Appendix to the paper "A Unified Theory of Firm Selection and Growth"*

Costas Arkolakis†
Yale University, NBER & CESifo

March 16, 2015

Abstract

This appendix provides a set of additional results and proofs for the paper "A Unified Theory of Firm Selection and Growth"

Contents

1 Alternative Proof for the Double Pareto Distribution **3**

2 Existence and Uniqueness of the Equilibrium **7**

3 Finite Integrals with O-U **14**

4 Proof for expected time to reach a certain size **16**

5 Additional result for the cohort survival rate **19**

*I am indebted to Jonathan Eaton, Marcela Eslava, Maurice Kugler, and James Tybout for kindly sharing moments from their Colombian exporting data. Olga Timoshenko provided outstanding research assistance. All remaining errors are mine.

†Contact: Department of Economics, Yale University, 28 Hillhouse ave., New Haven, CT, 06511. Email: costas.arkolakis@yale.edu. web: http://www.econ.yale.edu/˜ka265/research.htm

6 Hazard Rates **21**
6.1 Firm Hazard Rates 21
6.2 Cohort Hazard Rates 22

7 Firm and Cohort Sales **25**
7.1 Firm Expected Sales 25
7.2 Cohort Market Shares 26
7.3 Size and Age in the US Data 27

8 Detailed derivations regarding proposition 2 of the paper **28**

9 An Extension: Multivariate Brownian Motion **30**

10 Numerical Procedure for Computing the Equilibrium **33**

11 Robustness: Size Distribution **33**

12 Appendix: Properties of the Normal Distribution **36**

1 Alternative Proof for the Double Pareto Distribution

Given the assumptions stated in the main paper we will show an alternative way to illustrate the emergence of a double Pareto distribution from the assumptions of continuous entry at a certain rate and the stochastic process of a geometric Brownian motion. To do so we will look directly at the distribution of sizes of the cross-section of ideas with different productivities which involves properly weighting the distributions of ideas of different generations.

We assume that all ideas initiate with productivity $\bar{z}$. The logarithm of a geometric Brownian motion follows a simple Brownian motion and for simplicity we will only consider the process for the natural logarithm of the productivities. We let the drift be defined as μ and the standard deviation as σ_z. The probability density for ideas with a productivity ϕ, that start from a point $\bar{z}$ and follow a simple Brownian motion is simply

$$
\begin{aligned}
p(\phi, a|\bar{z}) &= \frac{1}{\sigma_z\sqrt{a2\pi}} \exp\left\{-\frac{\left(\frac{\phi-\bar{z}-\mu a}{\sigma_z\sqrt{a}}\right)^2}{2}\right\} = \\
& \frac{1}{\sigma_z\sqrt{a2\pi}} \exp\left\{-\frac{(\phi-\bar{z}-\mu a)^2}{\sigma_z^2 2a}\right\} .
\end{aligned}
$$

with support $(-\infty, +\infty)$.

The entry rate of new ideas is g_B which, in the balanced growth path, can be interpreted as the rate of overall growth of ideas plus the rate of exogenous death of ideas. Therefore, in order to find the cross-sectional distribution for ϕ's we have to integrate the above density by putting the appropriate weigth to each cohort of ideas. Namely we need a weight $\exp\{-g_B a\}$ for ideas of age a. Therefore the density of productivities ϕ is given bt:

$$g\left(\phi|\bar{z}\right)=\int_{0}^{+\infty}\frac{e^{-g_{B}a}}{\sigma_{z}\sqrt{a2\pi}}\exp\left\{-\frac{\left(\phi-\bar{z}-\mu a\right)^{2}}{\sigma_{z}^{2}2a}\right\}da=$$

$$\int_{0}^{+\infty}\frac{e^{-g_{B}(1-\alpha)a}}{\sigma_{z}\sqrt{a2\pi}}\exp\left\{-\frac{\left(\phi-\bar{z}\right)^{2}-2\left(\phi-\bar{z}\right)\mu a+\left(\mu a\right)^{2}}{\sigma_{z}^{2}2a}\right\}da=$$

$$\frac{1}{\sigma_{z}}\exp\left\{\frac{\left(\phi-\bar{z}\right)\mu}{\sigma_{z}^{2}}\right\}\int_{0}^{+\infty}\frac{1}{\sqrt{a2\pi}}\exp\left\{-\left(\frac{\left(\phi-\bar{z}\right)^{2}}{\sigma_{z}^{2}2a}+\frac{\left(\mu^{2}+\sigma_{z}^{2}2g_{B}\right)\left(a\right)^{2}}{\sigma_{z}^{2}2a}\right)\right\}da=$$

$$\frac{1}{\sigma_{z}}\exp\left\{\frac{\left(\phi-\bar{z}\right)\mu}{\sigma_{z}^{2}}\right\}\int_{0}^{+\infty}\frac{1}{\sqrt{a2\pi}}\exp\left\{-\left(\frac{\sqrt{\left(\phi-\bar{z}\right)^{2}}}{\sigma_{z}\sqrt{2a}}-\frac{\sqrt{\left(\mu^{2}+\sigma_{z}^{2}2g_{B}\right)a}}{\sigma_{z}\sqrt{2}}\right)^{2}-\frac{\sqrt{\left(\phi-\bar{z}\right)^{2}}}{\sigma_{z}}\frac{\sqrt{\left(\mu^{2}+\sigma_{z}^{2}2g_{B}\right)}}{\sigma_{z}}\right\}$$

$$\frac{1}{\sigma_{z}}\exp\left\{\frac{\left(\phi-\bar{z}\right)\mu}{\sigma_{z}^{2}}-\frac{\sqrt{\left(\phi-\bar{z}\right)^{2}}}{\sigma_{z}}\frac{\sqrt{\left(\mu^{2}+\sigma_{z}^{2}2g_{B}\right)}}{\sigma_{z}}\right\}\int_{0}^{+\infty}\frac{1}{\sqrt{a2\pi}}\exp\left\{-\left(\frac{\sqrt{\left(\phi-\bar{z}\right)^{2}}}{\sigma_{z}\sqrt{2a}}-\sqrt{\frac{\left(\mu^{2}+\sigma_{z}^{2}2g_{B}\right)a}{\sigma_{z}^{2}2}}\right)^{2}\right\}$$

Now notice that if we set $x=\sqrt{a}$, $dx=\frac{1}{2}\frac{1}{\sqrt{a}}da$ we have

$$\int_{0}^{+\infty}\frac{1}{\sqrt{a2\pi}}\exp\left\{-\frac{\left(\frac{\sqrt{\left(\phi-\bar{z}\right)^{2}}}{\sqrt{a}}-\sqrt{a}\sqrt{\left(\mu^{2}+\sigma_{z}^{2}2g_{B}\right)}\right)^{2}}{\sigma_{z}^{2}2}\right\}da\quad=$$

$$\int_{0}^{+\infty}\frac{2}{\sqrt{2\pi}}\exp\left\{-\frac{\left(\frac{\sqrt{\left(\phi-\bar{z}\right)^{2}}}{x}-x\sqrt{\mu^{2}+\sigma_{z}^{2}2g_{B}}\right)^{2}}{\sigma_{z}^{2}2}\right\}dx\quad=$$

and thus

$$g\left(\phi|\bar{z}\right)=\frac{1}{\sigma_{z}}\exp\left\{\frac{\left(\phi-\bar{z}\right)\mu}{\sigma_{z}^{2}}\right\}\int_{0}^{+\infty}\frac{1}{\sqrt{a2\pi}}\exp\left\{-\left(\frac{\left(\phi-\bar{z}\right)^{2}}{\sigma_{z}^{2}2a}+\frac{\left(\mu^{2}+\sigma_{z}^{2}2g_{B}\right)\left(a\right)^{2}}{\sigma_{z}^{2}2a}\right)\right\}da=$$
$$\exp\left\{\frac{\left(\phi-\bar{z}\right)\mu}{\sigma_{z}^{2}}\right\}\frac{2}{\sigma_{z}\sqrt{2\pi}}\int_{0}^{+\infty}\exp\left\{-\frac{\left(\phi-\bar{z}\right)^{2}}{\sigma_{z}^{2}2x^{2}}-\frac{\left(\mu^{2}+\sigma_{z}^{2}2g_{B}\right)x^{2}}{\sigma_{z}^{2}2}\right\}dx.$$

Now we have the following result for the last integral

$$\int_{0}^{+\infty}e^{-a/x^{2}-bx^{2}}dx=\frac{1}{4\sqrt{b}}\left(\begin{array}{c}\sqrt{\pi}\left[e^{-2\sqrt{a}\sqrt{b}}\left(\operatorname{erf}\left(\sqrt{b}x-\frac{\sqrt{a}}{x}\right)+1\right)\right]+\\ e^{2\sqrt{a}\sqrt{b}}\left(\operatorname{erf}\left(\sqrt{b}x+\frac{\sqrt{a}}{x}\right)-1\right)\end{array}\right|_{0}^{+\infty}\right)$$

where $\mathrm{erf}(x)$ is the error function, $\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$ (and $\mathrm{erf}(+\infty) = 1$, $\mathrm{erf}(-\infty) = -1$). This means that

$$\begin{aligned} \int_0^{+\infty} e^{-a/x^2 - bx^2} dx &= \frac{1}{4\sqrt{b}} \begin{pmatrix} \sqrt{\pi} e^{-2\sqrt{a}\sqrt{b}} 2 + 0 \\ 0 + 0 \end{pmatrix} \\ &= \frac{1}{2\sqrt{b}} \left(\sqrt{\pi} e^{-2\sqrt{a}\sqrt{b}} \right) \end{aligned}$$

Given the above $g(\phi|\bar{z})$ becomes

$$g(\phi|\bar{z}) = \exp\left\{ \frac{(\phi - \bar{z})\mu}{\sigma_z^2} \right\} \frac{2}{\sigma_z \sqrt{2\pi}} \int_0^{+\infty} \exp\left\{ -\frac{(\phi - \bar{z})^2}{\sigma_z^2 2x^2} - \frac{\left(\mu^2 + \sigma_z^2 2g_B\right) x^2}{\sigma_z^2 2} \right\} dx =$$

$$\exp\left\{ \frac{(\phi - \bar{z})\mu}{\sigma_z^2} \right\} \frac{2}{\sigma_z \sqrt{2\pi}} \frac{1}{2\sqrt{\frac{(\mu^2 + \sigma_z^2 2g_B)}{\sigma_z^2 2}}} \left(\sqrt{\pi} e^{-2\sqrt{\frac{(\phi - \bar{z})^2}{\sigma_z^2 2}} \sqrt{\frac{(\mu^2 + \sigma_z^2 2g_B)}{\sigma_z^2 2}}} \right) =$$

$$\frac{1}{\sqrt{(\mu^2 + \sigma_z^2 2g_B)}} e^{(\phi - \bar{z})\frac{\mu}{\sigma_z^2} - |\phi - \bar{z}| \frac{1}{\sigma_z^2} \sqrt{(\mu^2 + \sigma_z^2 2g_B)}} =$$

$$\frac{1}{\sqrt{(\mu^2 + \sigma_z^2 2g_B)}} e^{(\phi - \bar{z})\frac{\mu}{\sigma_z^2} - |\phi - \bar{z}| \frac{1}{\sigma_z^2} \sqrt{(\mu^2 + \sigma_z^2 2g_B)}} =$$

$$\begin{cases} \frac{e^{(\phi - \bar{z})\left(\frac{\mu}{\sigma_z^2} - \frac{1}{\sigma_z^2} \sqrt{(\mu^2 + \sigma_z^2 2g_B)} \right)}}{\sqrt{\mu^2 + \sigma_z^2 2g_B}} & \text{if } \phi \geq \bar{z} \\ \frac{e^{(\phi - \bar{z})\left(\frac{\mu}{\sigma_z^2} + \frac{1}{\sigma_z^2} \sqrt{(\mu^2 + \sigma_z^2 2g_B)} \right)}}{\sqrt{\mu^2 + \sigma_z^2 2g_B}} & \text{if } \phi < \bar{z} \end{cases}$$

Thus, the density is

$$g(\phi|\bar{z}) = \frac{\min\left\{ e^{\theta_1(\phi - \bar{z})}, e^{-\theta_2(\phi - \bar{z})} \right\}}{\sqrt{\mu^2 + \sigma_z^2 2g_B}},$$

$$\theta_{1,2} = \pm \frac{\mu}{\sigma_z^2} + \sqrt{\frac{\mu^2}{\sigma_z^4} + 2\frac{g_B}{\sigma_z^2}}.$$

Of course this implies that $g_B > 0$ in order for the distribution not to explode.

