Intangibles and Endogenous Firm Volatility over the Business Cycle*

Pablo N D'Erasmo† — University of Maryland
Hernan J Moscoso Boedo‡ — University of Virginia

December 20, 2011
Preliminary

Abstract

We are interested in the endogenous determination of firm level idiosyncratic volatility and its evolution over the business cycle. Using data from the Kauffman Firm Survey and Compustat, we find that idiosyncratic volatility at the firm level is negatively correlated with intangible expenditures (e.g. advertising, marketing, brand development, R&D). We also find that intangible expenses are highly pro-cyclical and that firm level volatility is counter-cyclical. To understand this mechanism, we propose a firm dynamics model with endogenous market participation. Firms that incur higher intangible expenses are able expand the firm and end up diversifying market-specific demand shocks by servicing more markets. The model is driven only by first moment shocks (i.e. shocks to aggregate TFP) and is able to capture the relationship between intangibles and risk as well as their cyclical properties.

Keywords: Endogenous idiosyncratic risk.
JEL Classifications: D31, E4, L11, L26 O16

*We thank Rui Castro, Larry Christiano, John Willis and the seminar participants at the 2011 LACEA meeting in Chile, 2011 SED meeting in Ghent, and the University of Virginia for their very helpful comments. All errors are of course ours. Certain data included herein are derived from the Kauffman Firm Survey release 4.0. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Ewing Marion Kauffman Foundation.

†Correspondence: University of Maryland, Department of Economics, 3105 Tydings Hall, College Park, MD 20742, (301)405 3529, derasmo@econ.umd.edu

‡Correspondence: University of Virginia, Department of Economics, PO Box 400182, Charlottesville, VA 22904, (434)924 7654, hmoscoso@virginia.edu. Financial support from the Bankard Fund for Political Economy is gratefully acknowledged

1 Introduction

The objective of this paper is to understand the determinants of idiosyncratic volatility at the firm level and its evolution over the business cycle.[1] To that end we proceed as follows. First, we empirically document how firm level risk evolves over the business cycle and its relation to intangible expenditures. Second, we propose a theoretical model that is able to interpret the salient features of the data: an industry dynamics model in which firms use intangible expenditures to expand the reach of the firm.

The contribution of our empirical analysis is to document the relationship between firm-level risk and intangible expenses, as well as the cyclical properties of both. We use two sources that capture firms on opposite ends of the size distribution. Large, public firms are observed in Compustat, whereas small, entrepreneurial firms are captured in the Kauffman Foundation Firm Survey. As in Castro, Clementi, and MacDonald (2009), we measure firm level idiosyncratic risk as the portion of growth in sales that cannot be explained by firm level characteristics (such as age or size), industry, or year effects.[2] Intangible expenses correspond to selling and general expenses such as advertising, marketing, brand development, and research and development. The represent, on average, 17% of sales. Our findings are that the elasticity of firm-level volatility with respect to expenditures ranges between -8.5% and -30.1%, volatility is pro cyclical, and intangible expenditures are counter-cyclical.

In addition to our empirical analysis, we present a simple model to understand the findings. In this model, there is a representative consumer who derives utility from market-(or location-)specific consumption goods that are imperfect substitutes. The consumer faces iid location-specific taste shocks that results in stochastic market demands. There is also a continuum of competitive firms that differ in their level of productivity and can sell to

[1]We will refer to firm level volatility, risk, and uncertainty interchangeably.

[2]In the appendix, we present an alternative based on TFPR (total factor productivity revenue based) with similar results. However, measurement issues associated with physical capital and factor shares determines that our preferred firm level volatility criterion is based on sales growth.

many markets by incurring intangible expenses. Intangible expenses are increasing in the total number of markets that a firm services, and are incurred before taste shocks are realized. Thus, these endogenous sunk costs determine the pool of suppliers in each market. By reaching more markets, firms both increase their revenues and diversify market-specific shocks. This generates the negative relationship between intangible expenses and firm-level risk that is observed in the data. Moreover, as is also consistent with the data, high productivity firms will expand to more markets than low-productivity firms, making them less volatile. Finally, besides demand shocks, the model also incorporates an aggregate productivity shock. The incentives to expand are higher in good times. Hence, the model captures the pro-cyclicality of intangible expenses and, because higher intangible expenses result in lower firm-level volatility, the observed dispersion of firm-level risk is counter-cyclical.

