Figure C.2: Mapping moments to parameters



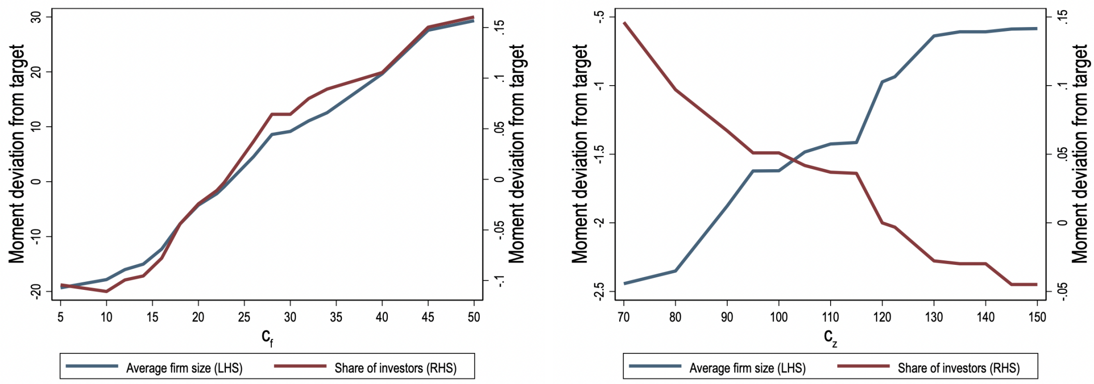
NOTES: This figure reports how each moment react to changes in parameters they are meant to identify. The red dashed line refer to the values of the empirical target.

Operating vs. innovation cost. In the estimation of the baseline model, the operating costs c_f is identified by targeting the average firm size (Figure C.2, panel A): by reducing profits, larger operating costs discourages lower-productivity, lower-amenities agents to become entrepreneurs. The innovation cost c_x is instead identified by targeting the share of firms conducting productivity-enhancing innovation (Figure C.2, panel B): everything else equal, larger fixed costs of innovation reduce the net benefits of innovation

Figure C.3 reports how both moments, i.e., the average firm size and the share of firms that innovate, respond to changes in c_f (panel A) and c_x (panel B), relative to the respective empirical targets.

Figure C.3: Contamination in the estimates of operating and innovation costs.

(a) Operating cost **(b)** Innovation cost



NOTES: This figure plots how the average firm size (blue line, left-hand side) and share of firms that innovate (red line, right-hand side) change with different value of c_f (panel A) and c_z (panel B), relative to their respective empirical targets.

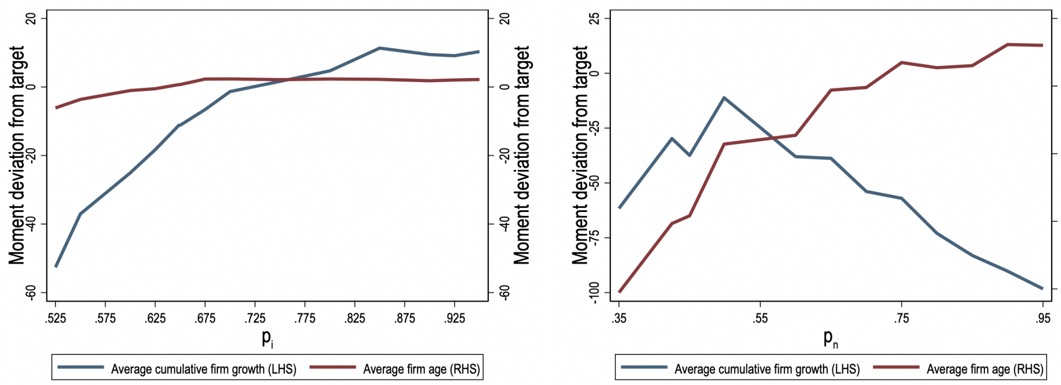
Changes in operating costs produce a monotonic increase in both moments. As we increase c_f, low-productivity, low-amenities agents are less likely to become entrepreneurs, and the distribution of active employer shifts towards entrepreneurs that can provide better amenities or with higher productivity. These entrepreneurs are also those who to run larger firms and more likely to engage in innovation. On the other hand, changes in innovation costs have an asymmetric effect: as we increase c_x, the share of firms that innovate declines, while the average firm size increases. While the first effect is trivial, the second one is again due to selection: when the innovation cost increases, average profits decline, and only high-productivity, high-amenities agents are have incentives to become entrepreneurs.

Productivity jumps: p_i vs. p_n. We repeat a similar analysis for the probabilities of productivity jump, p_i and p_n. In the estimation of the baseline model, these two parameters are identified by the average cumulative firm growth and the average firm age, respectively (Figure C.2, Panels E and F).

Figure C.4 reports how both moments, average cumulative firm size growth and average firm age, respond to changes in p_i (panel A) and $_n$ (panel B), relative to the respective empirical targets.

Figure C.4: Contamination in the estimates of productivity jumps

(a) Growth prob. of innovators **(b)** Growth prob. of non-innovators



NOTES: This figure plots how the average cumulative firm size growth (blue line, left-hand side) and the average firm age (red line, right-hand side) change with different value of p_i (panel A) and p_n (panel B), relative to their respective empirical targets.

Cross-contamination in the identification of p_i and p_n appears to be limited as well. Panel A shows that shifting p_i significantly affects the average cumulative firm size growth, i.e., the blue line rises monotonically, while it has a small effect on average firm age, i.e., the red line remains almost flat. This confirms that the former is informative about the parameter p_i, while the latter is not.

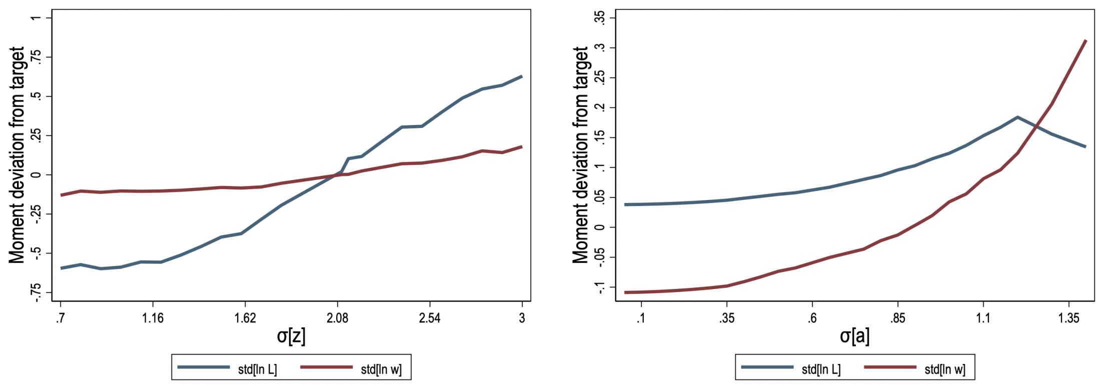
Viceversa, panel B shows that shifting p_n significantly affects the average firm age, i.e., the red line increases steeply (the right-hand axis), whereas changes in the average cumulative firm growth are large but not monotonic over the parametric space, i.e., the blue line first rises and then declines as we increase p_n. This confirms that the former moment is more informative about the parameters p_n compared to the latter.