We can also express it as a probability density

$$p(\phi|\bar{z}) = \frac{\theta_1 \theta_2}{\theta_1 + \theta_2} \min\left\{ e^{\theta_1(\phi - \bar{z})},\ e^{-\theta_2(\phi - \bar{z})} \right\},$$

where

$$\frac{\theta_1\theta_2}{\theta_1+\theta_2} = \frac{2\frac{g_B}{\sigma_z^2}}{\sqrt{\frac{\mu^2}{\sigma_z^4}+2\frac{g_B}{\sigma_z^2}}} = \frac{2g_B}{\sqrt{\mu^2+2\sigma_z^2 g_B}} \ .$$

Notice that in order for the distribution function to be properly defined a restriction in θ_1, θ_2 has to be imposed and in particular that $\theta_1, \theta_2 > 0$. This implies that $g_B > 0$.

Finally, if we assume that the initial $\bar{z}$ is drawn randomly from a distribution $G(\bar{z})$ we can also derive

$$p(\phi) = \int p(\phi|\bar{z})\, G(\bar{z})\, d\bar{z} \ .$$

This is done in Reed (2002).

2 Existence and Uniqueness of the Equilibrium

The proof is a direct application of the algorithm, Lemmas and Theorems in Allen, Arkolakis, and Li (2014) (henceforth AAL). Notice that for convenience we omit t throughout the proof since the results are obtained for any given time period t. We first start by stating the equilibrium conditions.

Zero-profit cutoff: To derive the zero-profit cutoff notice that from the first order condition for n_{ij} we obtain

$$\frac{L_j y_j}{\sigma} \frac{p_{ij}^{1-\sigma}}{P_j^{1-\sigma}} - \frac{w_i L_j^{\alpha}}{\psi} (1 - n_{ij})^{-\beta} \geq 0.$$

For the cutoff firms, $n_{ij} = 0$ and the above equation holds with equality. Thus, given their constant markup price choice, $p_{ij}\left(z_{ij}^*\right) = \frac{\tilde{\sigma}\tau_{ij} w_i}{z_{ij}^*}$, we obtain the cutoff as

$$z_{ij}^* \quad = \quad c_{ij} \frac{w_i^{\frac{\sigma}{\sigma-1}}}{P_j y_j^{\frac{1}{\sigma-1}}}, \tag{1}$$

where $c_{ij} \equiv \tilde{\sigma}\tau_{ij} \left[\frac{\psi L_j^{1-\alpha}}{\sigma}\right]^{\frac{1}{1-\sigma}}$ is a constant.

Notice also that total bilateral sales for country i from country j can be written as

$$y_{ij} \equiv \frac{\sigma w_i L_j^{\alpha}}{\psi} \int_{z_{ij}^*}^{\infty} \left[\left(\frac{z}{z_{ij}^*}\right)^{\sigma-1} - \left(\frac{z}{z_{ij}^*}\right)^{(\sigma-1)\frac{\beta-1}{\beta}} \right] f_i(z) dz. \tag{2}$$

Budget Balance: Budget balance implies that the revenue earned by all firms from country j should equal to total expenditure $y_j L_j$, that is

$$y_j L_j = \sum_k y_{kj} \tag{3}$$

This equation essentially defines the price index P_j.

Labor market clearing: The labor market clearing condition implies that the total labor income should equal to the sum of labor income from production and marketing spending

$$w_i L_i = \frac{\sigma-1}{\sigma}\sum_j y_{ij} + \sum_j \frac{w_i L_j^{\alpha}}{\psi(1-\beta)} \int_{z_{ij}^*}^{\infty} \left[1 - \left(\frac{z}{z_{ij}^*}\right)^{(\sigma-1)\frac{\beta-1}{\beta}}\right] f_i(z)dz. \tag{4}$$

Current account balance: The current account balance condition implies that the total expenditure should equal to the total revenue, i.e.

$$y_i L_i = \sum_j y_{ij}. \tag{5}$$

Given the above definitions and to facilitate the proof, we introduce the following notation:

i. excess expenditure

$$E_j^e = \sum_k y_{kj} - y_j L_j, \tag{6}$$

ii. excess wage

$$E_i^w = \frac{\sigma-1}{\sigma}\sum_j y_{ij} + \sum_j \frac{w_i L_j^{\alpha}}{\psi(1-\beta)} \int_{z_{ij}^*}^{\infty} \left[1 - \left(\frac{z}{z_{ij}^*}\right)^{(\sigma-1)\frac{\beta-1}{\beta}}\right] f_i(z)dz - w_i L_i, \tag{7}$$

iii. excess income

$$E_i^I = \sum_j y_{ij} - y_i L_i. \tag{8}$$

The general equilibrium $\mathcal{F}$ is equivalent to E_j^e, E_i^w and E_i^I equal to zero for all i. Thus, $\mathcal{F}$ is the equilibrium as specified for Proposition 2 in the main paper. Following AAL we define a set of partial equilibria, $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$ to facilitate the proof: $\mathcal{F}_1$ is defined as the equilibrium where w and y are exogenously given and price P is endogenously solved by equation (3) i.e. $E_j^e = 0$; $\mathcal{F}_2$ is defined as the equilibrium where y are exogenously given and wage w are endogenously solved by equations (4) i.e. $E_i^w = 0$ in which price P is solved in $\mathcal{F}_1$; $\mathcal{F}_3$ is defined as the equilibrium where y is endogenously solved by equation (5) i.e. $E_i^I = 0$ in which P and w is solved in $\mathcal{F}_2$. $\mathcal{F}_3$ is equivalent to $\mathcal{F}$.

Technical preparation

In what follows we are going to differentiate E_j^e, E_i^w and E_i^I, which will be used in the proof of existence and uniqueness. Notice that differentiating bilateral sales we obtain

$$\begin{aligned} dy_{ij} &= -(\sigma-1)\frac{dz_{ij}^*}{z_{ij}^*}\frac{\sigma w_i L_j^\alpha}{\psi}\int_{z_{ij}^*}\left[\left(\frac{z}{z_{ij}^*}\right)^{\sigma-1}-\frac{\beta-1}{\beta}\left(\frac{z}{z_{ij}^*}\right)^{(\sigma-1)\frac{\beta-1}{\beta}}\right]f_i(z)\,dz+y_{ij}\frac{dw_i}{w_i} \\ &= -\left[(\sigma-1)\,y_{ij}+\sigma\triangle N_{ij}\right]\frac{dz_{ij}^*}{z_{ij}^*}+y_{ij}\frac{dw_i}{w_i} \end{aligned}$$

where we denote $\triangle N_{ij}=\frac{(\sigma-1)}{\beta}\frac{w_i L_j^\alpha}{\psi}\int_{z_{ij}^*}\left(\frac{z}{z_{ij}^*}\right)^{(\sigma-1)\frac{\beta-1}{\beta}}m_i(z)\,dz$. Furthermore if we denote $M_{ij}=(\sigma-1)\,y_{ij}+\sigma\triangle N_{ij}$, we can write

$$dE_j^e=\sum_k y_{kj}\frac{dw_k}{w_k}-\sum_k M_{kj}\frac{dz_{kj}^*}{z_{kj}^*}-y_jL_j\frac{dy_j}{y_j} \tag{9}$$

The differentiation of the excess wage gives

$$\begin{aligned} dE_i^w &= d\left\{\frac{\sigma-1}{\sigma}\sum_j y_{ij}+\sum_j\frac{w_iL_j^\alpha}{\psi(1-\beta)}\int_{z_{ij}^*}^\infty\left[1-\left(\frac{z}{z_{ij}^*}\right)^{(\sigma-1)\frac{\beta-1}{\beta}}\right]f_i(z)dz-w_iL_i\right\} \\ &= E_i^w\frac{dw_i}{w_i}-\frac{\sigma-1}{\sigma}\sum_j M_{ij}\frac{dz_{ij}^*}{z_{ij}^*}+\sum_j\frac{w_iL_j^\alpha}{\psi(1-\beta)}d\int_{z_{ij}^*}^\infty\left[1-\left(\frac{z}{z_{ij}^*}\right)^{(\sigma-1)\frac{\beta-1}{\beta}}\right]f_i(z)dz \end{aligned}$$

Notice also that

$$\begin{aligned} \frac{w_iL_j^\alpha}{\psi(1-\beta)}d\int_{z_{ij}^*}^\infty\left[1-\left(\frac{z}{z_{ij}^*}\right)^{(\sigma-1)\frac{\beta-1}{\beta}}\right]f_i(z)dz &= -\frac{dz_{ij}^*}{z_{ij}^*}\frac{(\sigma-1)}{\beta}\frac{w_iL_j^\alpha}{\psi}\int_{z_{ij}^*}^\infty\left(\frac{z}{z_{ij}^*}\right)^{(\sigma-1)\frac{\beta-1}{\beta}}f_i(z)dz \\ &= -\triangle N_{ij}\frac{dz_{ijt}^*}{z_{ijt}^*}. \end{aligned}$$

Thus,

$$dE_i^w = E_i^w\frac{dw_i}{w_i}-\sum_j M_{ij}^1\frac{dz_{ij}^*}{z_{ij}^*} \tag{10}$$

where $M^1_{ij} = \frac{\sigma-1}{\sigma} M_{ij} + \triangle N_{ij} = \frac{(\sigma-1)^2}{\sigma} y_{ij} + \sigma \triangle N_{ij}$. Similarly,

$$dE^I_i = -\sum_j M_{ij} \frac{dz^*_{ij}}{z^*_{ij}} + \sum_j y_{ij} \frac{dw_i}{w_i} - y_i L_i \frac{dy_i}{y_i}. \tag{11}$$

In all the above, $\frac{dz^*_{ij}}{z^*_{ij}}$ is given by differentiating equation 1,

$$\frac{dz^*_{ij}}{z^*_{ij}} = -\frac{dP_j}{P_j} + \frac{\sigma}{\sigma-1} \frac{dw_i}{w_i} + \frac{1}{1-\sigma} \frac{dy_j}{y_j}. \tag{12}$$

Insert the above equation into 9, and in $\mathcal{F}_1$ we have $dE^e_j = 0$, so that

$$\begin{aligned} \frac{dP_j}{P_j} &= \frac{\sum_k \tilde{\sigma} \left[\frac{(\sigma-1)^2}{\sigma} y_{kj} + \sigma \triangle N_{kj}\right] \frac{dw_k}{w_k} - \tilde{\sigma} \sum_k \triangle N_{kj} \frac{dy_j}{y_j}}{\sum_k \left[(\sigma-1) y_{kj} + \sigma \triangle N_{kj}\right]} \\ &= \frac{\sum_k \tilde{\sigma} M^1_{kj} \frac{dw_k}{w_k} - \tilde{\sigma} \sum_k \triangle N_{kj} \frac{dy_j}{y_j}}{\sum_k M_{kj}}. \end{aligned} \tag{13}$$

Existence and uniqueness

Now we proceed to prove proposition 2 in the main paper. We will state the requirements of the proposition as the following condition

Condition 1 *Let $G(z^*) = \int_{z^*}^{\infty} \left(e^{\bar{c}_1 \ln(z/z^*)} - e^{\bar{c}_2 \ln(z/z^*)}\right) f(z)\, dz$. Assume $G(z^*)$ is finite for any $z^* > 0$.*

In what follows we use two lemmas and a theorem regarding the existence and uniqueness of $\mathcal{F}_1$, $\mathcal{F}_2$ and $\mathcal{F}_3$ so that we prove the uniqueness and existence of a general equilibrium, given condition C.1. The strategy of the proof follows the main steps of AAL and makes use of two relevant lemmas and one theorem.