Our paper is related to the empirical literature analyzing firm level-risk. Castro, Clementi and Lee (2010) attribute differences in firm-level volatility to differences in the sectors in which firms operate. Differing from them, we uncover the relationship between firm-level volatility and intangible expenses (after controlling for industry effects). Comin and Phillipon (2005) document the increasing trend in firm level volatility using Compustat, whereas Davis et. al. (2006) show that the increase in firm-level risk is related to a selection issue, making public companies more volatile. Finally, Kehrig (2011) uses establishment-level data to document that risk is counter-cyclical and that this correlation is stronger in non-durable industries.

Regarding the cyclical properties of the model, this work can be understood as a theory of endogenous volatility. Along this dimension, the closest work is Bachman and Moscarini (2011), however our mechanism is different. In Bachman and Moscarini, firms experiment with prices in downturns, making them more volatile, whereas in this paper we introduce firm expansion and contraction as a function of total factor productivity shocks. The endogenous volatility literature also builds on the exogenous volatility papers that looked at the effects of changes in the second moments. This literature includes Bloom (2009), Arellano, Bai, and Kehoe (2011), and Christiano, Motto, and Rostagno (2011).

We build on the literature on multi-products firms such as in Bernard, Redding, and Schott (2010) in which they endogenously allow for the expansion of the firm but do not look at the risk dimension. Other related papers include Arkolakis (2010), Bloom et. al. (2010) and Gourio and Rudanko (2011). Arkolakis (2010) develops a model of customer capital through advertisement –which is one of our elements within intangible expenditures. Bloom et. al (2010) measures the effects of management expenditures (also within our definition of intangibles) on Indian firms and Gourio and Rudanko (2011) develop a search model to analyze how intangible expenses affect firm dynamics.

The paper is organized as follows: Section 2 presents the empirical facts regarding the risk distribution across firms and over the business cycle, using Compustat and Kauffman Foundation data. Section 3 presents a firm dynamics model with endogenous expansion and contraction of firms, to capture the evidence presented in Section 2. Section 4 calibrates the model to the distribution of firms in the US. Section 5 discusses the effects of total factor productivity changes. Section 6 simulates the model and compares it to the data presented in Section 2. Finally, Section 7 concludes.

2 Idiosyncratic Risk Facts

In this section, we first present evidence on the relationship between idiosyncratic risk and intangible expenses. Second, we show evidence on the relationship between expenses, idiosyncratic risk, and business cycles.

2.1 Determinants of Idiosyncratic Risk

We are interested in understanding the determinants of cross-sectoral variation in firm-level idiosyncratic risk and its relationship with intangibles expenses such as advertising, marketing, brand development, organizational development, worker training, and other expenses that do not correspond to standard factor input payments such as capital or labor. As in Castro, Clementi, and Lee (2010) our proxy for risk is the volatility of the portion of growth in sales or TFP which is not explained by either industry effects, time effects, or firm characteristics associated with growth such as the firm's age or size (measured by labor).

We use two sources: the Kauffman Firm Survey (KFS) and Compustat. The KFS is a large panel of "young" businesses. Firms in the sample were founded in 2004 and have been tracked annually.[3] This panel was created using a random sample from Dun and Bradstreet's database of new businesses. The target population consisted of new businesses that were started in 2004 in the United States, and excludes any branch or subsidiary that was owned by an existing business or was inherited from someone else. The sample for the first survey consisted of 4,928 businesses.

Compustat provides annual accounting data for publicly listed U.S. firms. We use data from 1960 to 2010, consisting of an unbalanced panel of more than 8400 firms with a total of 232,193 firm-year observations. Compustat data is subject to selection bias as described by Davis, Haltiwanger, Jarmin, and Miranda (2006). Because these firms are relatively larger and older than those which are not in Compustat, they are likely to be less volatile (see Castro, Clementi and Lee (2010)). Moreover, we try to address these differences by controlling for age and size.[4]

[3]Data is currently available for the years through 2008. Firms will continue to be tracked through 2011. See http://www.kauffman.org/kfs/ for a detailed description of the data and for the public-use microdata itself.

