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

Costas Arkolakis†
Yale University, NBER & CESifo

June 6, 2009

Abstract

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

Contents

*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. This paper previously circulated under the title "Market Penetration Costs and Trade Dynamics"

†Contact: Department of Economics, Yale University, 37 Hillhouse ave., New Haven, CT, 06511. Email: costas.arkolakis@yale.edu. web: http://www.econ.yale.edu/~ka265/research.htm

1 Proof for expected time to reach a certain size

We assume that $s_0 = 0$ and consider the time that it takes to reach $s_b > 0$. We want to compute the probability the first hitting time of an idea, conditional on that idea not being hit by an exogenous death shock. Essentially, we want to compute $P(T(s = s_b) > t|\text{Not Death})$ where $T(s = s_b)$ is the first time that $s = s_b$. To do that we simply have to compute

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

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

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

Thus, we are essentially looking at

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

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

$$\frac{s_b}{\sigma_z\sqrt{2\pi}}e^{\frac{s_b\mu}{\sigma_z^2}}\int_0^t e^{-\frac{1}{2}\frac{s_b^2}{\sigma_z^2 a}-\frac{1}{2}\frac{(\mu)^2+2\delta\sigma_z^2}{\sigma_z^2}a}\frac{1}{a^{\frac{3}{2}}}da =$$

$$-\frac{s_b}{\sigma_z\sqrt{2\pi}}e^{\frac{s_b\mu}{\sigma_z^2}}\int_0^t e^{-\frac{1}{2}\frac{s_b^2}{\sigma_z^2}\tilde{a}-\frac{1}{2}\frac{(\mu)^2+2\delta\sigma_z^2}{\sigma_z^2}\frac{1}{\tilde{a}}}\tilde{a}^{-\frac{1}{2}}d\tilde{a} =$$

$$-\frac{s_b}{\sigma_z\sqrt{2\pi}}e^{\frac{s_b\mu}{\sigma_z^2}}\frac{1}{2\sqrt{\frac{1}{2}\frac{s_b^2}{\sigma_z^2}}}e^{-2\sqrt{\frac{1}{2}\frac{s_b^2}{\sigma_z^2}}\sqrt{\frac{1}{2}\frac{(\mu)^2+2\delta\sigma_z^2}{\sigma_z^2}}}\sqrt{\pi}\times$$

$$\left| -\operatorname{erf}\left(\frac{\sqrt{\frac{1}{2}\frac{(\mu)^2+2\delta\sigma_z^2}{\sigma_z^2}}-\sqrt{\frac{1}{2}\frac{s_b^2}{\sigma_z^2}}x}{\sqrt{x}}\right)+e^{4\sqrt{\frac{1}{2}\frac{s_b^2}{\sigma_z^2}}\sqrt{\frac{1}{2}\frac{(\mu)^2+2\delta\sigma_z^2}{\sigma_z^2}}}\left[\operatorname{erf}\left(\frac{\sqrt{\frac{1}{2}\frac{s_b^2}{\sigma_z^2}}x+\sqrt{\frac{1}{2}\frac{(\mu)^2+2\delta\sigma_z^2}{\sigma_z^2}}}{\sqrt{x}}\right)-1\right]+1\right|_{+\infty}^{1/t} =$$

$$-e^{\frac{s_b\mu}{\sigma_z^2}-\sqrt{\frac{s_b^2}{\sigma_z^2}}\sqrt{\frac{(\mu)^2+2\delta\sigma_z^2}{\sigma_z^2}}}\frac{1}{2}\times$$

$$\left[\begin{array}{c} -\operatorname{erf}\left(\frac{\sqrt{\frac{1}{2}\frac{(\mu)^2+\delta\sigma_z^2}{\sigma_z^2}}-\sqrt{\frac{1}{2}\frac{s_b^2}{\sigma_z^2}\frac{1}{t}}}{\sqrt{\frac{1}{t}}}\right)+e^{4\sqrt{\frac{1}{2}\frac{s_b^2}{\sigma_z^2}}\sqrt{\frac{1}{2}\frac{(\mu)^2+\delta\sigma_z^2}{\sigma_z^2}}}\left[\operatorname{erf}\left(\frac{\sqrt{\frac{1}{2}\frac{s_b^2}{\sigma_z^2}\frac{1}{t}}+\sqrt{\frac{1}{2}\frac{(\mu)^2+\delta\sigma_z^2}{\sigma_z^2}}}{\sqrt{\frac{1}{t}}}\right)-1\right]- \\ -\left[-(-1)\right]+0 \end{array}\right]$$