Productivity dispersion vs. amenity dispersion. Finally, we study the identification of productivity and amenity dispersion, σ_z and σ_a. In the baseline estimation, σ_z is identified by targeting the standard deviation of log firm size, whereas σ_a is identified by the standard deviation of log wages (Figure C.2, Panels C and D). Higher probabilities rise the likelihood of productivity improvements over the firm life-cycle, increasing firm survival and employment growth.

Figure C.5 reports how both moments, i.e., log size and log wage standard deviation, respond to changes in σ_z (panel A) and σ_a (panel B), relative to the respective empirical targets.

Figure C.5: Contamination in the estimates of productivity and amenity dispersion.

(a) Productivity **(b)** Amenity



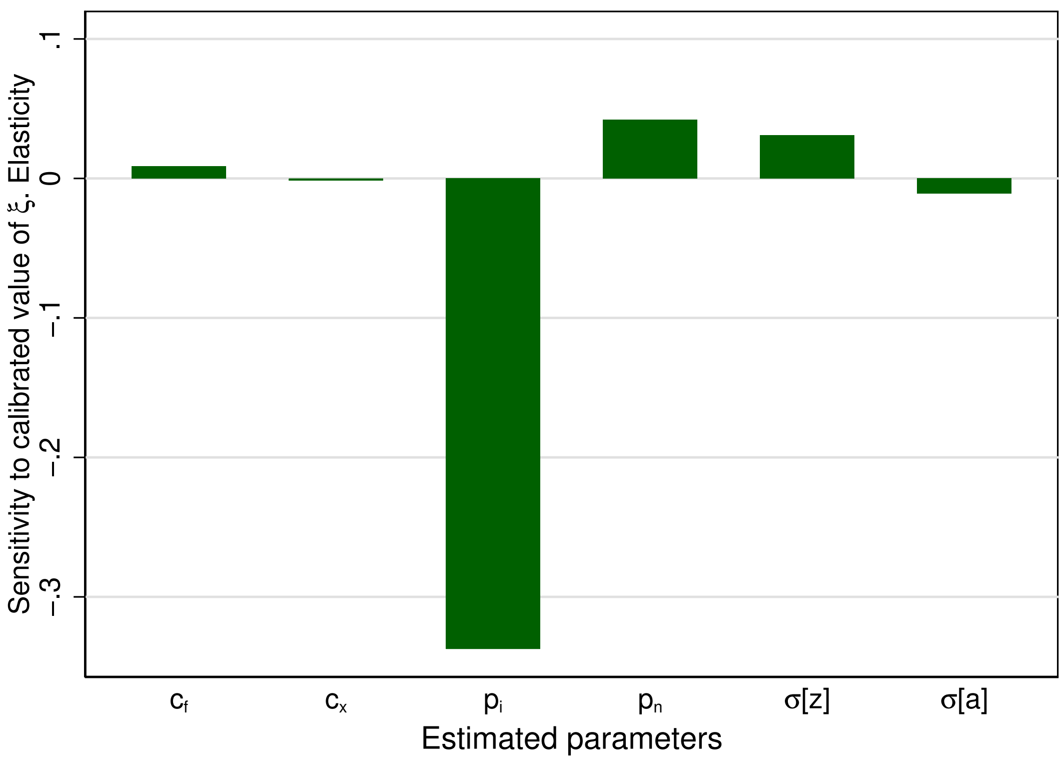
NOTES: This figure plots how the standard deviation of log firm size (blue line) and log wages (red line) change with different value of σ_z (panel A) and σ_a (panel B), relative to their respective empirical targets.

Cross-contamination in the identification of σ_z and σ_a appears to be limited. Panel A shows that shifting σ_z significantly affects firm size dispersion, i.e., the blue line rises steeply, while it only has a limited effect on wage dispersion, i.e., the red line remains almost flat. This confirms that the former is informative about the dispersion of productivity, while the latter is not.

Viceversa, panel B shows that shifting σ_a significantly affects wage dispersion, i.e., the red line increases steeply, whereas changes in firm size dispersion are not as much pronounced and are not monotonic over the parametric space, i.e., the blue line first rises by a little amount and then declines as we increase σ_a. This confirms that the former is more informative about the dispersion of amenities compared to the latter.

Sensitivity to the revenue elasticity. In the baseline model, we calibrate the revenue elasticity of labor ξ externally, using the corresponding estimate obtained in Section 2, i.e., 0.333. Here we assess how sensitive are the estimates to this assumption.

Figure C.6: Sensitivity to ξ



NOTES: This figure reports the sensitivity of each parameter estimates as in Jørgensen (2023) to the calibrated value of the revenue elasticity of labor, ξ.

We follow Jørgensen (2023) and compute a vector of sensitivity to ξ, defined as

$$S = -(J'WJ')^{-1}J'WD,$$

where W is the weighing matrix used in the estimation, J is the Jacobian matrix, whose $(i,j)^{\text{th}}$-entry is equal to $\partial d(\vartheta(i))/\partial\vartheta(j)$, and D is a vector of partial derivatives of moment conditions with respect to ξ, whose i^{th}-entry is $\partial d(\vartheta(i))/\partial\xi$.

Figure C.6 reports the elasticities of each estimated parameters k with respect to ξ; that is,

$$s(j) = S(j)\frac{\xi}{\vartheta(k)},$$

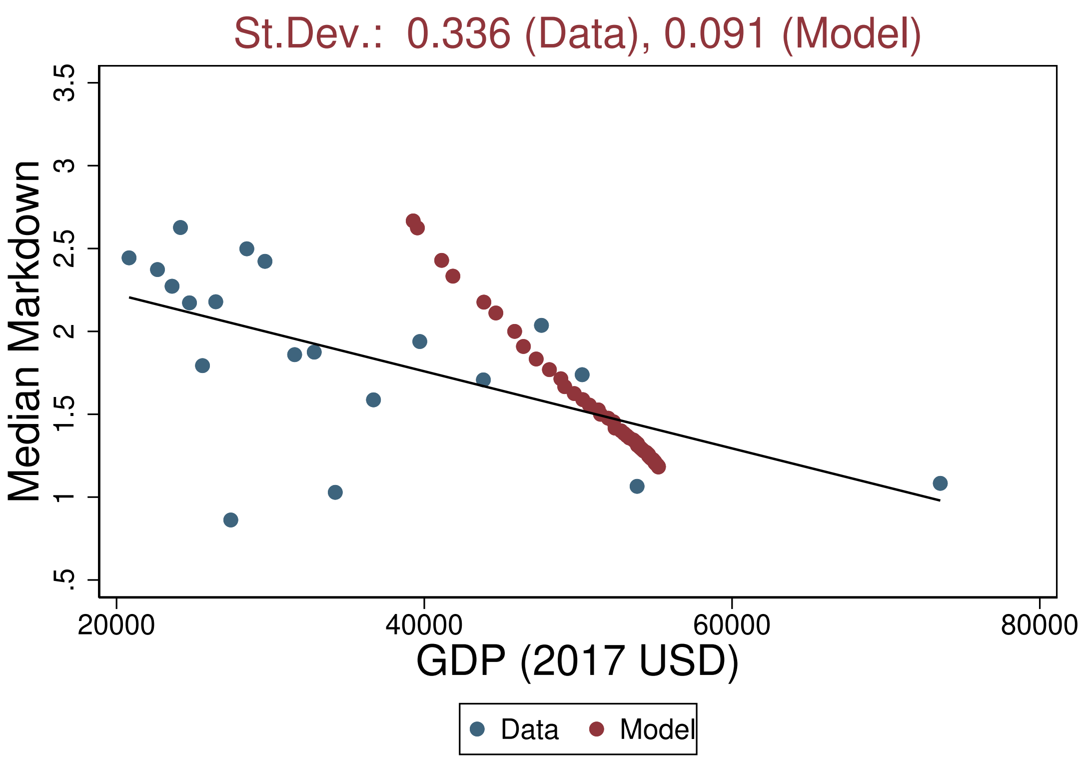
where $\vartheta(j)$ is the j^{th}-entry in the vector of estimated parameters, while $S(j)$ denotes the j^{th}-entry in the vector S.