[4]Appendix A1 provides a detailed description of our sample and of the construction of the key variables in both data sets.

We consider these two data sources for the following reasons. Both data sets, the KFS and Compustat, have unique information on firm-level expenses that do not correspond to factor input payments, such as expenses on marketing, advertising, and research and development. Moreover, the KFS provides us with a unique opportunity to study a panel of new businesses from startup, with available data on the revenues, workers, products, services, and innovations that these businesses possess and develop in their early years. However, the short duration of the panel does not allow us to completely disentangle age effects or simple learning-by-doing effects. Compustat, on the other hand, is a sample of relatively large and mature firms for which these effects are not as strong, and the access to an unbalanced panel allows us to control for age effects. Because our findings are robust to the different data sets, we are confident in the existence of the relationship between firm-level volatility and intangible expenses.

Before presenting our estimates of firm-level risk, Table 1 presents summary statistics for key variables in both samples.

Table 1: Distribution Sales, Int. Expenses and Capital (%)

	KFS Sample		
(in thousands)	Sales	Int. Expenses	Capital
$ < 3	6.80	37.87	16.94
$ 3 − 10	8.71	28.02	16.59
$ 10 − 50	8.81	13.83	19.02
$ 50 − 100	22.27	15.21	17.91
$ > 100	53.40	5.07	29.54
# Firms	1,940	1,381	1,977
	Compustat Sample		
(in millions)	Sales	Int. Expenses	Capital
$ < 10	12.73	21.70	20.42
$ 10 − 20	5.42	12.74	8.75
$ 20 − 50	10.12	16.85	11.84
$ 50 − 100	10.20	13.14	9.26
$ 100 − 250	14.08	13.65	12.31
$ > 250	47.45	21.92	37.41
# Firms	5,186	4,741	5,084

Note: Data is for 2008. All variables are in real terms. Sales and Int. Expenses are deflated using the BEAs 2-digit sector-specific price deflator for value added. Capital is deflated using the price deflator for investment, following Hall (1990).

Table 1 shows a clear difference in size between the firms in the KFS sample and those in the Compustat sample. Whereas approximately half of the firms in the KFS sample have real sales of less than $100K, about 85% of Compustat firms have sales of more than $10 million.

We now turn to the analysis of firm-level idiosyncratic dispersion. Our proxy idiosyncratic risk is derived from the portion of the variation in firms' sales growth that is not accounted for by aggregate disturbances or by other factors that vary systematically with growth, such as a firm's size or age.[5] The first step towards obtaining our measure of idiosyncratic volatility

[5]We are able to explicitly control for age in our Compustat sample; however, because all firms in the KFS are of the same age (all firms are born in 2004), this effect is already factored in.

is to estimate the following equation:

$$\Delta \ln(sales_{ijt}) = \mu_i + \delta_{jt} + \beta_{1j} \ln(size)_{ijt} + \beta_{2j} \ln(age)_{ijt} + \epsilon_{ijt+1}, \tag{1}$$

where $\Delta \ln(sales)_{ijt}$ is the growth of real sales for firm i, in industry j, in period t.[6] The variable μ_i is a firm fixed effect that accounts for unobserved persistent heterogeneity at the firm level (such as higher productivity, higher human capital of the entrepreneur, etc). The variable δ_{jt} denotes a full set of time- and industry-specific fixed effects.[7] We allow for industry-specific size effects. The estimation of equation (1) is done using the fixed effects panel estimator with robust standard errors. In the KFS sample, we use revenues from sales of goods, services or intellectual properties as our measure of sales. In the Compustat sample, our measure of sales is item #12, *net sales*.[8] Size, as it is standard in the literature, is defined in both samples as the number of employees. Age corresponds to the time since a firm first appeared in the sample.

Once equation (1) is estimated, we can compute the error, or the pure idiosyncratic and unpredictable component of firms' sales growth, $\hat{\epsilon}_{ijt+1}$. Given ϵ_{ijt+1}, we can study how its variance at the firm level, ϵ^2_{ijt+1}, is related to expenses in marketing, advertising, brand development, or organization development once industry-specific factors are accounted for. In particular, we estimate the following log-linear equation:

$$\ln(\epsilon^2_{ijt+1}) = \gamma_i + \theta_{tj} + \alpha_1 \ln(expense_{ijt}) + \alpha_2 t + u_{ijt}, \tag{2}$$

[6]In the appendix, we present an alternative based on TFPR (total factor productivity revenue based) with similar results. However, measurement issues associated with physical capital and factor shares determines that our preferred firm level volatility criterion is based on sales growth.