which finally gives the probability we are looking for is

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

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

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 $t = 1, 2, ...$

conditional on its sales at $t = 0$ the numbers would be very similar, as numerical simulations imply. In addition, this derivation can be done analytically using the cdf of the probability distribution of productivities.

2 Additional result for the cohort survival rate

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

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

First, we prove that the survival function is decreasing in a, for $\mu < 0$. To show that it suffices to show that $DS_{ij}(a) < 0$. Notice that the derivative with respect to a of the part of the expression $S_{ij}(a)$ 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 $S_{ij}(a)$ decreases in a when $\mu < 0$ it suffices to show that the term in the bracket is negative which implies that

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

Given that $\sigma_z^2\theta_2 + \mu > 0$, as implied by assumption 2, and $\theta_2 > 1$, we can make use of property F5. This property implies that expression (2) 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 $\mu < 0$.

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 $S_{ij}(a)$ with respect to μ, I want 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.

Regarding the derivative wrt 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.

3 Additional Results for the Cohort Hazard Rates

In order to derive the expression for the cohort hazard rate, first compute the derivative of the survival function $S_{ij}(a)$:

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

after some manipulation this reduces to

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

The cohort hazard rate is given by:

$$-\frac{DS_{ij}(a)}{S_{ij}(a)} = \delta + \frac{\dfrac{\theta_2\sigma_z\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}}{2\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\sqrt{a}}\dfrac{\varphi\left(\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\sqrt{a}\right)}{\Phi\left(-\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\sqrt{a}\right)} - \dfrac{\sigma_z^2}{2}\theta_2^2 - \mu\theta_2}{1+\exp\left\{-a\left(\frac{\sigma_z^2}{2}\theta_2^2+\mu\theta_2\right)\right\}\dfrac{\Phi\left(\frac{\mu\sqrt{a}}{\sigma_z}\right)}{\Phi\left(-\frac{\mu+\theta_2\sigma_z^2}{\sigma_z}\sqrt{a}\right)}} \tag{3}$$

The limit $\frac{\varphi(x)}{\Phi(-x)x}$ as $x \to \infty$ is 1 (see F5) so that the numerator in the expression is given by $-\mu\theta_2/2$. For $\mu < 0$ I have to use De l' Hospital to compute the same limit for the denominator. It is

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

and this implies the result:

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

4 More derivations regarding proposition 5 of the paper

Here we describe in more detail the derivations of proposition 5 of the paper. First, notice the definitions (for simplicity of notation supress subscript notation for s_{ija}): $h(s) = e^{s(\sigma-1)} - e^{s(\sigma-1)/\tilde{\beta}}$, $h'(s) = (\sigma-1)e^{s(\sigma-1)} - \frac{(\sigma-1)}{\tilde{\beta}}e^{s(\sigma-1)/\tilde{\beta}}$, $h''(s) = (\sigma-1)^2e^{s(\sigma-1)} - \frac{(\sigma-1)^2}{(\tilde{\beta})^2}e^{s(\sigma-1)/\tilde{\beta}}$, $\frac{h'(s)}{h(s)} = \frac{(\sigma-1)e^{s(\sigma-1)}-\frac{(\sigma-1)}{\tilde{\beta}}e^{s(\sigma-1)/\tilde{\beta}}}{e^{s(\sigma-1)}-e^{s(\sigma-1)/\tilde{\beta}}} = (\sigma-1)\frac{1-\frac{1}{\tilde{\beta}}e^{-s(\sigma-1)(1)/\beta}}{1-e^{-s(\sigma-1)(1)/\beta}}$.