In fact, the main differences lie in the different definitions of M_{ij}, M^1_{ij} and $\triangle N_{ij}$ used in AAL (their case corresponds to the limiting case $\beta \to 0$). However, since the properties of these objects are exactly the same, (in particular $0 < \triangle N_{ij} < M^1_{ij} < M_{ij}$) they are used in the proof fashion as in AAL to prove the uniqueness. The existence part is very similar as well but requires slight modifications that we illustrate below in further detail. We proceed with three steps proving existence and uniqueness of equilibrium for $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$ in each respective step.

Step 1:

We first prove the following lemma

Lemma 2 *Assume C.1 holds. Then there exists a unique price given wage and income solving $\mathcal{F}_1$.*

Proof. For existence, notice that according to equation (1), for any given P_i, w_i and y_i there always exists a cutoff z_{ij}^*. And also given any w_j and y_j, when $P_j \to 0$, $z_{kj}^* \to \infty$, so

$$E_j^e = \sum_k \frac{\sigma w_k L_j^\alpha}{\psi} G_k(z_{kj}^*) - y_j L_j, \to -y_j L_j;$$

and when $P_j \to \infty$, $z_{kj}^* \to 0$, the above condition implies that $E_j^e \to \infty$. Besides, obviously E_j^e is continuous with respect t P_j, according to the mean value theorem, there must exist a price P_j satisfying (3) i.e. $E_j^e = 0$ holds .

The uniqueness proof can be found in AAL given that $M_{ij}, M_{ij}^1, \triangle N_{ij}$ satisfy $0 < \triangle N_{ij} < M_{ij}^1 < M_{ij}$. This completes the first step. ■

Step 2:

We now prove the second lemma that establishes the existence of a unique solution in $\mathcal{F}_2$.

Lemma 3 *Assume C.1 holds. There exists a unique wage given income solving $\mathcal{F}_2$.*

Proof. We start again with the existence part. To simplify notation, we first define a few functions:

$$Z_{ij}\left(w_i, P_j, y\right) = c_{ij} \frac{\left(w_i\right)^{1+\frac{1}{\sigma-1}}}{P_j y_j^{\frac{1}{\sigma-1}}},$$

which is increasing w.r.t w_i;

$$H_i\left(z^*\right) = \left(\sigma - 1\right) G_i\left(z^*\right) + \frac{1}{1-\beta} \int_{z^*}^{\infty} \left[1 - \left(\frac{z}{z^*}\right)^{(\sigma-1)\frac{\beta-1}{\beta}}\right] f_i(z) dz,$$

$H_i(z)$ which is decreasing with respect to z^* and moreover if $z^* \to 0$, $H_i(z^*) > (\sigma-1)\, G_i(z^*) \to \infty$; if $z^* \to \infty$, $H_i(z^*) < \left(\sigma - 1 + \frac{1}{1-\beta}\right) G_i(z^*) \to 0$; $P_j(w,y)$ is the implicit function of equation (3) according to lemma 1 and from equation (13) we know that $P(w,y)$ increases with w and decreases with y. The rest of the existence proof follows AAL and we illustrate the details of its implementation.

Define a continuous operator $W : w^{old} \to w^{new}$ where w_i^{new} is determined in $L_i = \sum_j \frac{L_j^\alpha}{\psi} H_i\left(Z_{ij}\left(w_i^{new}, P_j\left(w^{old}, y\right), y\right)\right)$. The existence of a fixed point $w \in R_{++}^N$ of operator W is equivalent with the existence of w in $\mathcal{F}_2$. Notice that $\sum_j \frac{L_j^\alpha}{\psi} H_i\left(Z_{ij}\left(w_i^{new}, P_j, y\right)\right)$ is decreasing w.r.t w_i^{new}. If $w_i^{new} \to 0$,$Z_{ij}\left(w_i^{new}, P_j, y\right) \to 0$ thus $\sum_j \frac{L_j^\alpha}{\psi} H_i\left(Z_{ij}\left(w_i^{new}, P_j, y\right)\right) \to \infty$; $w_i^{new} \to \infty$,$Z_{ij}\left(w_i^{new}, P_j, y\right) \to \infty$, $\sum_j \frac{L_j^\alpha}{\psi} H_i\left(Z_{ij}\left(w_i^{new}, P_j, y\right)\right) \to 0$. Then for any given w^{old} and y, there exists a unique w_i^{new}, thus W is a well defined continuous operator (function). Furthermore, W is increasing w.r.t w^{old}.

From lemma 1 of AAL, to prove the existence of fixed point, it is sufficient is to verify negative feedback condition i.e. there exist constants $m > 0$, $M > 0$ and w^0 such that $mw^0 \le W\left(mw^0\right) \le W\left(Mw^0\right) \le Mw^0$, as monotonicity conditions of W holds from above.

We will first see how the price index change if the wage w^0 becomes $tw^0 (t \ne 1)$. If t increase, to keep equation (3) $y_j L_j = \sum_k \frac{\sigma t w_k^0 L_j^\alpha}{\psi} \int_{z_{kj}^*}^{\infty} \left(\left(\frac{z}{z_{kj}^*}\right)^{\sigma-1} - \left(\frac{z}{z_{kj}^*}\right)^{(\sigma-1)\frac{\beta-1}{\beta}}\right) f_k(z)dz$ holds, $z_{kj}^* = c_{kj} \frac{\left(tw_k^0\right)^{1+\frac{1}{\sigma-1}}}{P_j(tw^0,y) y_j^{\frac{1}{\sigma-1}}}$ has to increase. Moreover, if $t \to 0$, $z_{kj}^* \to 0$; if $t \to \infty$, $z_{kj}^* \to \infty$. Define $\nabla_p(t) = \frac{P\left(tw^0,y\right)}{t^{\tilde{\sigma}} P\left(w^0,y\right)}$, thus it is an decreasing function, $\nabla_p(1) = 1$, and if $t \to 0$, $\nabla_p(t) \to \infty$; if $t \to \infty$, $\nabla_p(t) \to 0$.

Denote $\nabla_w(t) = \frac{W\left(tw^0\right)}{tw^0}$. Then $Z_{ij}\left(W\left(tw^0\right)_i, P_j, y\right) = \frac{(\nabla_w(t))_i}{(\nabla_p(t))_j} Z_{ij}\left(w^0, P\left(w_0, y\right), y\right)$. Obviously, $\frac{(\nabla_w(t))_i}{\max_j (\nabla_p(t))_j} \le \frac{(\nabla_w(t))_i}{(\nabla_p(t))_j} \le \frac{(\nabla_w(t))_i}{\min_j (\nabla_p(t))_j}$. Thus

$$\sum_j \frac{L_j^\alpha}{\psi} H_i\left(\frac{(\nabla_w(t))_i}{\min_j (\nabla_p(t))_j} z_{ij}^0\right) \le L_i = \sum_j \frac{L_j^\alpha}{\psi} H_i\left(\frac{(\nabla_w(t))_i}{(\nabla_p(t))_j} z_{ij}^0\right) \le \sum_j \frac{L_j^\alpha}{\psi} H_i\left(\frac{(\nabla_w(t))_i}{\max_j (\nabla_p(t))_j} z_{ij}^0\right).$$

Also notice that there exists a constant s_i satisfying $L_i = \sum_j \frac{L_j^\alpha}{\psi} H_i\left(s_i z_{ij}^0\right)$. Thus

$$\sum_j \frac{L_j^\alpha}{\psi} H_i\left(\frac{(\nabla_w(t))_i}{\min\nabla_p(t)} z_{ij}^0\right) \leq \sum_j \frac{L_j^\alpha}{\psi} H_i\left(s_i z_{ij}^0\right) \leq \sum_j \frac{L_j^\alpha}{\psi} H_i\left(\frac{(\nabla_w(t))_i}{\max\nabla_p(t)} z_{ij}^0\right)$$

which means that $\frac{(\nabla_w(t))_i}{\max\limits_j(\nabla_p(t))_j} \leq s_i \leq \frac{(\nabla_w(t))_i}{\min\limits_j(\nabla_p(t))_j}$ i.e.

$$s_i \min_j (\nabla_p(t))_j \leq (\nabla_w(t))_i \leq s_i \max_j (\nabla_p(t))_j$$

As the range of $\nabla_p(t)$ is $(0, \infty)$, so is $(\nabla_w(t))_i$, specifically, if $t \to 0$, $(\nabla_w(t))_i \to \infty$; if $t \to \infty$, $(\nabla_w(t))_i \to 0$. Then there must exist $t = M$ such that for all i $(\nabla_w(t))_i < 1$; and also there must exist $t = m$ such that for all i $(\nabla_w(t))_i > 1$. So $mw^0 \leq W\left(mw^0\right) \leq W\left(Mw^0\right) \leq Mw^0$. Existence holds.

The uniqueness proof can be found in AAL given that $M_{ij}, M_{ij}^1, \triangle N_{ij}$ satisfy $0 < \triangle N_{ij} < M_{ij}^1 < M_{ij}$. This completes the second step. ■

Step 3:

The following Theorem is proved in AAL and its statement is equivalent to Proposition 2 in the paper:

Theorem 4 *Assume C.1 holds. Then $\mathcal{F}$ has a unique solution.*

and $\mathcal{F}_1$,$\mathcal{F}_2$ have a unique solution.

This last theorem completes the proof of Proposition 2.

3 Finite Integrals with O-U

Below we prove that the integrals for bilateral sales that arise from the O-U process are finite with $\mu, \rho < 0$. We also assume that the rate of entry of new ideas is positive, $g_B > 0$. The density function of z in this case is given by

$$f(z,a) = \frac{1}{z}\frac{1}{\sqrt{2\pi\frac{\sigma^2}{2\rho}\left(1-e^{-2\rho a}\right)}}\exp\left(-\frac{\left[\ln z-\mu a\right]^2}{2\frac{\sigma^2}{2\rho}\left(1-e^{-2a}\right)}\right)$$

The relevant integrals in the paper have the form

$$\int_{z^*}^{\infty} z^{\sigma-1}\int_0^{\infty} g_B e^{-g_B a} f(x,a)\, dadx =$$

$$\int_{z^*}^{\infty} z^{\sigma-1}\int_0^{\infty} g_B e^{-g_B a}\frac{1}{z}\frac{1}{\sqrt{2\pi\frac{\sigma^2}{2\rho}\left(1-e^{-2\rho a}\right)}}\exp\left(-\frac{\left[\ln z-\mu a\right]^2}{2\frac{\sigma^2}{2\rho}\left(1-e^{-2a}\right)}\right) dadz =$$

$$\int_0^{\infty} g_B e^{-g_B a}\int_{z^*}^{\infty} z^{\sigma-2}\frac{1}{\sqrt{2\pi\frac{\sigma^2}{2\rho}\left(1-e^{-2\rho a}\right)}}\exp\left(-\frac{\left[\ln z-\mu a\right]^2}{2\frac{\sigma^2}{2\rho}\left(1-e^{-2a}\right)}\right) dadz.$$

Denote by ,$D = \frac{\sigma^2}{2\rho}\left(1-e^{-2a}\right)$,$y = \ln z$,$e^y = z$, we have

$$\int_{z^*}^{\infty} z^{\sigma-2}\frac{1}{\sqrt{2\pi\frac{\sigma^2}{2\rho}\left(1-e^{-2\rho a}\right)}}\exp\left(-\frac{\left[\ln z-\mu a\right]^2}{2\frac{\sigma^2}{2\rho}\left(1-e^{-2a}\right)}\right) dz =$$

$$\int_{\ln z^*}^{\infty}\frac{1}{\sqrt{2\pi D}}\exp\left(-\frac{\left(y-\mu a\right)^2}{2D}+\left(\sigma-1\right)y\right) dy =$$

$$\int_{\ln z^*}^{\infty}\frac{1}{\sqrt{2\pi D}}\exp\left(-\frac{\left(y-\mu a-\left(\sigma-1\right)D\right)^2}{2D}+\mu a\left(\sigma-1\right)+\frac{\left(\sigma-1\right)^2 D}{2}\right) dy <$$

$$< \exp\left(\mu a\left(\sigma-1\right)+\frac{\left(\sigma-1\right)^2 D}{2}\right).$$

All in all,

$$\begin{aligned}\int_{z^*}^{\infty} z^{\sigma-1}\int_0^{\infty} g_B e^{-g_B a} g(x,a)\, dadz &< \int_0^{\infty} g_B e^{-g_B a}\exp\left(\mu a\left(\sigma-1\right)+\frac{\left(\sigma-1\right)^2 D}{2}\right) da \\ &= \int_0^{\infty} g_B \exp\left(\left[\mu\left(\sigma-1\right)-g_B\right]a+\frac{\left(\sigma-1\right)^2 D}{2}\right) da.\end{aligned}$$

Notice that $\frac{(\sigma-1)^2 D}{2} \leq \frac{(\sigma-1)^2\sigma^2}{4\rho}$ and, thus, as long as $\mu(\sigma-1) - g_B < 0$ the integral is bounded. But we have assumed $\mu < 0$ and $g_B > 0$ and thus$\mu(\sigma-1) - g_B < 0$. This last step completes the proof.