The productivity jump for firms that innovate, p_i, is the most sensitive parameter: increasing the calibrated value of ξ by 100 percent would reduce the estimate of p_i by about 30 percent. This is because productivity-enhancing investments deliver a higher return to entrepreneurs when the revenue elasticity of labor is higher. As ξ increases, we require a lower productivity jump for firms that innovate, p_i, to match the same average cumulative firm size growth. We find that all the other parameters are relatively insensitive to the revenue elasticity of labor.

C.3 Cross-country income variation

Figure C.7 scatters the observed GDP per capita of each country in our sample against the estimated wage markdown (blue dots) and it compares them to the model-based GDP per capita obtained in counterfactual economies that feature the labor supply elasticity in line with our estimates of wage markdown reported in Section 2 (red dots). As before, all other parameters are kept fixed at their benchmark values. Both observed and simulated GDP per capita are reported in USD at 2017 constant price level.

Figure C.7: Cross-Country Income Differences: Model vs. Data



NOTES: Blue dots refer to the observed GDP p.c. from panel A of Figure 4. Red dots refer to the model-based predicted GDP p.c. obtained by changing the value of ϵ^L to match the corresponding markdown in the data. Both observed and predicted GDP p.c. are expressed in USD at 2017 constant price. SOURCE: WDI and authors' calculation.

The standard deviation in observed GDP across countries is equal to 0.336. This implies that the model accounts for about 27 percent (=0.091/0.336*100) of the observed variation in GDP across countries.

C.4 Counterfactual experiment: The Netherlands vs. Greece

Table C.3 reports various outcomes for benchmark economy (The Netherlands, in column 1) and counterfactual economy (Greece, in column 2), and compares the latter to their empirical counterparts (column 3).

Table C.3: The Netherlands vs. Greece

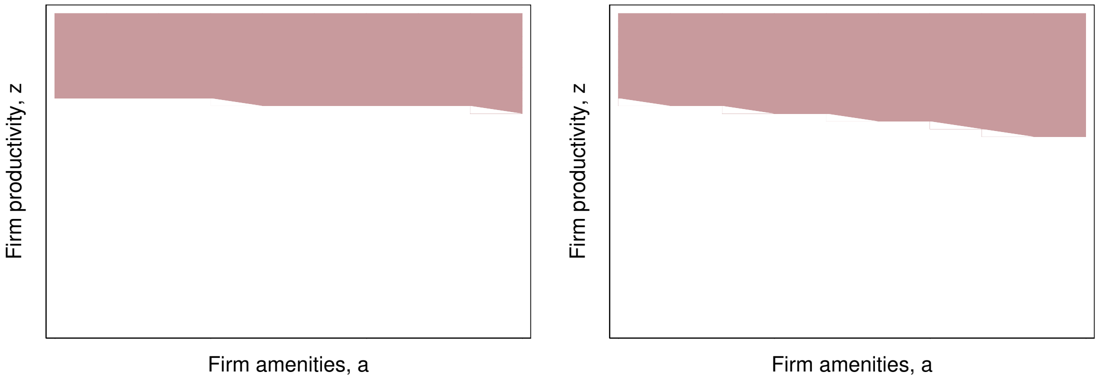
	Netherlands Benchmark (1)	Greece Counterfactual (2)	Greece Data (3)	Explained (4)
Elasticity of labor supply	3.318	0.616	-	-
Average firm size	57.13	47.45	25.79	29.1%
Average cumulative growth rate (%)	142.3	58.4	84.5	123.3%
Average firm age	31.30	29.24	22.50	27.0%
Firms investing in R&D	0.432	0.262	0.177	73.9%
TFP	1	0.825	0.500	35.0%
GDP p.c.	1	0.732	0.537	57.9%

NOTES: Column (1) reports the average life cycle firm growth, the average firm age, the share of entrepreneurs who innovate, the value of TFP and GDP p.c. in the benchmark economy (Netherlands). Column (2) shows the same model-based moments in a counterfactual economy where ϵ^L is chosen to match the median markdown observed in Greece (leaving other parameters unchanged). Column (3) reports the empirical counterparts of these moments for Greece. Column (4) reports how much (in percent) of the difference between the Netherlands and Greece is explained by differences in labor supply elasticity, and it is computed as the differences between columns (1) and (2) in this table, divided by the difference between column (1) in table C.2 and column (3) in this table. GDP p.c. in column (4) is computed as the differences between columns (1) and (2) in this table, divided by the difference between columns (1) and (3) in this table. Measured TFP is constructed using data from the Penn Word Table, v. 10.01; that is, TFP$\equiv \left(\frac{rgdpna}{(emp \times avh)^{2/3} rnna^{1/3}}\right)$, where *rgdpna* denotes real GDP at constant 2017 USD, *emp* is the number of persons engaged, *avh* is the average annual hours worked by persons engaged, and *rnna* is the capital stock at constant 2017 USD. Model-based predicted TFP refers to GDP per worker.

Figures C.8 and C.9 shows the policy functions for entrepreneurship and technology adoption in the baseline and counterfactual scenarios, i.e. The Netherlands vs. Greece. The colored areas refer to pairs of productivity and amenities where policy function takes value 1, while the white areas refers to to pairs of productivity and amenities where the policy functions takes value 0.