[7]We use two digits NAICS codes for firms in our KFS and Compustat sample.

[8]The sample selection and the definition of all variables used in the analysis are described in detail in Appendix A1. Nominal variables are deflated using the BEA's 2-digit sectoral-specific price deflator for value added.

where α_i is a firm fixed effect, θ_{tj} is an industry- and year-specific component, $\ln(expense_{ijt})$ is our measure of intangible expenses for firm i in sector j at time t, t is a time trend.[9] Expenses is defined as expenses that do not correspond to inputs of production. In the KFS sample, it is constructed as real total intangible expenses that include expenses in, for example, new product design, brand development, advertising, marketing, organizational development, and management consulting. In the Compustat sample, expenses is constructed as selling, general and administrative expenses (SGA).[10] These include expenses in advertising, marketing, and engineering.

Table 2 presents the estimates from equation (2) for our two samples.

Table 2: Firm level Idiosyncratic Volatility from Sales Growth

	KFS Sample	Compustat Sample
	Dependent Variable $\ln(\epsilon^2_{ijt+1})$	
$\ln(expense_{ijt})$	-0.085	-0.301
Std Error	0.036**	0.009***
Firm Fixed Effect	Yes	Yes
Industry-Year Controls	Yes	Yes
Time trend	Yes	Yes
N obs	2547	177178
R-squared	0.16	0.0435

Note: *** denotes significant at the 1% level, ** at the 5% level and * at the 10% level.

Table 2 shows that, in both samples, there is a significant and negative relationship between firm-level idiosyncratic volatility and intangible expenses (even after controlling for industry-time fixed effects and a time trend). The elasticity of volatility with respect to

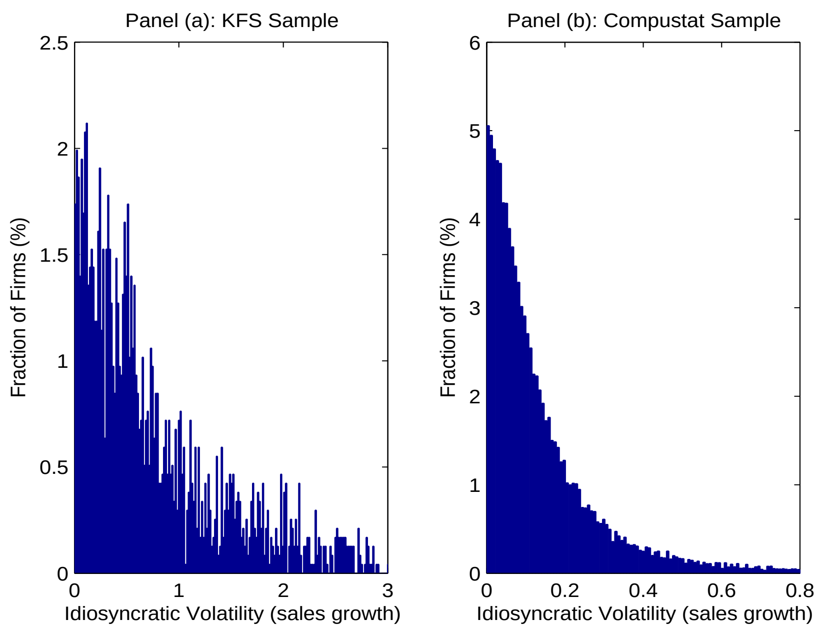
[9]We will show below that, consistent with the evidence presented in Comin and Phillipon (2005), a time trend is necessary because the variance of idiosyncratic risk for public firms has been rising during the last 30 years.

[10]Our results are robust to different definitions of intangible expenses. We experiment with two alternative specifications. In the first, we use Selling, General, and Administrative expenses (SGA) minus Research and Development Expenditures (XLR). In the second, it is defined as Advertising Expenses. Moreover, the results are also robust to incorporating age as an additional factor to account for learning effects.

expenses lies between -0.08 to -0.30.