4 Proof for expected time to reach a certain size

As in the main paper we denote by s_b the log ratio of the productivity of the firm to the cutoff of entry. An exogenous death shock arrives at a rate δ. We let $s_0 = 0$ and characterize the expected time required for the firm to reach a certain size $s_b > 0$, conditional on not being hit by a death shock. Essentially, we want to compute $P\left(T\left(s = s_b\right) > t | \text{Not Death}\right)$ where $T\left(s = s_b\right)$ is the first time that $s = s_b$. To do that we simply have to compute

$$\int_0^t P\left(T\left(s_b\right) = a | \text{Not Death by } a\right) da =$$
$$\int_0^t P\left(T\left(s_b\right) = a\right) \Pr\left(\text{Not Death by } a\right) da =$$
$$\int_0^t P\left(T\left(s_b\right) = a\right) e^{-\delta a} da \ ,$$

since the probability that a firm is not hit by an exogenous death shock by time a is $e^{-\delta a}$. It is well known that (see for example Harrison (1985), p. 14):

$$P\left(T\left(s_b\right) = a\right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\frac{s_b^2 + (\mu a)^2 - 2 s_b \mu a}{\sigma_z^2 a}} \frac{s_b}{a^{\frac{3}{2}} \sigma_z} \ .$$

Thus, we have to compute

$$\int_0^t P\left(T\left(s_b\right) = a | \text{Not Death by } a\right) da = \int_0^t \frac{e^{-\frac{1}{2}\frac{s_b^2 + (\mu a)^2 - 2 s_b \mu a}{\sigma_z^2 a}}}{\sqrt{2\pi}} \frac{s_b}{a^{\frac{3}{2}} \sigma_z} e^{-\delta a} da =$$
$$\int_0^t \frac{e^{-\frac{1}{2}\frac{s_b^2 + (\mu a)^2 - 2 s_b \mu a + 2\delta \sigma_z^2 a^2}{\sigma_z^2 a}}}{\sqrt{2\pi}} \frac{s_b}{a^{\frac{3}{2}} \sigma_z} da =$$
$$\frac{s_b}{\sigma_z \sqrt{2\pi}} e^{\frac{s_b \mu}{\sigma_z^2}} \int_0^t e^{-\frac{1}{2}\frac{s_b^2}{\sigma_z^2 a} - \frac{1}{2}\frac{(\mu)^2 + 2\delta \sigma_z^2}{\sigma_z^2} a} \frac{1}{a^{\frac{3}{2}}} da \ .$$

Using change of variables we have that

$$\begin{aligned} \tilde{a} &= \frac{1}{a} \implies d\tilde{a} = -\frac{1}{a^2} da \\ &\implies \tilde{a}^{-\frac{1}{2}} d\tilde{a} = -\frac{1}{a^2} a^{1/2} da \implies \\ \tilde{a}^{-\frac{1}{2}} d\tilde{a} &= -\frac{1}{a^{3/2}} da, \end{aligned}$$

which implies that

$$\frac{s_b}{\sigma_z\sqrt{2\pi}}e^{\frac{s_b\mu}{\sigma_z^2}}\int_0^t e^{-\frac{1}{2}\frac{s_b^2}{\sigma_z^2 a}-\frac{1}{2}\frac{(\mu)^2+2\delta\sigma_z^2}{\sigma_z^2}a}\frac{1}{a^{\frac{3}{2}}}da =$$
$$-\frac{s_b}{\sigma_z\sqrt{2\pi}}e^{\frac{s_b\mu}{\sigma_z^2}}\int_0^t e^{-\frac{1}{2}\frac{s_b^2}{\sigma_z^2}\tilde{a}-\frac{1}{2}\frac{(\mu)^2+2\delta\sigma_z^2}{\sigma_z^2}\frac{1}{\tilde{a}}}\tilde{a}^{-\frac{1}{2}}d\tilde{a} =$$
$$-\frac{s_b}{\sigma_z\sqrt{2\pi}}e^{\frac{s_b\mu}{\sigma_z^2}}\frac{1}{2\sqrt{\frac{1}{2}\frac{s_b^2}{\sigma_z^2}}}e^{-2\sqrt{\frac{1}{2}\frac{s_b^2}{\sigma_z^2}}\sqrt{\frac{1}{2}\frac{(\mu)^2+2\delta\sigma_z^2}{\sigma_z^2}}}\sqrt{\pi}\times$$
$$\left|-erf\left(\frac{\sqrt{\frac{1}{2}\frac{(\mu)^2+2\delta\sigma_z^2}{\sigma_z^2}}-\sqrt{\frac{1}{2}\frac{s_b^2}{\sigma_z^2}x}}{\sqrt{x}}\right)+e^{4\sqrt{\frac{1}{2}\frac{s_b^2}{\sigma_z^2}}\sqrt{\frac{1}{2}\frac{(\mu)^2+2\delta\sigma_z^2}{\sigma_z^2}}}\left[\operatorname{erf}\left(\frac{\sqrt{\frac{1}{2}\frac{s_b^2}{\sigma_z^2}x}+\sqrt{\frac{1}{2}\frac{(\mu)^2+2\delta\sigma_z^2}{\sigma_z^2}}}{\sqrt{x}}\right)-1\right]+1\right|_{+\infty}^{1/t} =$$
$$-e^{\frac{s_b\mu}{\sigma_z^2}-\sqrt{\frac{s_b^2}{\sigma_z^2}}\sqrt{\frac{(\mu)^2+2\delta\sigma_z^2}{\sigma_z^2}}}\frac{1}{2}\times$$
$$\left[\begin{array}{c} -erf\left(\frac{\sqrt{\frac{1}{2}\frac{(\mu)^2+\delta\sigma_z^2}{\sigma_z^2}}-\sqrt{\frac{1}{2}\frac{s_b^2}{\sigma_z^2}\frac{1}{t}}}{\sqrt{\frac{1}{t}}}\right)+e^{4\sqrt{\frac{1}{2}\frac{s_b^2}{\sigma_z^2}}\sqrt{\frac{1}{2}\frac{(\mu)^2+\delta\sigma_z^2}{\sigma_z^2}}}\left[\operatorname{erf}\left(\frac{\sqrt{\frac{1}{2}\frac{s_b^2}{\sigma_z^2}\frac{1}{t}}+\sqrt{\frac{1}{2}\frac{(\mu)^2+\delta\sigma_z^2}{\sigma_z^2}}}{\sqrt{\frac{1}{t}}}\right)-1\right]- \\ -\left[-(-1)\right]+0 \end{array}\right].$$

Thus, the probability that we are looking for is given by:

$$P\left(T\left(s=s_b\right)>t|\text{Not Death}\right)=e^{\frac{s_b\mu}{\sigma_z^2}-\sqrt{\frac{s_b^2}{\sigma_z^2}}\sqrt{\frac{(\mu)^2+2\delta\sigma_z^2}{\sigma_z^2}}}\times$$
$$\left[1-\Phi\left(\frac{-\sqrt{(\mu)^2+\delta\sigma_z^2}t+s_b}{\sigma_z\sqrt{t}}\right)+e^{2\sqrt{\frac{s_b^2}{\sigma_z^2}}\sqrt{\frac{(\mu)^2+2\delta\sigma_z^2}{\sigma_z^2}}}\Phi\left(\frac{-s_b-\sqrt{(\mu)^2+2\delta\sigma_z^2}t}{\sigma_z\sqrt{t}}\right)\right].$$

Notice that if we simply study the case $\mu<0$, $\delta=0$ the probability is given by

$$e^{\frac{s_b\mu}{\sigma_z^2}+\frac{s_b\mu}{\sigma_z^2}}\times$$
$$\left[1-\Phi\left(\frac{\mu t+s_b}{\sigma_z\sqrt{t}}\right)+e^{-2\frac{s_b\mu}{\sigma_z^2}}\Phi\left(\frac{-s_b+\mu t}{\sigma_z\sqrt{t}}\right)\right]=$$
$$e^{2\frac{s_b\mu}{\sigma_z^2}}\left[1-\Phi\left(\frac{\mu t+s_b}{\sigma_z\sqrt{t}}\right)\right]+\Phi\left(\frac{-s_b+\mu t}{\sigma_z\sqrt{t}}\right)=$$
$$1-\Phi\left(\frac{s_b-\mu t}{\sigma_z\sqrt{t}}\right)+e^{2\frac{s_b\mu}{\sigma_z^2}}\left[\Phi\left(\frac{-\mu t-s_b}{\sigma_z\sqrt{t}}\right)\right],$$

which is the well known hitting time for a Brownian motion without the exogenous death (see for example Harrison (1985), p. 14).

It is worth pointing out that the derivation above refers to the expected hitting time of a Brownian motion with a drift which is a continuous time process. When we report the sales of firms in the model we look at their sales at discrete points of time, $t = 0, 1, 2, ...$ a common approach in this literature (see for example Klette and Kortum (2004) and Luttmer (2007)). If we were reporting the probability that a firm surpasses a sales level at discrete points of time $t = 1, 2, ...$ the numbers would be very similar, as numerical simulations indicate. In addition, this derivation can be done analytically using the cdf of the probability distribution of productivities.