Figure C.8: Entrepreneurship Policy Functions

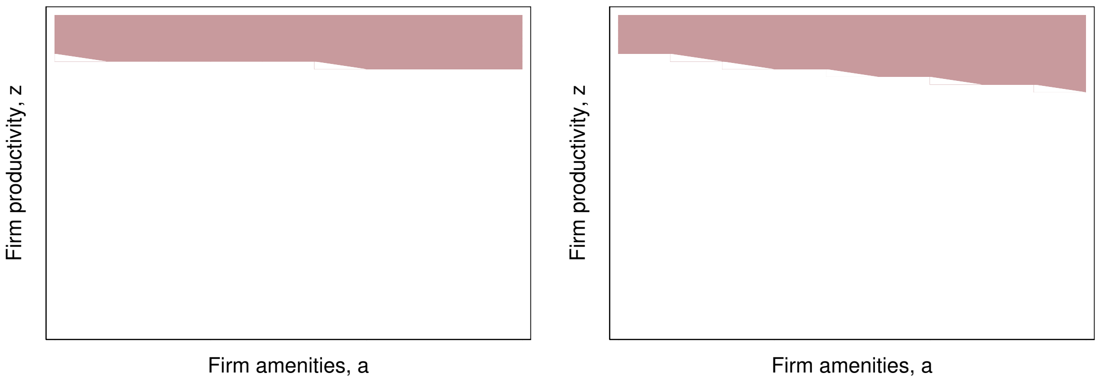
(a) Baseline **(b)** Counterfactual



NOTES: Panels A and B show the policy function for entrepreneurship in the baseline and the counterfactual respectively. The colored areas refer to pairs of productivity and amenities where policy function takes value 1, while the white areas refers to to pairs of productivity and amenities where the policy functions takes value 0.

Figure C.9: Innovation Policy Functions

(a) Baseline **(b)** Counterfactual



NOTES: Panels A and B show the policy function for investment in the baseline and counterfactual respectively. The colored areas refer to pairs of productivity and amenities where policy function takes value 1, while the white areas refers to to pairs of productivity and amenities where the policy functions takes value 0.

Table C.4 reports the model-based TFP in the baseline scenario, i.e. the Netherlands (column 1); in a counterfactual scenario where the elasticity of labor supply is chosen to match the median markdown observed in Greece (column 4); in a scenario where the elasticity of labor supply is chosen to match the markdown observed in the baseline (the Netherlands), while keeping fixed the policy functions for investment at its counterfactual level (Greece), and leaving other parameters unchanged (column 3); and in a scenario where the elasticity of labor supply is chosen to match the markdown observed in the baseline (the Netherlands), while keeping fixed the policy functions for investment and entrepreneurship at their counterfactual levels (Greece), leaving other parameters unchanged (column 4).

Table C.4: Sources of TFP losses

	Baseline (1)	Baseline with counterfactual investment (2)	Baseline with counterfactual investment and entry (3)	Counterfactual (4)
ϵ^L	3.318	3.318	3.318	0.616
TFP	1	0.849	0.835	0.825
%	0	86.3	94.3	100

NOTES: Column (1) reports the value of TFP in the benchmark economy (the Netherlands). Column (4) shows the value of TFP in the counterfactual where we set ϵ^L to match the median markdown observed in Greece. Column (2) shows the value of TFP in the baseline economy (the Netherlands) when the investment policy is fixed at its counterfactual level (Greece). Column (3) shows the value of TFP in the baseline economy (the Netherlands) when both investment and entrepreneurship policy functions are fixed at their counterfactual level (Greece).

C.5 Role of productivity-amenity correlation.

In the baseline model, new entrants draw a tuple of characteristics (z, a) from two independent log-normal distributions, $\Psi_z(z)$, and $\Psi_a(a)$, with mean 0 and variances σ_z^2 and σ_a^2, respectively. In this section, we solve a version of the model where new entrants draw their initial productivity and amenities from a bivariate zero-mean log-normal distribution, $\Psi(z, a)$, with variances equal to σ_z^2 and σ_a^2, and correlation σ_{za}.

We estimate σ_{za} using indirect inference. Specifically, using the equilibrium solution of the model, we simulate a large panel of agents, and construct a model-based conditional correlation between firm-specific (log) wage and satisfaction premia, $\hat{\kappa}$, as

$$\hat{\kappa} = \frac{\text{cov}(\log w_{jt}, a_{jt})}{\text{var}(\log w_{jt})}$$

where a_{jt} and w_{jt} denote amenities and wages of firm j at time t in the simulated panel. Therefore, we choose σ_{za} such that $\hat{\kappa} = 0.622$. All other moments are kept the same as in the baseline version. Table C.5 reports the set of estimates, empirical moments used as targets and simulated counterparts. We obtain a value for σ_{za} of 0.296.

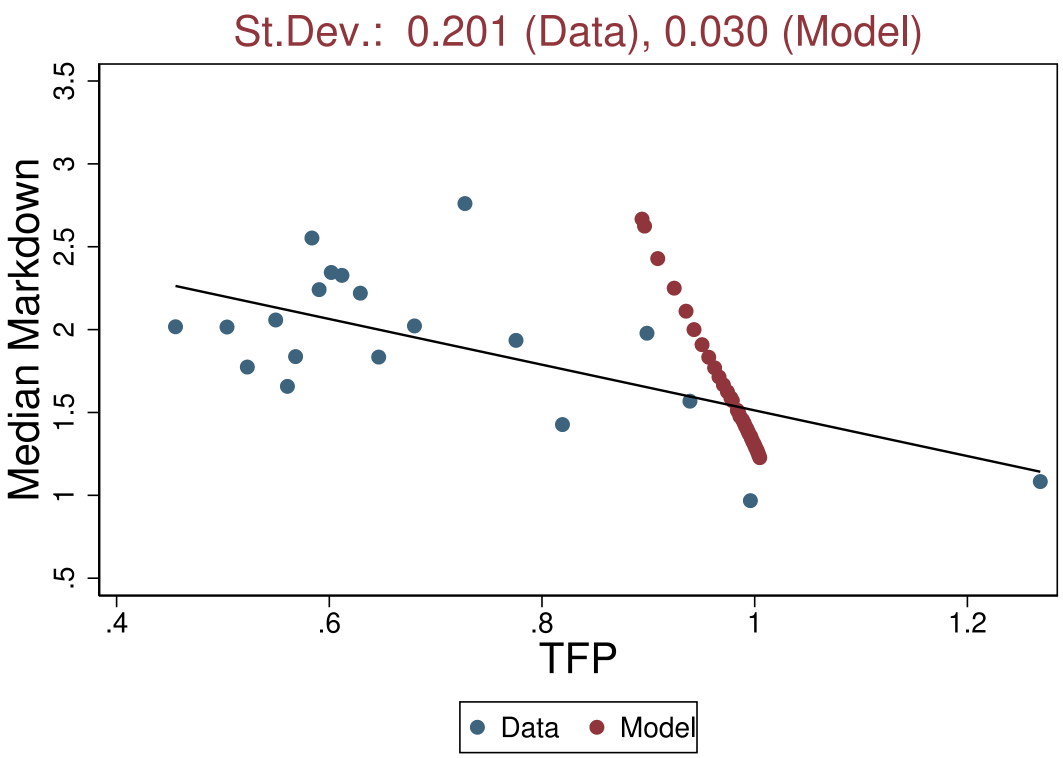
Table C.5: Estimates and Targets - Model with Correlated Amenities