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,

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

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

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

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

It is also true that,

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

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

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

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

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

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

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

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

5 An Extension: Multivariate Brownian Motion

In this section I 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. 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.[1] 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. I will thus now consider directly the process for the logarithm of the productivity of each idea the process for the proxy of firm size,

$$s_a = \bar{s}_i + (g_I - g_E)\,a + \sigma_z W_a \ , \tag{4}$$

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 $\boldsymbol{\mu}_i^T = [\mu_{i1}, ..., \mu_{iN}]$.

Consider the process of the logarithms of the productivity of a given idea from i selling to different destination markets,

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

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

$$s_{ija} = \bar{s}_{ij} + \mu_{ij}a + v_{j1}W_{1a} + ... + v_{jN}W_{Na} \text{ for } j = 1, ..., N \ . \tag{5}$$

Standard results for the normal distribution imply that s_{ij} is normally distributed as $s_{ija} \sim \mathcal{N}\left(\bar{s}_{ij}+\mu_{ij}a, \left[(v_{j1})^2+...+(v_{jN})^2\right]a\right)$ as long as it is considered independent from the other s_i's.

[1] 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.

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

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

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

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 I 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 (5) 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 (2008) 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.

6 Quantitative Results for the CES Benchmark

In this section I present the quantitative results for the CES benchmark (marked as $\beta = 0$) in the main paper. The results for the CES demand ($\beta \rightarrow 0$) are illustrated with black dots while the results for the calibrated model of the paper with red dots. The pictures that I include are the ones where the two models give different predictions (the predictions of the model for firm entry and exit are independent of β). In particular the large differences in the initial market shares of new entrants that the model with $\beta = 0, 1$ imply (figures 1, 2) are explained by the fact that the model with $\beta = 0$ implies counterfactually large size and lower growth rates for new firms (figure 4).

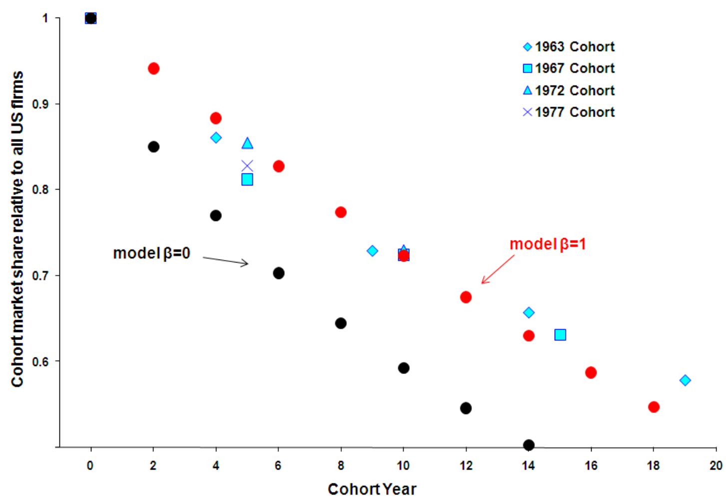

Figure 1: Market share of Cohort Survivors and new entrants in US manufacturing Census and Model. Data: (Dunne, Roberts, and Samuelson 1988)

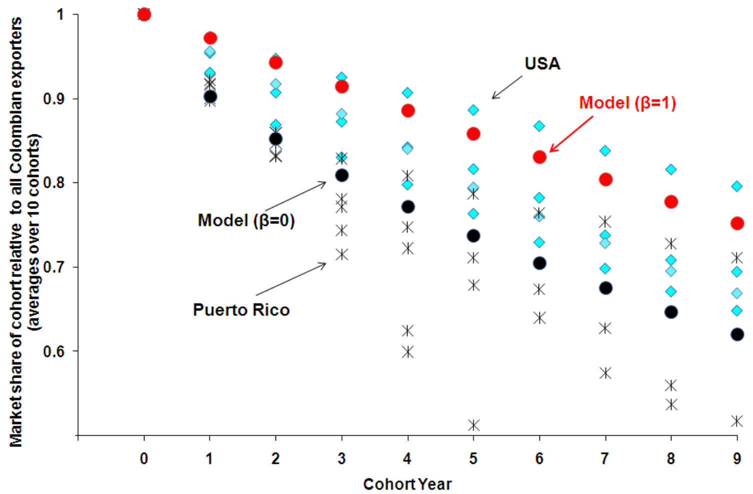

Figure 2: Market share of Cohort Survivors and new entrants into individual destinations for Colombian exporters. Data kindly provided by (Eaton, Eslava, Kugler, and Tybout 2008).