5 Additional result for the cohort survival rate

In this section we prove various results for the cohort survival rate

$$S_{ij}\left(a\right)=e^{-\delta a}\left[\Phi\left(\frac{\mu}{\sigma_z}\sqrt{a}\right)+e^{a\left(\frac{\sigma_z^2}{2}\theta_2^2+\mu\theta_2\right)}\Phi\left(-\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\sqrt{a}\right)\right]. \tag{14}$$

a) First, we prove that the survival function is decreasing in a, for $\mu<0$. To show that it suffices to show that $DS_{ij}\left(a\right)<0$. Notice that the derivative with respect to a of the part of the expression $S_{ij}\left(a\right)$ inside the brackets is given by

$$\begin{aligned}&\varphi\left(\frac{\mu\sqrt{a}}{\sigma_z}\right)\frac{\mu}{2\sigma_z\sqrt{a}}+\\ +\;\;& e^{a\left(\frac{\sigma_z^2}{2}\theta_2^2+\mu\theta_2\right)}\left[\left(\frac{\sigma_z^2}{2}\theta_2^2+\mu\theta_2\right)\Phi\left(-\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\sqrt{a}\right)-\frac{\mu+\theta_2\sigma_z^2}{2\sigma_z\sqrt{a}}\varphi\left(-\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\sqrt{a}\right)\right]\end{aligned}$$

The first term is always negative. To show that $S_{ij}\left(a\right)$ decreases in a when $\mu<0$ it suffices to show that the term in the bracket is negative which implies that

$$\frac{\left(\frac{\sigma_z^2}{2}\theta_2^2+\mu\theta_2\right)}{\frac{1}{2}\left(\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\right)^2}\left[1-\Phi\left(\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\sqrt{a}\right)\right]<\frac{\varphi\left(\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\sqrt{a}\right)}{\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\sqrt{a}} \tag{15}$$

Given that $\sigma_z^2\theta_2+\mu>0$, as implied by assumption 2, and $\theta_2>1$, we can make use of property P5 (See at the end of this Appendix for definitions). This property implies that expression (15) is negative if

$$\frac{\frac{\sigma_z^2}{2}\theta_2^2+\mu\theta_2}{\frac{1}{2}\left(\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\right)^2}<1\implies 0<\mu^2\ ,$$

which holds for $\mu<0$.

b) Second, one can show that higher drift implies higher survival probability for a given age, a. In order to characterize the sign of the derivative the expression $S_{ij}\left(a\right)$ with respect to μ, we need to characterize the sign of the following expression:

$$\varphi\left(\frac{\mu\sqrt{a}}{\sigma_z}\right)\frac{\sqrt{a}}{\sigma_z}+e^{\left(\frac{\sigma_z^2}{2}\theta_2^2+\mu\theta_2\right)a}\left[a\theta_2\Phi\left(-\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\sqrt{a}\right)-\frac{\sqrt{a}}{\sigma_z}\varphi\left(-\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\sqrt{a}\right)\right].$$

Notice that this equation can be decomposed in three terms. The first and the third term of these terms cancel out since completing the square it can be shown that:

$$e^{-\frac{1}{2}\left(\frac{\mu\sqrt{a}}{\sigma_z}\right)^2} = e^{\left(\frac{\sigma_z^2}{2}\theta_2^2+\mu\theta_2\right)a} e^{-\frac{1}{2}\left(\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\sqrt{a}\right)^2} .$$

The remaining second term is always positive which implies the result.

c) Finally, the derivative of the survival probability with respect to σ_z, this derivative is negative meaning that more variability implies smaller fraction of firms surviving for each given time. The proof of this result is provided upon request.

6 Hazard Rates

6.1 Firm Hazard Rates

In the main paper we prove the expression for the survival probability of a firm

$$S_{ij}\left(a|s_{ij0}=s_0\right)=e^{-\delta a}\Phi\left(\frac{s_0+\mu a}{\sigma_z\sqrt{a}}\right)\ . \tag{16}$$

Now we characterize the following proposition about the hazard rate of survival – Properties (P) and Assumptions (A) can be found at the end of this appendix–.

Lemma 5 *Given A1-A2, the (instantaneous) hazard rate of survival for a firm of a given size s_{ij0} in market j after time a has elapsed is given by*

$$-\frac{DS_{ij}\left(a|s_{ij0}\right)}{S_{ij}\left(a|s_{ij0}\right)}=\delta+m\left(-\frac{s_{ij0}+\mu a}{\sigma_z\sqrt{a}}\right)\frac{s_{ij0}-\mu a}{2a\sigma_z\sqrt{a}}\ , \tag{17}$$

where $m(x)=\varphi(x)/\Phi(-x)$ is the inverse Mills ratio, with $\varphi(x)$, $\Phi(x)$ are the pdf and the cdf of the standard normal distribution. If $\mu<0$, the long run hazard rate converges to

$$-\lim_{a\to\infty}\frac{DS_{ij}\left(a|s_{ij0}\right)}{S_{ij}\left(a|s_{ij0}\right)}=\delta+\left(\mu/\sigma_z\right)^2/2.$$

Proof. From expression (16) we can compute the instantaneous (conditional) hazard rate which is defined as the rate of change of the survivor function, $-DS_{ij}\left(a|s_{ij0}\right)/S_{ij}\left(a|s_{ij0}\right)$. Simple substitution for the definition of $S_{ij}\left(a|s_{ij0}\right)$ gives equation (17). Notice that the limit of the expression (17) $s_{ij0}\to\infty$ (for a given a), is given by

$$\lim_{s_{ij0}\to\infty}\left[-\frac{DS_{ij}\left(a|s_{ij0}\right)}{S_{ij}\left(a|s_{ij0}\right)}\right]=\delta+\frac{1}{2\sigma_z\sqrt{a}a}\frac{\lim_{s_{ij0}\to\infty}\varphi\left(\frac{s_{ij0}+\mu a}{\sigma_z\sqrt{a}}\right)\left(s_{ij0}-\mu a\right)}{\lim_{s_{ij0}\to\infty}\Phi\left(\frac{s_{ij0}+\mu a}{\sigma_z\sqrt{a}}\right)}=\delta+0\ ,$$

which implies that the instantaneous hazard rate for large firms is only the exogenous death rate of ideas.[1] I also consider two limits of the hazard rate $a\to\infty$ and $a\to 0$, for $\mu<0$.

[1] In this result and other results of the appendix I use the fact that exponential growth is faster than polynomial growth without further discussion.

Since the first term of equation (17) is always δ we will derive what happens to the second term in these two cases. First, consider the limit of the second term of equation (17) for $a \to \infty$. This limit is 0/0 and thus applying l' Hospital rule:

$$\lim_{a\to+\infty} \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{s_{ij0}+\mu a}{\sigma_z\sqrt{a}}\right)^2} \frac{s_{ij0}+\mu a}{\sigma_z\sqrt{a}} \left(\frac{s_{ij0}-\mu a}{a2\sigma_z\sqrt{a}}\right)^2 + \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{s_{ij0}+\mu a}{\sigma_z\sqrt{a}}\right)^2} \frac{a\mu-3s_{ij0}}{4a^{\frac{5}{2}}\sigma_z}}{\frac{-s_{ij0}+\mu a}{a2\sigma_z\sqrt{a}} \varphi\left(\frac{s_{ij0}+\mu a}{\sigma_z\sqrt{a}}\right)} =$$

$$\lim_{a\to+\infty} \frac{s_{ij0}+\mu a}{\sigma_z\sqrt{a}} \frac{-s_{ij0}+\mu a}{a2\sigma_z\sqrt{a}} + \lim_{a\to+\infty} \frac{a\mu - 3s_{ij0}}{(-s_{ij0}+\mu a)\, 4a\sigma_z} = \frac{1}{2}\left(\frac{\mu}{\sigma_z}\right)^2 + 0$$

This derivation gives the result of lemma 5. ■

Also it is easy to derive that the hazard rate goes to zero when $a \to 0$, i.e. derive the hazard rate for very small ages: Notice that $\frac{s_{ij0}+\mu a}{\sigma_z\sqrt{a}}$ always declines with age and thus also $\Phi\left(\frac{s_{ij0}+\mu a}{\sigma_z\sqrt{a}}\right)$. Thus,

$$-\lim_{a\to 0} \frac{\varphi\left(\frac{s_{ij0}+\mu a}{\sigma_z\sqrt{a}}\right) \frac{-s_{ij0}+\mu a}{a2\sigma_z\sqrt{a}}}{\Phi\left(\frac{s_{ij0}+\mu a}{\sigma_z\sqrt{a}}\right)} = \frac{0}{1} = 0$$

Finally, remains to show that in the case where $\mu < 0$, $s_{ij0} > 0$, the firm hazard in expression (17) ends up not being monotonic in age a. Simply notice that P4 implies that the term $m(\cdot)$ is decreasing in its argument, $\frac{s_{ij0}+\mu a}{\sigma_z\sqrt{a}}$, where the latter is initially decreasing and eventually increasing with age. The term $\frac{s_{ij0}-\mu a}{a2\sigma_z\sqrt{a}}$ is always decreasing in age.

6.2 Cohort Hazard Rates

Lemma 6 *Cohorts hazard rates for $a \to \infty$ are*

$$-\lim_{a\to\infty} \frac{DS_{ij}(a)}{S_{ij}(a)} = \delta + \frac{1}{2}\left(\frac{\mu}{\sigma_z}\right)^2 \tag{18}$$

Proof: The cohort hazard rate is given by:

$$-\frac{DS_{ij}(a)}{S_{ij}(a)} = \delta + \frac{\frac{\theta_2\sigma_z\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}}{2\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\sqrt{a}}\frac{\varphi\left(\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\sqrt{a}\right)}{\Phi\left(-\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\sqrt{a}\right)} - \frac{\sigma_z^2}{2}\theta_2^2 - \mu\theta_2}{1+\exp\left\{-a\left(\frac{\sigma_z^2}{2}\theta_2^2+\mu\theta_2\right)\right\}\frac{\Phi\left(\frac{\mu\sqrt{a}}{\sigma_z}\right)}{\Phi\left(-\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\sqrt{a}\right)}} \tag{19}$$

where the derivation of the numerator and the denominator is as follows:

$$DS_{ij}(a) = e^{-\delta a}\frac{\mu a^{-1/2}}{2\sigma_z}\varphi\left(\frac{\mu\sqrt{a}}{\sigma_z}\right) + e^{-\delta a}\left(\frac{\sigma_z^2}{2}\theta_2^2+\mu\theta_2\right)\exp\left\{a\left(\frac{\sigma_z^2}{2}\theta_2^2+\mu\theta_2\right)\right\}\Phi\left(-\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\sqrt{a}\right) +$$
$$-\frac{\mu+\theta_2\sigma_z^2}{2\sigma_z}a^{-1/2}\exp\left\{a\left(\frac{\sigma_z^2}{2}\theta_2^2+\mu\theta_2\right)\right\}\varphi\left(-\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\sqrt{a}\right)e^{-\delta a} - \delta S_{ij}(a)$$

after some manipulation this reduces to

$$DS_{ij}(a) = e^{-\delta a}\left\{-\frac{\theta_2\sigma_z}{2\sqrt{a}}\frac{e^{-\frac{1}{2}\frac{\mu^2}{\sigma_z^2}a}}{\sqrt{2\pi}} + \left(\frac{\sigma_z^2}{2}\theta_2^2+\mu\theta_2\right)e^{a\left(\frac{\sigma_z^2}{2}\theta_2^2+\mu\theta_2\right)}\Phi\left(-\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\sqrt{a}\right)\right\} - \delta S_{ij}(a)$$

Now looking at expression (19) the limit $\frac{\varphi(x)}{\Phi(-x)x}$ as $x\to\infty$ is 1 (see P5) so that the numerator in the expression is given by $-\mu\theta_2/2$. If $\mu<0$, we can use De l' Hospital to compute the same limit for the denominator. It is

$$1+\frac{\mu^2+\sigma_z^2\theta_2^2\sigma_z^2+2\mu\theta_2\sigma_z^2}{-\left(\mu+\theta_2\sigma_z^2\right)\mu} = \frac{\sigma_z^2\theta_2^2\sigma_z^2+\mu\theta_2\sigma_z^2}{-\left(\mu+\theta_2\sigma_z^2\right)\mu} = -\frac{\theta_2}{\mu}\sigma_z^2$$

and this implies equation (18).

Lemma 7 ***I****f* $\mu<0$ *cohort hazard rates are monotonic in* a.