Parameters	Description	Estimates	Targets	Data	Model
c_f	Operating costs	78.97	Average firm size	59.02	57.95
c_x	Innovation costs	227.3	Share of firms investing in R&D	0.407	0.424
p_i	Growth prob. of innovators	0.588	Average cumulative growth rate (%)	152.6	140.9
p_n	Growth prob. of non-innovators	0.540	Average firm age	30.13	30.67
σ_z	Productivity dispersion	2.832	Log firm size dispersion	0.988	1.103
σ_a	Amenities dispersion	0.956	Log wage dispersion	0.418	0.425
σ_{za}	Correlation productivity-amenities	0.296	(log) Wage-amenities correlation	0.622	0.622

NOTES: The table shows the list of estimated parameters, their point estimates, the value of empirical targets, and model-based simulated moments. In this version of the model, the correlation between productivity and amenities is internally estimated to match the relation between firms' wage and job amenities estimated in Sockin (2024).

Endowed with new estimates, we repeat the same analysis of Section 5.1. Figure C.10 scatters a measure of TFP of each country in our sample against the estimated wage markdown (blue dots) and compares it to the model-based

Figure C.10: Cross-Country Productivity Differences: Model with Correlated Amenities vs. Data



NOTES: Blue dots refer to measured TFP, constructed using data from the Penn Word Table, v. 10.01; that is, TFP$\equiv \left(\frac{rgdpna}{(emp \times avh)^{2/3} rnna^{1/3}}\right)$, where *rgdpna* denotes real GDP at constant 2017 USD, *emp* is the number of persons engaged, *avh* is the average annual hours worked by persons engaged, and *rnna* is the capital stock at constant 2017 USD. Red dots refer to model-based predicted TFP (i.e., GDP per worker), obtained by changing the value of ϵ^L to match the corresponding markdown in the data. Both variables are reported as fraction of the value measured in the Netherlands. In this version of the model, the correlation between productivity and amenities is internally estimated to match the relation between firms' (log) wage and job amenity estimated in Sockin (2024). SOURCE: WDI and authors' calculation.

TFP obtained in counterfactual economies that feature counterfactual labor supply elasticities. As before, all other parameters are kept fixed at their benchmark values. Both measured and simulated TFP are reported as fraction of the value measured for the Netherlands. We find a standard deviation in model-based TFP of 0.03. Which implies that the model can explain about 15 percent (=0.030/0.201*100) of the observed variation of measured TFP across countries.

C.6 Role of entry and investment costs.

In this section, we discuss a version of the model where both entry and investment costs are defined in terms of labor. In this model, the value to an agent i of being an entrepreneur is still equal to:

$$V(z_i, a_i) = \max\{V^I(z_i, a_i), V^N(z_i, a_i)\}$$

where $V^I(z_i, a_i)$ is the value of investing in productivity, equal to

$$\begin{aligned}V^I(z_i, a_i) &= \epsilon^L \ln(\pi^I(z_i, a_i)) + a_i \\ &+ \beta(1-\delta_w)\big(p_i \max\{V(z_{i+}, a_i), \tilde{U}(z_{i+}, a_i)\} + (1-p_i)\max\{V(z_{i-}, a_i), \tilde{U}(z_{i-}, a_i)\}\big)\end{aligned}$$

while $V^N(z_i, a_i)$ is value of not investing,

$$\begin{aligned}V^N(z_i, a_i) &= \epsilon^L \ln(\pi^N(z_i, a_i)) + a_i \\ &+ \beta(1-\delta_w)\big(p_n \max\{V(z_{i+}, a_i), \tilde{U}(z_{i+}, a_i)\} + (1-p_n)\max\{V(z_{i-}, a_i), \tilde{U}(z_{i-}, a_i)\}\big).\end{aligned}$$

On the other hand, as opposed to the baseline version, the profit flows net of fixed costs for entrepreneurs investing in productivity are the solution to the following problem:

$$\begin{aligned}&\max_{w_j} \pi_j^I(z_j, a_j) = z_j L_j^{\xi} - w_j(L_j + c_z + c_f) \\ &\textit{subject to} \quad L_j + c_z + c_f = L\Theta \exp\left(\epsilon^L \ln(w_j) + a_j\right),\end{aligned}$$

while the profit flows net of fixed costs for entrepreneurs who do not invest solve a different problem:

$$\begin{aligned}&\max_{w_j} \pi_j^N(z_j, a_j) = z_j L_j^{\xi} - w_j(L_j + c_f) \\ &\textit{subject to} \quad L_j + c_f = L\Theta \exp\left(\epsilon^L \ln(w_j) + a_j\right).\end{aligned}$$

All the other details of the model remain the same as in Section 3.

A stationary recursive equilibrium for this economy is a list of value functions $V(z_i, a_i)$, $U(z_i, a_i, z_j, a_j)$ and $\tilde{U}(z_i, a_i)$, an associated entrepreneurship policy function $\rho^e(z_i, a_i)$ and innovation policy function $\rho^z(z_i, a_i)$, a wage schedule $W(z_i, a_i)$, an allocation of labor $L(z_i, a_i)$, an aggregate measure of workers L, a distribution of agents over productivity and amenities, $\Omega(z_i, a_i)$, and distributions of wage workers and entrepreneurs over productivity and amenities, $\phi(z_i, a_i)$ and $\psi(z_i, a_i)$, such that:

- The labor supply to firm i, $L(z_i, a_i)$, satisfies equation (3);
- $\rho^e(z_i, a_i)$ and $\rho^z(z_i, a_i)$ solve the entrepreneurial and the innovation choices, and the value functions $V(z_i, a_i)$, $U(z_i, a_i, z_j, a_j)$ and $\tilde{U}(z_i, a_i)$ attain their maxima;
- The aggregate measure of workers is consistent with the entrepreneurial choices:

$$
\begin{aligned}
L &= \int_{\mathcal{Z}\times\mathcal{A}} [L(z_i, a_i) + c_f + c_z \rho^z(z_i, a_i)] \psi(z_i, a_i) dz_i da_i \\
&= \int_{\mathcal{Z}\times\mathcal{A}} (1 - \rho^e(z_i, a_i)) \Omega(z_i, a_i) dz_i da_i.
\end{aligned}
$$

- The distribution of agents over productivity and amenities, $\Omega(z_i, a_i)$ is stationary and replicates itself through entry and exit, and the policy functions, as in equations (14), (15) and (16), defined in Appendix B.2.
- The distributions of wage workers and entrepreneurs over productivity and amenities are stationary and defined as

$$
\phi(z_i, a_i) = \frac{(1 - \rho^e(z_i, a_i)) \Omega(z_i, a_i)}{\int_{\mathcal{Z}\times\mathcal{A}} (1 - \rho^e(z_i, a_i)) \Omega(z_i, a_i) dz_i da_i},
$$

  and

$$
\psi(z_i, a_i) = \frac{\rho^e(z_i, a_i) \Omega(z_i, a_i)}{\int_{\mathcal{Z}\times\mathcal{A}} \rho^e(z_i, a_i) \Omega(z_i, a_i) dz_i da_i},
$$

  respectively.