Figure 1 presents the estimated distribution of idiosyncratic dispersion for both samples, based on the sales growth equation.

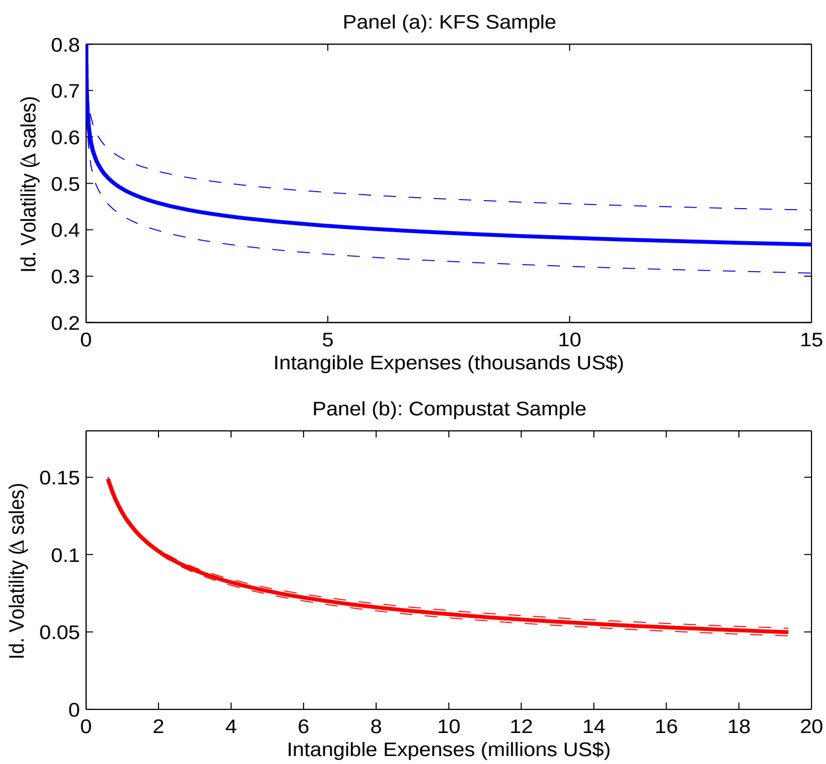
Figure 1: Idiosyncratic Dispersion



Note: Data from KFS and Compustat. Idiosyncratic Dispersion based on sales growth.

We note that firms in the KFS sample show considerably more volatility that those in the Compustat sample. This is consistent with the evidence presented in Haltiwanger et al (2010). The median dispersion in the KFS is more than 5 times the median dispersion in the Compustat sample. The estimated dispersion for the Compustat sample is consistent with the estimates in Castro et. al (2011). Finally, in Figure 2 we show the estimated relationship between intangible expenses and volatility for a representative firm in each of the samples.

Figure 2: Idiosyncratic Dispersion and Int. Expenses



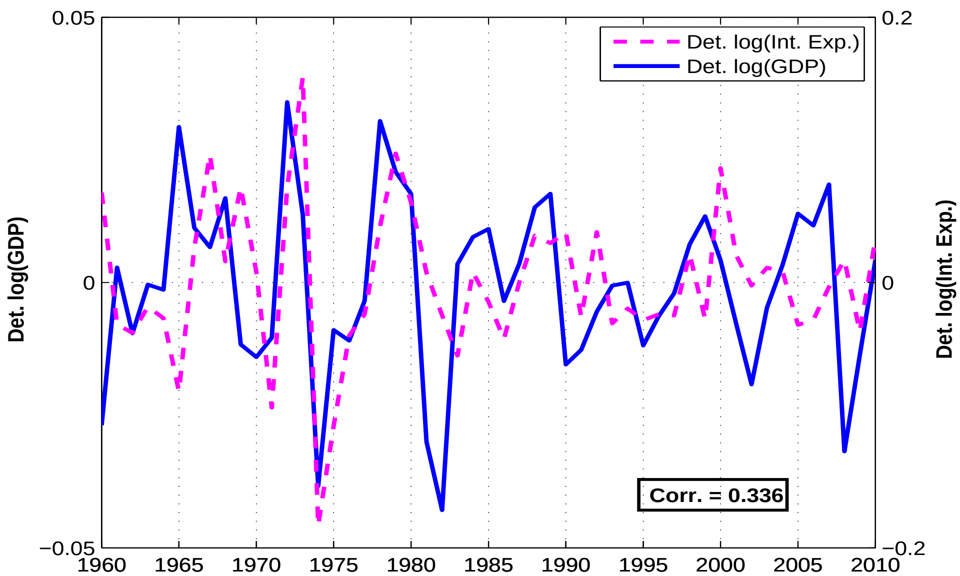
Note: Data from KFS and Compustat. Idiosyncratic Dispersion based on sales growth. Solid line corresponds to the estimated elasticity of idiosyncratic dispersion with intangible expenses. Dotted lines are the 95% confidence intervals.