Proof: To prove monotonicity notice that by property P5 the numerator of expression (19) decreases with time. For the denominator we simply have to prove that it increases. The derivative of the denominator wrt to a:

$$\frac{e^{-a\left(\frac{\sigma_z^2}{2}\theta_2^2+\mu\theta_2\right)}\Phi\left(\frac{\mu\sqrt{a}}{\sigma_z}\right)}{\Phi\left(-\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\sqrt{a}\right)} = \frac{e^{-a\left(\frac{\sigma_z^2}{2}\theta_2^2+\mu\theta_2+\left(\frac{\mu}{\sigma_z}\right)^2\right)}}{-a\left(\frac{\mu}{\sigma_z}\right)^2}\frac{\Phi\left(\frac{\mu\sqrt{a}}{\sigma_z}\right)}{\Phi\left(-\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\sqrt{a}\right)}$$

$$= \frac{m\left(\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\sqrt{a}\right)}{m\left(-\frac{\mu}{\sigma_z}\sqrt{a}\right)} \tag{20}$$

where m is again the inverse mills ratio. Define $\tilde{c}_1 \equiv \frac{\mu+\theta_2\sigma_z^2}{\sigma_z} > \tilde{c}_2 \equiv -\frac{\mu}{\sigma_z}$. For $\mu < 0$ and because of assumption A1 (see at the end of this appendix) we have $\tilde{c}_1, \tilde{c}_2 > 0$. In order for the derivative of (20) wrt to $\sqrt{a}$ to be positive it must be:

$$\frac{m'\left(\tilde{c}_1\sqrt{a}\right)}{m\left(\tilde{c}_1\sqrt{a}\right)}\tilde{c}_1\sqrt{a} > \frac{m'\left(\tilde{c}_2\sqrt{a}\right)}{m\left(\tilde{c}_2\sqrt{a}\right)}\tilde{c}_2\sqrt{a} \ .$$

This condition is equivalent to the following auxiliary lemma which completes the proof.

Lemma 8 *Let $y > x > 0$. Then*

$$y\frac{m'(y)}{m(y)} > x\frac{m'(x)}{m(x)} \ ,$$

Proof. Detailed derivations can be found in an online appendix. Here, we sketch the proof. Notice that $m' = m(m-x)$. Using that, $\left(x\frac{m'}{m}\right)' = mx(m-x) - 2x + m$, it suffices to show

$$mx(m-x) - 3m > 2x - 4m \ . \tag{21}$$

It is also true that $m-x > 0$ from property P5 and that (see a lenghty proof by Barrow and Cohen (1954) p. 406 and online appendix) $mx(m-x) - 3m > -\frac{2}{m-x}$. In combination with (21) the last two inequalities imply that it suffices to show

$$-\frac{2}{m-x} > 2x - 4m \implies 0 < (2m-x)(m-x) - 1.$$

This inequality has been proven by Sampford (1953), which completes the proof. ■

7 Firm and Cohort Sales

7.1 Firm Expected Sales

The goal is to compute the expected value of the expression

$$r_{ijt}(z) \equiv p_{ijt}(z)\, q_{ijt}(z) = \begin{cases} L_{jt}^{\alpha} y_{jt} \frac{1}{\tilde{\psi}} \left[e^{\bar{c}_1 \ln\left(z/z_{ijt}^*\right)} - e^{\bar{c}_2 \ln\left(z/z_{ijt}^*\right)} \right] & \text{if } z \geq z_{ijt}^* \\ 0 & \text{otherwise} \end{cases}, \tag{22}$$

where we use the definition of $s_{ija} = \ln z_{\tilde{t}+a}/z_{ij\tilde{t}+a}^*$ (for simplicity denoted by s_a) as a proxy for the relative size of a firm in a given market j. We focus on deriving the expected value of the two terms inside the brackets since the terms outside the brackets are deterministic. Of course, since the term s_a follows a simple Brownian motion with a drift μ and a volatility σ_z we can simply consider each term separately and calculate $E\left(e^{\bar{c}_i s_a} | s_a \geq 0, s_0 \geq 0\right)$ where $\bar{c}_1 = \sigma - 1$, $\bar{c}_2 = (\sigma - 1)/\tilde{\beta}$ with $\tilde{\beta} = \beta/(\beta - 1)$. Then the terms can be combined and multiplied by the values of the deterministic parameters.

Regarding the expected values for $\bar{c}_i$, $i = 1, 2$, we have

$$E\left(e^{\bar{c}_i s_a} | s_a \geq 0, s_0 = s_0\right) = \int_0^{+\infty} \frac{e^{\bar{c}_i s_a}}{\sigma_z \sqrt{a2\pi}} \frac{e^{-\frac{\left(\frac{s_a - s_0 - \mu a}{\sigma_z \sqrt{a}}\right)^2}{2}}}{\Phi\left(\frac{s_0 + \mu a}{\sigma_z \sqrt{a}}\right)} ds_a$$

$$= \frac{e^{\frac{-(s_0)^2 - 2\mu a s_0 - \mu^2 a^2}{\sigma_z^2 a2}}}{\Phi\left(\frac{s_0 + \mu a}{\sigma_z \sqrt{a}}\right) \sigma_z \sqrt{a2\pi}} \int_0^{+\infty} \exp\left\{ -\frac{1}{\sigma_z^2 a2} (s_a)^2 + \frac{2s_0 + 2\mu a + \bar{c}_i \sigma_z^2 a2}{\sigma_z^2 a2} s_a \right\} ds_a$$

using property P9 and P10 the above expression equals to:

$$\frac{e^{\frac{-(s_0)^2 - 2\mu a s_0 - \mu^2 a^2}{\sigma_z^2 a2}}}{\Phi\left(\frac{s_0 + \mu a}{\sigma_z \sqrt{a}}\right) \sigma_z \sqrt{a2\pi}} e^{\frac{\left(\frac{2s_0 + 2\mu a + \bar{c}_i \sigma_z^2 a2}{\sigma_z^2 a2}\right)^2 \sigma_z^2 a2}{4}} \frac{\sqrt{\pi}}{2\sqrt{1/\left(\sigma_z^2 a2\right)}} \left[1 + nerf\left(\frac{s_0 + \mu a + \bar{c}_i \sigma_z^2 a}{\sqrt{\sigma_z^2 a2}} \right) \right]$$

Using P8 this last expression gives

$$E\left(e^{\bar{c}_i s_a} | s_a \geq 0, s_0 = s_0\right) = \exp\left\{ \frac{(\bar{c}_i)^2 \sigma_z^2}{2} a + \bar{c}_i \mu a + \bar{c}_i s_0 \right\} \frac{\Phi\left(\frac{s_0 + \mu a + \bar{c}_i \sigma_z^2 a}{\sigma_z \sqrt{a}}\right)}{\Phi\left(\frac{s_0 + \mu a}{\sigma_z \sqrt{a}}\right)} \tag{23}$$

7.2 Cohort Market Shares

To compute cohort market shares we need to compute the expected sales of a cohort and divide by the total sales in this market by all firms. This ratio at year 0 is 1. The growth rate of the total sales of all firms equals the rate of change of nominal GDP, $g_\kappa + g_\eta$.

The expected sales of the cohort can be computed by finding the expected sales of firms of different sizes, weighted by the density of initial sizes of these firms (integrating 23, unconditional on survival, over the productivity density). To compute the implied integral we split it into two parts (in a manner similar to what was done in the previous subsection) and compute the following integral for $i = 1, 2$:

$$\int_0^\infty \theta_2 e^{-\theta_2(s_0)} e^{a\left(\frac{\bar{c}_i^2\sigma_z^2}{2}+\bar{c}_i\mu\right)+\bar{c}_i s_0} \Phi\left(\frac{s_0+\mu a+\bar{c}_i\sigma_z^2 a}{\sigma_z\sqrt{a}}\right) ds_0$$

$$= \frac{\theta_2 e^{a\left(\frac{\bar{c}_i^2\sigma_z^2}{2}+\bar{c}_i\mu\right)}}{\theta_2-\bar{c}_i}\left\{\Phi\left(\frac{\mu a+\bar{c}_i\sigma_z^2 a}{\sigma_z\sqrt{a}}\right)+\int_0^\infty \frac{e^{(\bar{c}_i-\theta_2)s_0}}{\sigma_z\sqrt{a}}\varphi\left(\frac{s_0+\mu a+\bar{c}_i\sigma_z^2 a}{\sigma_z\sqrt{a}}\right) ds_0\right\} \quad (24)$$

where the second expression is derived using integration by parts. The integral inside the brackets equals to

$$\frac{e^{\frac{-(\mu a)^2-\sigma_z^4(\bar{c}_i a)^2-2\mu a\bar{c}_i\sigma_z^2 a}{2\sigma_z^2 a}}}{\sigma_z\sqrt{2\pi}\sqrt{a}}\int_0^\infty \exp\left\{-\frac{1}{2\sigma_z^2 a}(s_0)^2-\frac{\theta_2\sigma_z^2+\mu}{\sigma_z^2}s_0\right\} ds_0 =$$

$$\frac{e^{\frac{-(\mu a)^2-\sigma_z^4(\bar{c}_i a)^2-2\mu a\bar{c}_i\sigma_z^2 a}{2\sigma_z^2 a}}}{\sigma_z\sqrt{2\pi}\sqrt{a}} e^{\left(\frac{\theta_2\sigma_z^2+\mu}{\sigma_z^2}\right)^2/\left[4\left(\frac{1}{2\sigma_z^2 a}\right)\right]} \frac{\sqrt{\pi}}{2\sqrt{\frac{1}{2\sigma_z^2 a}}}\left[1-\operatorname{erf}\left(\frac{\frac{\theta_2\sigma_z^2+\mu}{\sigma_z^2}}{2\sqrt{\frac{1}{2\sigma_z^2 a}}}\right)\right] =$$

$$e^{\left[\frac{\sigma_z^2}{2}(\theta_2+\bar{c}_i)+\mu\right](\theta_2-\bar{c}_i)a}\Phi\left(-\frac{\left(\theta_2\sigma_z^2+\mu\right)\sqrt{a}}{\sigma_z}\right)$$

where we used the definition of the normal for the first line, the property P10 for the first equality, and property P8 for the second equality. The expression is derived by replacement in expression (24). Notice that the assumption $\theta_2 \geq (\sigma - 1)$ is required that derives from assumption A2. Expression (24) will be added for $i = 1$ and substracted for $i = 2$. Finally, we need to multiply by $e^{g_\kappa + g_\eta \alpha}$ to capture the rate of growth of the average sales of the incumbents and $e^{-\delta a}$ to discount by the exogenous death rate of ideas. This completes the derivation.

7.3 Size and Age in the US Data

In addition to the results on the size distribution of firms of different ages the model implies that firms reach a large size at a reasonable age compared to what is observed in the US data. Considering firms of age 100 years or less, the benchmark model predicts that around 1.2 in 1,000 firms are of the size of 2,500 employees or more. The corresponding number for the population of the US manufacturing firms is around 3.8 in every 1,000. The data are from the Small Business Administration, sba.gov, for 2005. The model was calibrated to the mean employment size of US manufacturing firms in 2005 (47.3 employees) to generate this number by appropriately choosing the value of the parameter ψ.

8 Detailed derivations regarding proposition 2 of the paper

Here we describe in more detail the derivations of proposition 2 of the paper. First, notice the definitions (for simplicity of notation supress subscript notation for s_{ija}): $h(s) = e^{s(\sigma-1)} - e^{s(\sigma-1)/\tilde{\beta}}$, $h'(s) = (\sigma-1)\, e^{s(\sigma-1)} - \frac{(\sigma-1)}{\tilde{\beta}} e^{s(\sigma-1)/\tilde{\beta}}$, $h''(s) = (\sigma-1)^2 e^{s(\sigma-1)} - \frac{(\sigma-1)^2}{(\tilde{\beta})^2} e^{s(\sigma-1)/\tilde{\beta}}$, $\frac{h'(s)}{h(s)} = \frac{(\sigma-1)e^{s(\sigma-1)} - \frac{(\sigma-1)}{\tilde{\beta}} e^{s(\sigma-1)/\tilde{\beta}}}{e^{s(\sigma-1)} - e^{s(\sigma-1)/\tilde{\beta}}} = (\sigma-1)\, \frac{1-\frac{1}{\tilde{\beta}} e^{-s(\sigma-1)(1)/\beta}}{1-e^{-s(\sigma-1)(1)/\beta}}$.