Table C.6 reports the set of estimates for this version of the model, empirical moments used as targets and simulated counterparts.

Table C.6: Estimates and Targets - Model with Costs defined in terms of Labor

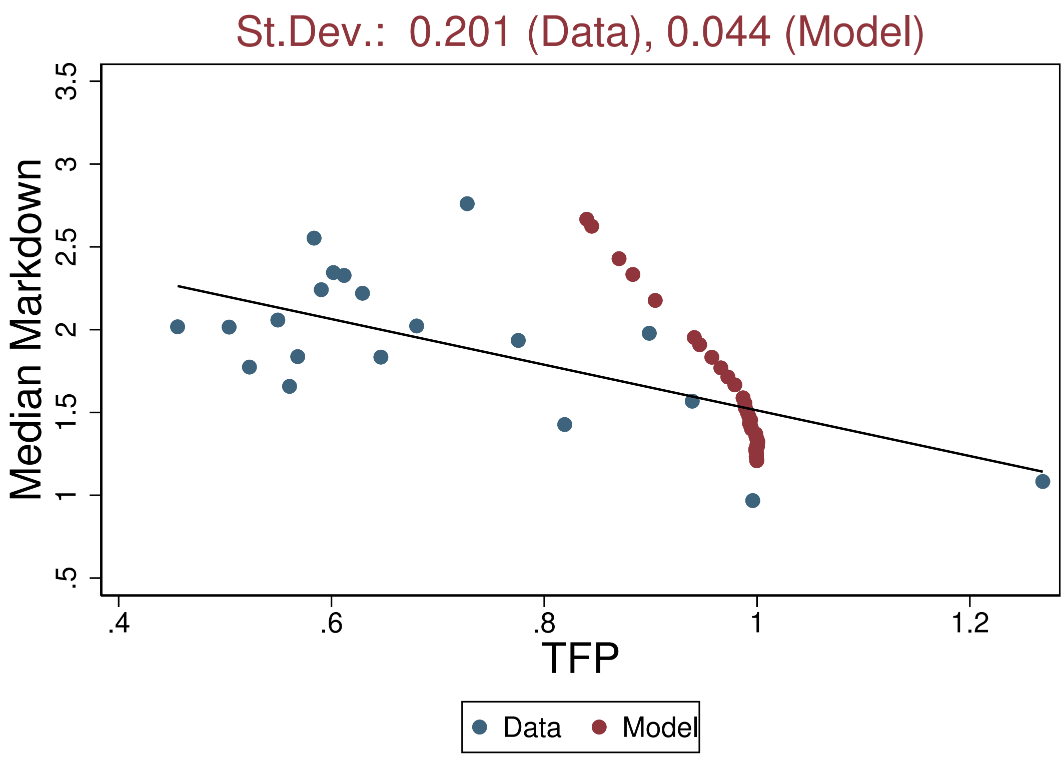
Parameters	Description	Estimates	Targets	Data	Model
c_f	Operating costs	2.274	Average firm size	59.02	58.10
c_x	Innovation costs	18.42	Share of firms investing in R&D	0.407	0.432
p_i	Growth prob. of innovators	0.603	Average cumulative growth rate (%)	152.6	150.9
p_n	Growth prob. of non-innovators	0.526	Average firm age	30.13	29.18
σ_z	Productivity dispersion	2.789	Log firm size dispersion	0.988	1.274
σ_a	Amenities dispersion	0.669	Log wage dispersion	0.418	0.475

NOTES: The table shows the list of estimated parameters, their point estimates, the value of empirical targets, and model-based simulated moments. In this version of the model, operating costs and investment costs are expressed in terms of labor.

Figure C.11 scatters the total factor productivity of each country in our sample against the estimated wage markdown (blue dots) and compares it to the model-based TFP obtained in counterfactual economies that feature counterfactual labor supply elasticities. As before, all other parameters are kept fixed at their benchmark values. Both measured and simulated TFP are reported as fraction of the value measured for the Netherlands.

Using this version of the model, we find a standard deviation in model-based total factor productivity of 0.044. This value implies that the model can explain about 22 percent (=0.044/0.201*100) of the observed variation in TFP across countries.

Figure C.11: Cross-Country Productivity Differences: Model with Costs defined in terms of Labor vs. Data



NOTES: Blue dots refer to measured TFP, constructed using data from the Penn Word Table, v. 10.01; that is, TFP$\equiv \left(\frac{rgdpna}{(emp \times avh)^{2/3} rnna^{1/3}}\right)$, where *rgdpna* denotes real GDP at constant 2017 USD, *emp* is the number of persons engaged, *avh* is the average annual hours worked by persons engaged, and *rnna* is the capital stock at constant 2017 USD. Red dots refer to model-based predicted TFP (i.e., GDP per worker), obtained by changing the value of ϵ^L to match the corresponding markdown in the data. Both variables are reported as fraction of the value measured in the Netherlands. In this version of the model, operating costs and investment costs are expressed in terms of labor. SOURCE: WDI and authors' calculation.

C.7 Estimation using imputed moments

The WB-ES data only provides information for firms with at least 5 employees, making the sample biased towards relatively larger firms. In this section, we re-estimate our baseline model using firm-level moments that are imputed to cover the entire span of *existing firms* in the Netherlands.

To this purpose, we impute the average firm size, the (log) firm size dispersion, and the share of firms investing in R&D as follows. The average firm size is imputed by fitting a Pareto mean to the sample average among *observed firms* from the WB-ES, imposing a left truncation at 5 employees. Specifically, let the firm size be a random variable following a Pareto distribution, with scale parameter $x_0 > 0$ and shape parameter $\alpha > 1$, i.e., firm size $\sim \text{Par}(x_0, \alpha)$. Then, its expected value is equal to:

$$\frac{\alpha}{\alpha - 1} x_0.$$

With a left truncation at 5 employees, we can set $x_0 = 5$, and use the method of moments to estimate α by matching a sample average among *observed firms* of 59.02 employees. We obtain a value of $\alpha = 1.0926$. Therefore, we impute the average firm size for *existing firms* by computing the expected value of a Pareto random variable with $\alpha = 1.0926$ and $x_0 = 1$. We obtain a value of 11.80 employees.

Using our estimate of the shape parameters α, we impute the dispersion in (log) firm size by acknowledging that a log Pareto distribution is equal to an exponential distribution with a rate equal to the Pareto shape. Specifically, let again the firm size be a Pareto random with scale parameter $x_0 = 1$ and shape parameter $\alpha = 1.0926$. Then the (log) of firm size follows an exponential distribution with a rate $\lambda = \alpha = 1.0926$. The standard deviation of an exponential r.v. is equal to $\sqrt{1/\lambda^2}$, implying a dispersion of (log) firm size equal to 0.915.