2.2 Expenses, Volatility and Business Cycles

We are interested in understanding the endogenous determination of idiosyncratic dispersion over the business cycle. We start by presenting evidence on the relationship between intangible expenses and the business cycle. We use our Compustat sample because the KFS only extends for 5 years. Figure 3 shows the evolution of detrended real log-GDP and median real intangible expenses.[11]

[11]Real GDP is from FRED economic data at the St. Louis Federal Reserve. Both series are detrended using the H-P filter with a parameter of 6.25. GDP data is available since 1947.

Figure 3: Int. Expenses and Business Cycles

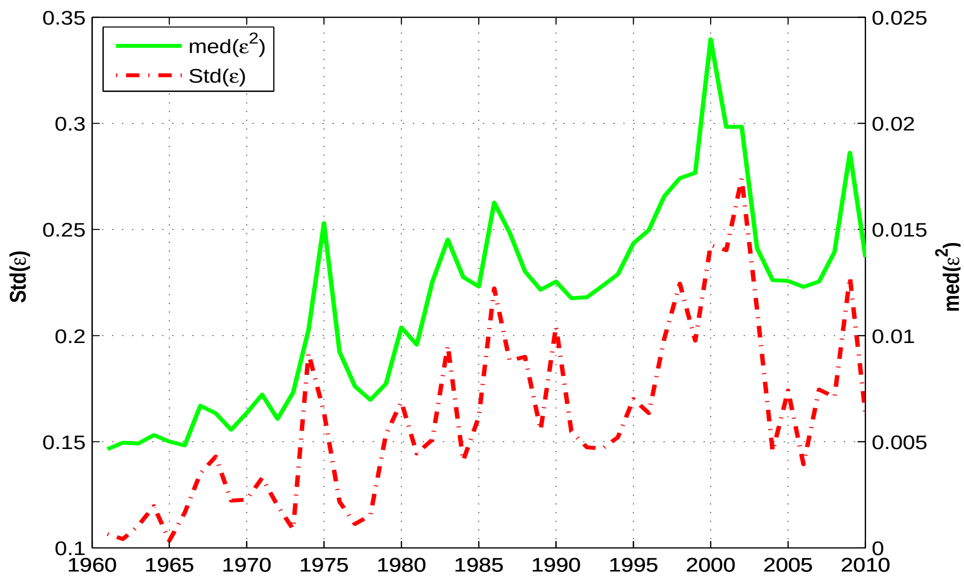


Note: Compustat Sample. Real intangible expenses corresponds to the median of the observed distribution in any given year. Real log GDP from FRED Economic Data (St. Louis Federal Reserve Bank). Both series are detrended using H-P filter with parameter 6.25. GDP data is available since 1947.

This figure shows that real GDP and intangible expenses are highly correlated. The correlation is 0.336 (significant at the 5% level).

Consistent with the estimates in Comin and Philipppon (2005), we find that idiosyncratic risk for publicly traded firms has been increasing over the last 5 decades. Figure 4 shows the evolution of the cross sectional deviation of ϵ_{ijt+1} (weighted by sales) and of median $\ln(\epsilon^2_{ijt+1})$.