We present results for $\beta \in (0,1) \cup (1,+\infty)$. Results for $\beta = 0, 1$ can be derived by taking limits. We have that,

$$\frac{\partial\left((\sigma-1)\,\frac{1-\frac{1}{\tilde{\beta}} e^{-s(\sigma-1)/\beta}}{1-e^{-s(\sigma-1)(1)/\beta}}\right)}{\partial s} =$$

$$= (\sigma-1)\,(\sigma-1)\,/\beta \frac{\frac{1}{\tilde{\beta}} e^{-s(\sigma-1)/\beta}\left(1-e^{-s(\sigma-1)(1)/\beta}\right) - \left(1-\frac{1}{\tilde{\beta}} e^{-s(\sigma-1)(1)/\beta}\right) e^{-s(\sigma-1)(1)/\beta}}{\left(1-e^{-s(\sigma-1)(1)/\beta}\right)^2}$$

$$= (\sigma-1)\,\frac{(\sigma-1)}{\beta}\left[\frac{\frac{1}{\tilde{\beta}} e^{-s(\sigma-1)/\beta} - \frac{1}{\tilde{\beta}} e^{-s(\sigma-1)/\beta} e^{-s(\sigma-1)/\beta} - e^{-s(\sigma-1)/\beta} + \frac{1}{\tilde{\beta}} e^{-s(\sigma-1)/\beta} e^{-s(\sigma-1)/\beta}}{\left(1-e^{-s(\sigma-1)(1)/\beta}\right)^2}\right]$$

$$= (\sigma-1)\,\frac{(\sigma-1)}{\beta}\left(\frac{\frac{1-\tilde{\beta}}{\tilde{\beta}} e^{-s(\sigma-1)/\beta}}{\left(1-e^{-s(\sigma-1)(1)/\beta}\right)^2}\right)$$

It is also true that,

$$\frac{\partial\left(\mu \frac{h'(s)}{h(s)} + \frac{\sigma_z^2}{2}\frac{h''(s)}{h(s)}\right)}{\partial s} \leq 0 \Leftrightarrow$$

$$\mu(\sigma-1)\,\frac{\partial\left(\frac{1-\frac{1}{\tilde{\beta}} e^{-s\frac{(\sigma-1)}{\beta}}}{1-e^{-s\frac{(\sigma-1)}{\beta}}}\right)}{\partial s} + \frac{\sigma_z^2}{2}(\sigma-1)^2\,\frac{\partial\left(\frac{1-\left(\frac{1}{\tilde{\beta}}\right)^2 e^{-s\frac{(\sigma-1)}{\beta}}}{1-e^{-s\frac{(\sigma-1)}{\beta}}}\right)}{\partial s} \leq 0 \Leftrightarrow$$

$$\mu(\sigma-1)\,\frac{\frac{(\sigma-1)}{\beta}\left(\frac{1-\tilde{\beta}}{\tilde{\beta}}\right) e^{-s\frac{(\sigma-1)}{\beta}}}{\left(1-e^{-s\frac{(\sigma-1)}{\beta}}\right)^2} + \frac{\sigma_z^2}{2}(\sigma-1)^2\,\frac{\frac{(\sigma-1)}{\beta}\frac{1-\tilde{\beta}^2}{\tilde{\beta}^2} e^{-s\frac{(\sigma-1)}{\beta}}}{\left(1-e^{-s\frac{(\sigma-1)}{\beta}}\right)^2} \leq 0 \Leftrightarrow$$

$$
\begin{aligned}
\mu\left(\sigma-1\right)\frac{\left(\sigma-1\right)}{\beta}\left(\frac{1-\tilde{\beta}}{\tilde{\beta}}\right)e^{-s\frac{(\sigma-1)}{\beta}}+\frac{\sigma_z^2}{2}\frac{\left(\sigma-1\right)}{\beta}\left(\sigma-1\right)^2\frac{1-\tilde{\beta}^2}{\tilde{\beta}^2}e^{-s\frac{(\sigma-1)}{\beta}} &\leq 0\Leftrightarrow\\
\mu\left(1-\tilde{\beta}\right)\tilde{\beta}e^{-s\frac{(\sigma-1)}{\beta}}+\frac{\sigma_z^2}{2}\left(\sigma-1\right)^2\left(1-\tilde{\beta}^2\right)e^{-s\frac{(\sigma-1)}{\beta}} &\leq 0\Leftrightarrow\\
\left(\sigma-1\right)\mu\left(\tilde{\beta}-\tilde{\beta}^2\right)+\frac{\sigma_z^2}{2}\left(\sigma-1\right)^2\left(1-\tilde{\beta}^2\right) &\leq 0\Leftrightarrow
\end{aligned}
$$

$$
\begin{aligned}
\left[\frac{\sigma_z^2}{2}\left(\sigma-1\right)^2+\left(\sigma-1\right)\mu\right]\tilde{\beta}^2-\left(\sigma-1\right)\mu\left(\tilde{\beta}\right)-\frac{\sigma_z^2}{2}\left(\sigma-1\right)^2 &\geq 0\Leftrightarrow\\
\left[\frac{\sigma_z^2}{2}\left(\sigma-1\right)^2+\left(\sigma-1\right)\mu\right]\left(\frac{\beta}{\beta-1}\right)^2-\left(\sigma-1\right)\mu\left(\frac{\beta}{\beta-1}\right)-\frac{\sigma_z^2}{2}\left(\sigma-1\right)^2 &\geq 0\Leftrightarrow\\
\left[\frac{\sigma_z^2}{2}\left(\sigma-1\right)^2+\left(\sigma-1\right)\mu\right]\left(\beta\right)^2-\left(\sigma-1\right)\mu\beta\left(\beta-1\right)-\left(\beta-1\right)^2\frac{\sigma_z^2}{2}\left(\sigma-1\right)^2 &\geq 0\Leftrightarrow\\
\left[\frac{\sigma_z^2}{2}\left(\sigma-1\right)^2+\left(\sigma-1\right)\mu\right]\left(\beta\right)^2-\left(\sigma-1\right)\mu\left(\beta^2-\beta\right)-\left(\beta^2-2\beta+1\right)\frac{\sigma_z^2}{2}\left(\sigma-1\right)^2 &\geq 0\Leftrightarrow
\end{aligned}
$$

$$
\begin{aligned}
-\left(\sigma-1\right)\mu\left(-\beta\right)-\left(-2\beta+1\right)\frac{\sigma_z^2}{2}\left(\sigma-1\right)^2\geq 0\Leftrightarrow\\
\beta\left[\left(\sigma-1\right)\mu+\sigma_z^2\left(\sigma-1\right)^2\right]\geq\frac{\sigma_z^2}{2}\left(\sigma-1\right)^2
\end{aligned}
\tag{26}
$$

Notice that if $\left[\left(\sigma-1\right)\mu+\sigma_z^2\left(\sigma-1\right)^2\right]<0$ this condition cannot be true. When the opposite inequality holds, $\left[\left(\sigma-1\right)\mu+\sigma_z^2\left(\sigma-1\right)^2\right]>0$, we have that the condition is true if

$$
\beta\geq\frac{\frac{\sigma_z^2}{2}\left(\sigma-1\right)^2}{\left(\sigma-1\right)\mu+\sigma_z^2\left(\sigma-1\right)^2}\ .
$$

9 An Extension: Multivariate Brownian Motion

In this section we develop a multi-country generalization of the model with universal productivity advances outlined in the previous section. The extension allows for the productivity of producing the good in different markets to be imperfectly correlated. Of course, since productivity and demand are isomorphic in terms of sales in this model this process can be interpreted as giving foundations for partially correlated demand across markets. The purpose of this section is to lay out the theoretical foundations of a generalized framework useful for future research planning to use this firm-level panel data information.[2] Thus, the objective is to simply facilitate future related work given that the paper operates at a more shallow level, i.e. moments of this firm-level data.

The modeling of the firm's optimization problem is similar to the one introduced in the paper. The difference will be the stochastic process for the productivity of the firm in each country. We will thus now consider directly the process for the logarithm of the productivity of each idea the process for the proxy of firm size,

$$s_a = \bar{s}_i + (g_I - g_E)\, a + \sigma_z W_a \ , \tag{27}$$

to a process that will be potentially different across different countries.

Define the process $\mathbf{W}(a)^T = [W_{1a}, ..., W_{Na}]$ composed of independent simple Brownian Motions where the superscript T denotes the transpose of the matrix. Let $\tilde{V}$ an $N \times N$ covariance matrix that is symmetric and positive definite. Given that $\tilde{V}$ is positive definite it can be written as $\tilde{V} = VV^T$ where $V = \{v_{jk}\}$ is an $N \times N$ nonsingular matrix (the reverse statement is also true, see theorem 23.18 Simon and Blume (1994)). $\bar{\mathbf{s}}_i^T = [\bar{s}_{i1}, ..., \bar{s}_{iN}]$ is a matrix of the logarithm of initial productivity sizes of the ideas in each destination country and $\mu_i^T = [\mu_{i1}, ..., \mu_{iN}]$.

Consider the process of the logarithms of the productivity of a given idea from i selling

[2] As shown in Luttmer (2007), and extending the reasoning to multiple countries, the only difference that this will imply is that the sales of the firm will be depending on the product of the country specific demand shock to the productivity shock. I denote this product with a single term.

to different destination markets,

$$\mathbf{S}_{ia} = \bar{\mathbf{s}}_i + \mu_i a + V\mathbf{W}$$

where $\mathbf{S}_{ia}^T = [s_{i1a}, ..., s_{iNa}]$. This means that

$$s_{ija} = \bar{s}_{ij} + \mu_{ij} a + v_{j1} W_{1a} + ... + v_{jN} W_{Na} \text{ for } j = 1, ..., N \ . \tag{28}$$

Standard results for the normal distribution imply that s_{ija} is normally distributed as

$$s_{ija} \sim \mathcal{N}\left(\bar{s}_{ij} + \mu_{ij} a, \left[(v_{j1})^2 + ... + (v_{jN})^2\right] a\right) \ , \tag{29}$$

as long as it is considered independent from the other s_i's. Using the moment generating function of the distribution (see Serfozo (1994) p. 345) the joint distribution at time a is given by a multivariate normal

$$f(s_{i1a}, ..., s_{iNa}) = \frac{1}{\sqrt{(2\pi a)^n \left|\tilde{V}\right|}} \exp\left\{-(\mathbf{S}_{ia} - \bar{\mathbf{s}}_i - \mu_i a)^T \frac{\tilde{V}^{-1}}{2} (\mathbf{S}_{ia} - \bar{\mathbf{s}}_i - \mu_i a)\right\}$$

where $\left|\tilde{V}\right|$ is the determinant, and $\tilde{V}^{-1}$ the inverse of matrix $\tilde{V}$.

A few points are in order. First, notice that since the distribution of s_{ija} is normal, the cross-sectional distribution of the exponential of these shocks will be double-Pareto, since the proof and the analysis discussed in Section 2.5 of the main paper applies. Second, notice that the matrix $\tilde{V}$ is the variance-covariance matrix for each process s_{ija}, $j = 1, ...N$, which determines how the growth rate of the processes is correlated across markets. Third, notice that this process combined with firm entry, results to a stationary distribution cross-sectional of shocks that is double Pareto for each market. Thus, firm outcomes at the market are the same as in the main text (e.g. growth as a function of initial firm size) and the aggregate equilibrium is the same as well (since the equilibrium at each date t depends on the cross-sectional distribution of shocks).

Balanced Growth Path Extending the analysis to more general correlation matrices $\tilde{V}$ might require respecifying the model in order to solve for the balanced growth path.