Finally, the overall share of innovating firms is obtained multiplying the imputed share of firms with more than 5 employees times the share of investors among *observed firms* from the WB-ES, under the assumption that firms with less than 5 employees do not engage in productivity-enhancing investment. Specifically, the probability that a Pareto random variable is larger or equal than a value x is $(x_0/x)^{\alpha}$. Using $x_0 = 1$ and $\alpha = 1.0926$, it follows that the probability of observing a firm with at least 5 employees is equal to 0.1723. Under the assumption that firms with less than 5 employees do not innovate, and given a share of *observed firms* that innovate of 0.407, the share of *existing firms* performing R&D will be equal to 0.1723 $\times$ 0.407 = 0.0701.

In the estimation, we keep the remaining three moments, i.e., the average cumulative growth rate, the average firm age, and the (log) wage dispersion, unchanged, and again apply the same data truncation to our simulated moments. Table C.7 summarizes the imputed moments used as targets, their simulated counterparts, and reports the parameter estimates.

Table C.7: Estimates and Targets - Baseline model estimated using imputed moments

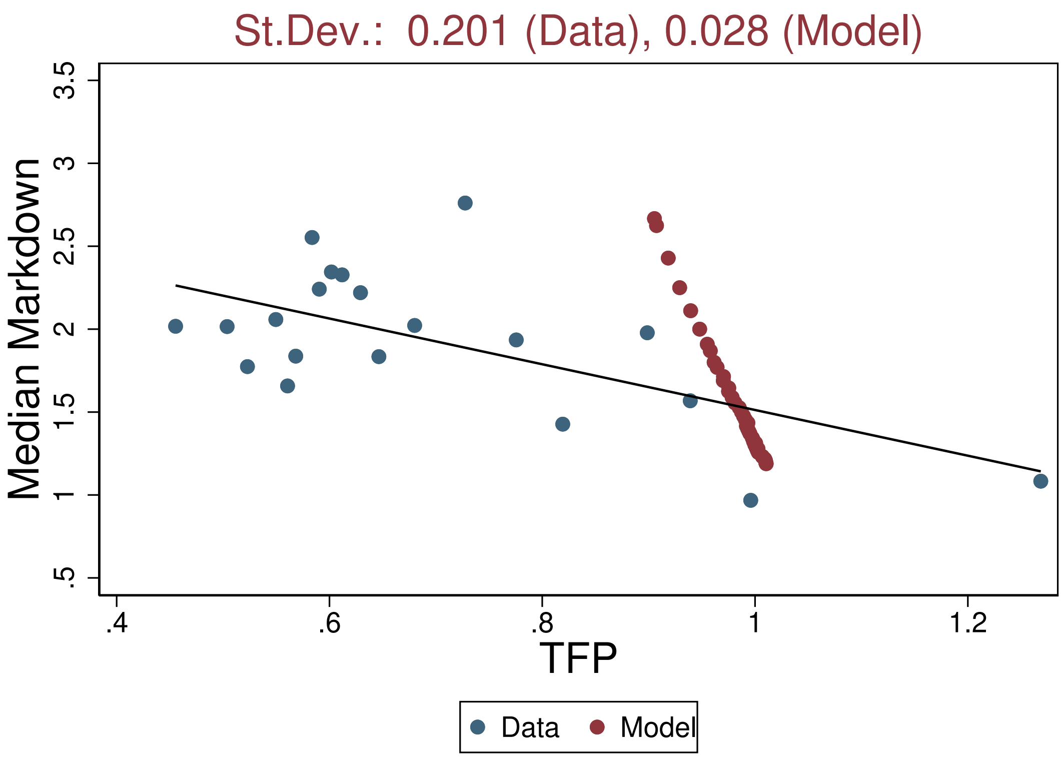
Parameters	Description	Estimates	Targets	Data	Model
			Imputed		
c_f	Operating costs	5.478	Average firm size	11.80	11.18
c_x	Innovation costs	98.34	Share of firms investing in R&D	0.070	0.066
σ_z	Productivity dispersion	2.012	Log firm size dispersion	0.915	1.064
			Observed		
p_i	Growth prob. of innovators	0.639	Average cumulative growth rate (%)	152.7	124.5
p_n	Growth prob. of innovators	0.542	Average firm age	30.13	27.78
σ_a	Amenities dispersion	0.838	Log wage dispersion	0.418	0.403

NOTES: The table shows the list of estimated parameters, their point estimates, the value of empirical targets, and model-based simulated moments. In this version of the model, we target the imputed average firm size, the imputed (log) firm-size dispersion, and the imputed share of firms investing.

Figure C.12 scatters total factor productivity of each country in our sample against the estimated wage markdown (blue dots) and compares it to the simulated TFP obtained in counterfactual economies that feature counterfactual labor supply elasticities. As before, all other parameters are kept fixed at their benchmark values. Both measured and simulated TFP are reported as fraction of the value measured for the Netherlands.

We find a standard deviation in model-based total factor productivity of 0.028. It follows that the model can account for about 14 percent (=0.028/0.201*100) of the observed variation in TFP across countries.

Figure C.12: Cross-Country Productivity Differences: Model estimated using imputed moments vs. Data



NOTES: Blue dots refer to measured TFP, constructed using data from the Penn Word Table, v. 10.01; that is, TFP$\equiv\left(\frac{rgdpna}{(emp\times avh)^{2/3}rnna^{1/3}}\right)$, where *rgdpna* denotes real GDP at constant 2017 USD, *emp* is the number of persons engaged, *avh* is the average annual hours worked by persons engaged, and *rnna* is the capital stock at constant 2017 USD. Red dots refer to model-based predicted TFP (i.e., GDP per worker), obtained by changing the value of ϵ^L to match the corresponding markdown in the data. Both variables are reported as fraction of the value measured in the Netherlands. In this version of the model, we target the imputed average firm size, the imputed (log) firm-size dispersion, and the imputed share of firms investing. SOURCE: WDI and authors' calculation.

C.8 Estimation using alternative identification

In this section we estimate the baseline model using an alternative identification strategy. In particular, we make the model over-identified by targeting more moments than parameters. First, we replace the average firm size with the shares of firms of different size (≤ 20, 21-100, and > 100 employees). Second, we replace the average firm age with the shares of firms of different ages (≤ 30 y.o., between 30-60 y.o., and > 60 y.o.), and the overall share of firms investing in R&D with the shares of firms investing in R&D by number of employees (≤ 20, 21-100, and > 100 employees). Finally, we keep the (log) firm size dispersion and the (log) wage dispersion in the set of moments, whereas we replace the average cumulative firm size growth rate with the average annualized growth rate, computed as in Section 2, and deflated by the annual growth rate in the average firm size observed in Netherlands in the past 20 years.