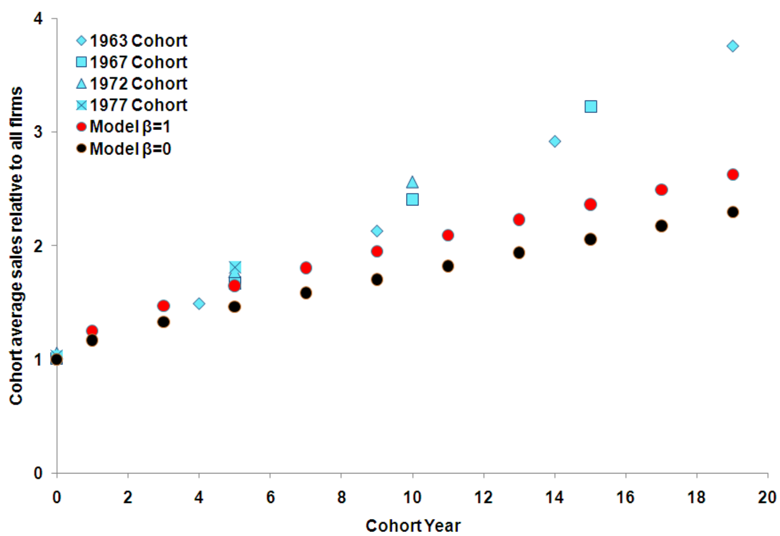

Figure 3: Average sales of Cohort Survivors and all census firms in US manufacturing Census and Model. Data: (Dunne, Roberts, and Samuelson 1988), combining table 9 and 10. (Numbers may contain small aproximation error 1%-2% due to rounding).

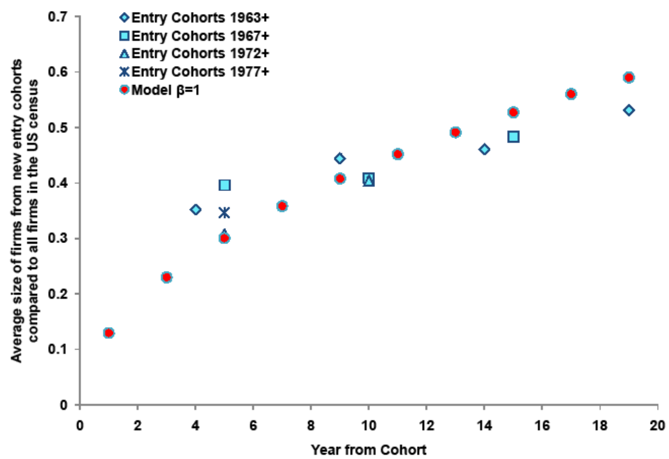

Figure 4: Average sales of firms that are new compared to starting Cohort Survivors and average sales of all census firms in US manufacturing Census and Model. Data: (Dunne, Roberts, and Samuelson 1988), combining table 9 and 10. (Numbers may contain small aproximation error 1%-2% due to rounding).

7 Appendix: Properties of the Normal

In the various proofs and derivations of the paper and the appendix I am going to use the following definitions and well known facts for the Normal distribution quoted as **properties F**.

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

F 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*

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

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

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

F 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* $\forall x \in (-\infty,+\infty)$

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

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

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

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

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

References

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.

EATON, J., M. ESLAVA, M. KUGLER, AND J. TYBOUT (2008): "The Margins of Entry Into Export Markets: Evidence from Colombia," in *Globalization and the Organization of Firms and Markets*, ed. by E. Helpman, D. Marina, and T. Verdier. Harvard University Press, Massachusetts.

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.

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.