To simplify the analysis, and create a direct generalization of the previous results consider that V is of the form

$$V = \begin{bmatrix} x & \bar{x} & ... & \bar{x} \\ \bar{x} & x & ... & \bar{x} \\ ... & ... & ... & ... \\ \bar{x} & \bar{x} & ... & x \end{bmatrix} = V^T ,$$

x, $\bar{x} \in (-\infty, +\infty)$, which is nonsingular if $x \neq \bar{x}$. Since V is nonsingular for $x \neq \bar{x}$, $\tilde{V}$ is positive definite for all $x \neq \bar{x}$. Its diagonal elements are given by $x^2 + \bar{x}^2 (N-1)$, while the off-diagonal ones are equal to $2x\bar{x} + \bar{x}^2 (N-2)$.

To create a simple correspondence with the previous model we now set

$$(x)^2 + (N-1)(\bar{x})^2 = \sigma_z^2$$

and $\mu_{ij} = \mu$, $\bar{s}_{ij} = \bar{s}_i$, $\forall i, j$. Given the covariance matrix $\tilde{V}$ the correlation of s_{ija}'s with $N \geq 2$ will be given by

$$\frac{2x\bar{x} + \bar{x}^2 (N-2)}{(x)^2 + (N-1)(\bar{x})^2} \in (-1, 1) \ .$$

Obsviously for $x = \bar{x}$ we are back to the model with perfect correlation of productivity shocks. In the simple example constructed above the analysis of the previous sections carries out intact for the more general case when productivities are imperfectly correlated: Given (28) and the assumption about the covariance matrix the new process will have identical implications to the one studied up to now for the cross-section and the growth of sales of ideas and firms in a given destination country. Thus, the mapping to the dynamic and static version of Chaney (2008) and Arkolakis (2010) is essentially intact. In addition, the correlation of sales, can be estimated by future researchers using panel data for the sales of firms to individual markets, ideally including the domestic economy.

10 Numerical Procedure for Computing the Equilibrium

A number of steps

Step 1 (generate a distribution of ideas): Simulate a generation of ideas. Start from $f(x, 0)$ as given, and define a stochastic process for the evolution of x_a. Track its distribution $f(x, a)$ over time a. Define some weight $g_B > 0$ and calculate stationary density

$$f(x) = \int_0^t g_B e^{-g_B a} f(x, a)$$

for some large t. Alternatively, an analytical form for $f(x, a)$ could be used to do the same procedure. Program: GEN_STATIONARY_DIST.m

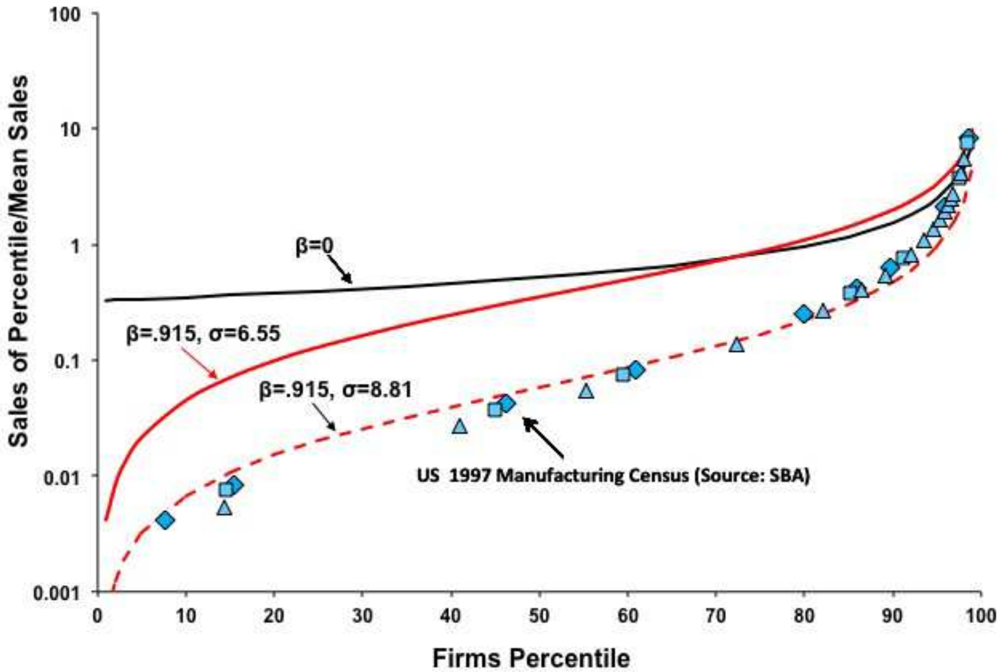
Step 2 (compute the aggregate model): Step 2 takes as an input $f(x)$ constructed in Step 1. This step can be run independent of step 1 for any given probability density function $f(x)$. Additional inputs to the program are aggregate shocks (aggregate productivity and trade costs) and all the rest of the model parameters. The program solve for the equilibrium in each period of time using the program and numerically integrating over the density and then computing aggregate bilateral sales and all equilibrium outcomes using the equilibrium conditions in Section 2.5. The model solves for a set of equilirbium wages, w_{it}, and incomes, y_{it}, for each country i and each date t. Notice that this part of the code does not require to simulate firms. Program: MELITZ_EQUILIBRIUM.m

Step 3 (compute firm statistics): Simulate a large of productivities of ideas using the stochastic process of choice. Use equation (11) in the main text and the aggregate outcomes from Step 2 to compute individual firm sales in each market. Then compute all relevant firm statistics. Programs: SIMULATE_FIRMS.m and STATISTICS.m

11 Robustness: Size Distribution

For robustness, I provide an alternative parameter specification of the endogenous cost model by explicitly calibrating it to the size distribution of exporter sales in both the

Figure 1: Distribution of sales; data and the model



Note: Data from the Small Business Administration (SBA) 1997, 2002, and 2007 censuses. The maximum point of each bin reported in the data and the corresponding number of firms is used to plot the sales size of the firms and the corresponding percentiles.

French and US data. Increasing β above the benchmark calibration value of .915 has small effects on the size distribution. Thus, given $\beta = .915$, setting $\tilde{\theta} = 1.06$ as in Luttmer (2007) provides a much better fit of the model for the firm size distribution of the smaller firms. The fit of the alternative model calibrations are illustrated in Figure 1. Given $\theta_2 = 8.28$, the alternative calibration of the endogenous cost model requires $\sigma = 8.81$.[3]

The alternative calibration of the endogenous cost model implies a similar distribution of growth rates as the baseline calibration with slightly higher growth rates for all per-

[3]Luttmer also sets $\tilde{\theta} = 1.06$ to ensure that the model matches the estimate of the upper tail of the size distribution by Axtell (2001).

centiles (because of the higher elasticity of substutitution, σ). However, it susbstantially overpredicts the market shares of surviving firms. For example, it predicts that the market share of surviving firms after 10 years amounts to about 87%. Dunne, Roberts, and Samuelson (1988) report an average of 73% across different manufacturing census cohorts in the US data.

12 Appendix: Properties of the Normal Distribution

The following assumptions are used in the paper

Assumption 1 *: The rate of innovation is positive,* $g_B > 0$.

Assumption 2 *: Productivity and sales parameters satisfy*

$$g_B > \max\left\{\mu + \sigma_z^2/2\ ,\ (\sigma-1)\,\mu + (\sigma-1)^2\,\sigma_z^2/2\right\}\ .$$

In the various proofs and derivations of the paper and the appendix we use the following definitions and well known facts for the Normal distribution quoted as **properties P**.

Property 1 *The simple normal distribution with mean 0 and variance 1 is given by* $\varphi(x) = \frac{e^{-x^2}/2}{\sqrt{2\pi}}$.

Property 2 *The cdf of the normal is given by* $\Phi\left(\frac{x-\mu}{\sigma_z}\right) = \frac{1}{\sigma_z\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left\{-\frac{(\tilde{x}-\mu)^2}{2\sigma_z^2}\right\}d\tilde{x}$. *Using change of variables* $\upsilon = (\tilde{x}-\mu)/\sigma_z$ *which implies* $d\upsilon = d\tilde{x}/\sigma_z$ *it is also true that*

$$\Phi\left(\frac{x-\mu}{\sigma_z}\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{x-\mu}{\sigma_z}} e^{-\frac{(\tilde{x})^2}{2}}d\tilde{x}$$

Property 3 *Because of the symmetry of the normal distribution,* $\varphi(x) = \varphi(-x)$ *and* $\Phi(x) = 1-\Phi(-x)$.

Property 4 *The inverse mill's ratio of the Normal,* $\varphi(x)/\Phi(-x)$, *is increasing in* x, $\forall x \in (-\infty,+\infty)$.

Property 5 $\varphi(x)/\Phi(-x)/x$ *is decreasing in* x, $\forall x \in (0,+\infty)$ *with* $\lim_{x\to\infty}\varphi(x)/\Phi(-x)/x = 1$. *This implies that* $\varphi(x)/(1-\Phi(x)) > x$ *for* $\forall x \in (-\infty,+\infty)$

Property 6 $\Phi(x+\tilde{c})/\Phi(x)$, *with* $\tilde{c} > 0$, *is decreasing in* x, $\forall x \in (-\infty,+\infty)$.

Property 7 *The error function is defined by:* $nerf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-(\tilde{x})^2} d\tilde{x}$.

Property 8 $\Phi(x) = \frac{1}{2}\left[1 + nerf\left(\frac{x}{\sqrt{2}}\right)\right]$, *where* $\Phi(x)$ *is the cdf of the standard normal cdf*

Property 9 *The error function is odd:* $nerf(-x) = -nerf(x)$. *Also* $\lim_{x\to+\infty} nerf\,(x) = 1$.

Property 10 $\int e^{-\tilde{c}_1 x^2 + \tilde{c}_2 x} dx = e^{(\tilde{c}_2)^2/4(\tilde{c}_1)} \sqrt{\pi} nerf\left(\frac{2\tilde{c}_1 x - \tilde{c}_2}{2\sqrt{\tilde{c}_1}}\right) / \left(2\sqrt{\tilde{c}_1}\right)$, *for some constants* $\tilde{c}_1,\ \tilde{c}_2 > 0$

References

Allen, T., C. Arkolakis, and X. Li (2014): "On the existence and uniqueness of trade equilibria," *mimeo, Northwestern and Yale Universities.*

Arkolakis, C. (2010): "Market Penetration Costs and the New Consumers Margin in International Trade," *Journal of Political Economy*, 118(6), 1151–1199.

Axtell, R. L. (2001): "Zipf Distribution of U.S. Firm Sizes," *Science*, 293(5536), 1818–1820.

Barrow, D. F., and A. C. Cohen (1954): "On Some Functions Involving Mill's Ratio," *Annals of Mathematical Statistics*, 25, 405–408.

Chaney, T. (2008): "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, 98(4), 1707–1721.

Dunne, T., M. J. Roberts, and L. Samuelson (1988): "Patterns of Firm Entry and Exit in US Manufacturing Industries," *The RAND Journal of Economics*, 19(4), 495–515.

Harrison, M. (1985): *Brownian Motion and Stochastic Flow Systems*. John Wiley and Sons, New York.

Klette, J., and S. Kortum (2004): "Innovating Firms and Aggregate Innovation," *Journal of Political Economy*, 112(5), 986–1018.

Luttmer, E. G. J. (2007): "Selection, Growth, and the Size Distribution of Firms," *Quarterly Journal of Economics*, 122(3), 1103–1144.

Reed, W. J. (2002): "On the Rank-Size Distribution for Human Settlements," *Journal of Regional Science*, 42(1), 1–17.

Sampford, M. P. (1953): "Some Inequalities on Mill's Ratio and Related Function," *The Annals of Mathematical Statistics*, 24(1), 130–132.

Serfozo, R. (1994): *Basics of Applied Stochastic Processes*, Probability and Its Applications. Springer, Berlin.

Simon, C. P., and L. Blume (1994): *Mathematics for Economists*. W. W. Norton and Company, New York.