Specifically, we compute the latter as follows. First, we construct the overall employment by multiplying the employment-population ratio from the World Bank Development Indicator database (variable "SL.EMP.TOTL.SP.ZS") and the working age population (variable "SL.TLF.TOTL.IN"). Hence, we construct wage and salary employment by subtracting those who are self-employed (variable "SL.EMP.SELF.ZS"). We find an annual growth rate in wage and salary employment of about 0.35 percent. Then, we use data from the Central Bureau For Statistics (CBS) on the number of active firms (retrieved from https://www.cbs.nl/en-gb/news/2023/06/enterprise-population-keeps-growing) to compute the annual growth rate in active firms. We find a value of 0.90 percent. Finally we obtain an average annual growth rate in average firm size of $0.35 - 0.90 = -0.55$ percent, and use this value to deflate the annualized growth rate in firm size obtained using the WB-ES data, i.e. 7.34 percent. We obtain and target an annualized deflated growth rate of $7.34 - (-0.55) = 7.89$ percent.

Table C.8 summarizes the parameter estimates together with empirical moments used as targets and their simulated counterparts. While there is no one-to-one mapping between parameters and targets, the identification strategy reflects the one used in the baseline estimation.

Figure C.13 relates the total factor productivity of each country in our sample

Table C.8: Estimates and Targets - Overidentified baseline model

Parameters	Description	Estimates	Target	Data	Model
c_f	Operating costs	14.76	Share of firms with ≤ 20 employees	0.576	0.605
c_x	Innovation costs	57.96	Share of firms with 21-100 employees	0.362	0.393
p_i	Growth prob. of innovators	0.623	Share of firms with > 100 employees	0.062	0.002
p_n	Growth prob. of innovators	0.546	Share of firms, ≤ 30 y.o.	0.615	0.685
σ_z	Productivity dispersion	2.024	Share of firms, 30-60 y.o.	0.286	0.191
σ_a	Amenities dispersion	0.778	Share of firms, > 60 y.o.	0.098	0.124
			Share of firms with ≤ 20 employees investing in R&D	0.331	0.333
			Share of firms with 21-100 employees investing in R&D	0.504	0.457
			Share of firms with > 100 employees investing in R&D	0.944	1.000
			Average annualized (deflated) growth rate (%)	7.891	6.701
			Log firm size dispersion	0.988	0.702
			Log wage dispersion	0.418	0.361

NOTES: This table reports the list of estimated parameters, their point estimates, the value of the empirical targets and their model-based simulated counterparts for the overidentified version of the baseline model.

(blue dots) and the simulated TFP (red dots) across counterfactual economies against wage markdowns, obtained keeping all other parameters fixed at their benchmark values. Both measured and simulated TFP are reported as fraction of the value measured for the Netherlands.

When the model is over-identified, we find a standard deviation in model-based TFP of 0.023. It follows the model can account for about 11 percent (=0.023/0.201*100) of the observed variation in TFP across countries.

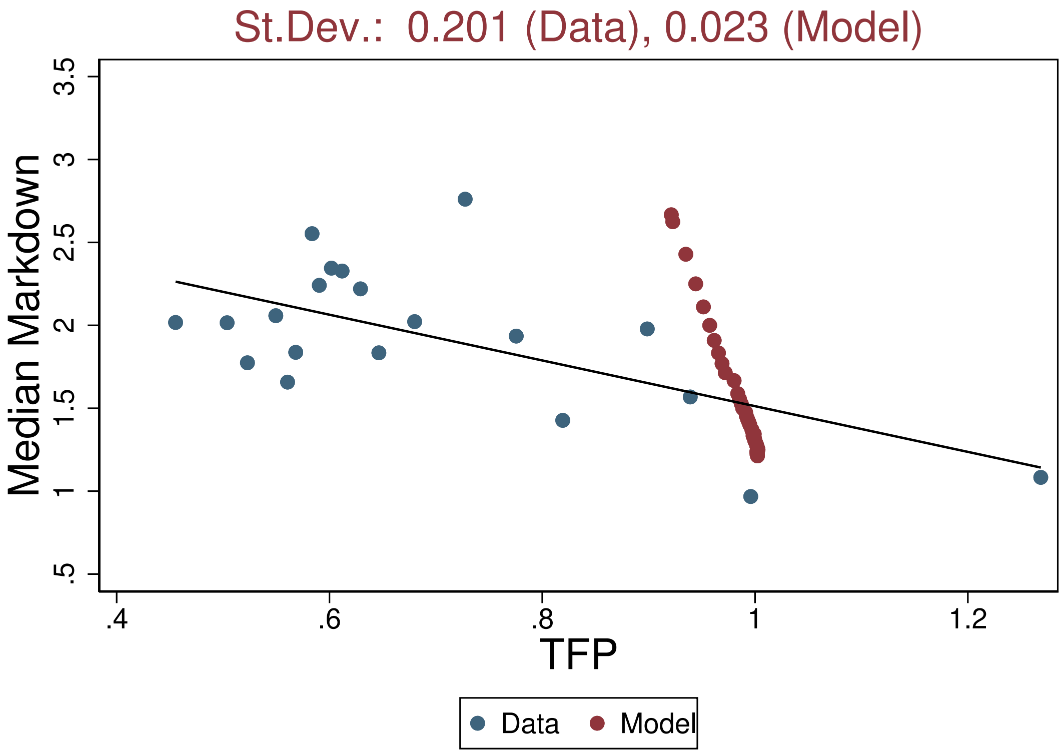
References

Amodio, Francesco, Emanuele Brancati, Peter Brummund, Nicolás de Roux, and Michele Di Maio, "Global Labor Market Power," *Unpublished Manuscript*, 2024. Retrieved from https://www.econstor.eu/bitstream/10419/282656/1/dp16529.pdf.

Jørgensen, Thomas H, "Sensitivity to calibrated parameters," *Review of Economics and Statistics*, 2023, *105* (2), 474–481.

Levinsohn, James and Amil Petrin, "Estimating production functions using inputs to control for unobservables," *The review of economic studies*, 2003, *70* (2), 317–341.

Figure C.13: Cross-Country Productivity Differences: Over-identified model vs. Data



NOTES: Blue dots refer to measured TFP, constructed using data from the Penn Word Table, v. 10.01; that is, TFP$\equiv \left(\frac{rgdpna}{(emp \times avh)^{2/3} rnna^{1/3}}\right)$, where *rgdpna* denotes real GDP at constant 2017 USD, *emp* is the number of persons engaged, *avh* is the average annual hours worked by persons engaged, and *rnna* is the capital stock at constant 2017 USD. Red dots refer to model-based predicted TFP (i.e., GDP per worker), obtained by changing the value of ϵ^L to match the corresponding markdown in the data. Both variables are reported as fraction of the value measured in the Netherlands. This version of the model is over-identified, as we target the share of firms of different sizes (# employees), the share of firms of different ages, and the share of firms investing in R&D by size, the average annualized (deflated) growth rate, (log) firm size dispersion and (log) wage dispersion. SOURCE: WDI, CBS and authors' calculation.

Sockin, Jason, "Show Me the Amenity: Are Higher-Paying Firms Better All Around?," *Unpublished Manuscript*, 2024. Retrieved from https://www.researchgate.net/profile/Jason-Sockin-2/publication.