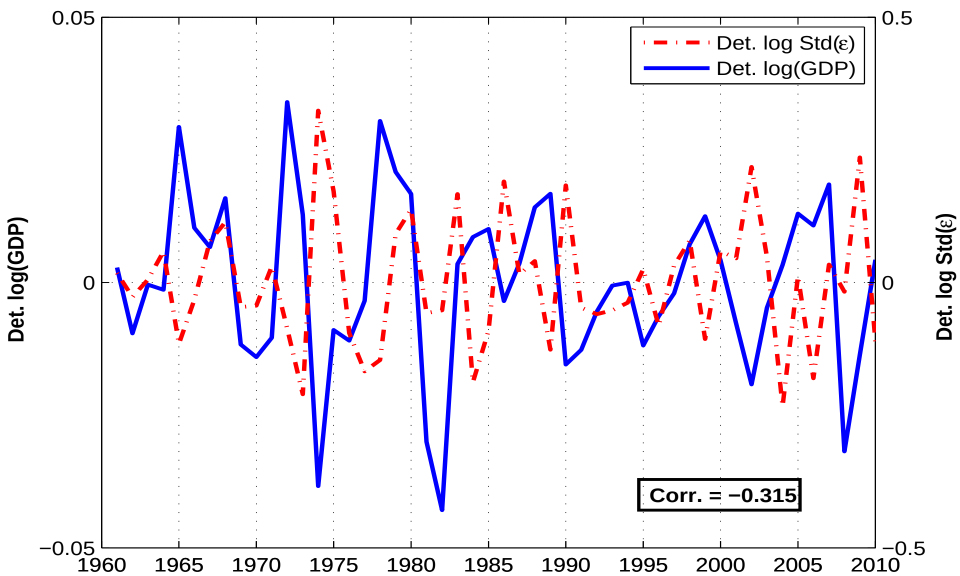
Figure 4: Evolution of Idiosyncratic Risk



Note: This figure shows the cross sectional standard deviation of ϵ_{ijt+1} (weighted by sales) and median $\ln(\epsilon^2_{ijt+1})$ where ϵ_{ijt+1} is the unexplained portion of sales growth.

We are interested in the cyclical properties of idiosyncratic risk. Because, as we showed in Figure 3, intangible expenses are highly pro-cyclical and, as we showed in the previous section, there is a significant negative relationship between intangible expenses and idiosyncratic risk (measured by the unexplained component of sales growth), it is not surprising that our estimated measure of dispersion is countercyclical. Figure 5 shows the relationship between real detrended log-GDP and detrended log weighted cross sectional standard deviation of ϵ_{ijt+1} (weighted by firm's sales).

Figure 5: Idiosyncratic Dispersion and Business Cycles



Note: $\ln(\epsilon^2_{ijt+1})$ estimated using data from Compustat and corresponds to the median of the estimated distribution. Real GDP from Fred Economic Data (St. Louis Federal Reserve Bank).

The correlation between real GDP and our estimated measure of idiosyncratic risk equals -0.315 and it is significant at the 5% level (the 5% confidence interval is $[-0.546, -0.041]$).[12] In what follows, we will explore the relationship between volatility and expenses in more detail through the lenses of our model.

3 Environment

We study an economy with N markets (where N is large but finite), a representative consumer and a continuum of competitive firms. Time is discrete. Firms can service each of

[12]As in most papers in the literature, we present the figure based on weighted Std. Deviation of ϵ_{ijt+1}. Results are similar if we consider median ϵ^2_{ijt+1}. In particular, the correlation with detrended GDP is -0.384 and significant at the 1% level.

the different markets by incurring sales and marketing expenses. We interpret the notion of a firm to be a brand, and a market is a geographic location.

3.1 Households Preferences and Endowments

The representative household derives utility from the consumption of the composite good C_t. More specifically, their preferences are given by

$$U(C_t) \tag{3}$$

where C_t is a composite of the consumption good at each location n:

$$C_t = \left[\sum_{n=1}^{N} \left(\epsilon_{n,t} c_{n,t}\right)^{\rho}\right]^{1/\rho}, \; 1 > \rho > 0, \tag{4}$$

where $c_{n,t}$ refers to consumption in location n, $\epsilon_{n,t}$ is a taste shock associated with location n in period t, and $1/(1-\rho) > 1$ is the elasticity of substitution among different locations. It is assumed that $\log(\epsilon_{n,t}^{\frac{\rho}{1-\alpha\rho}}) \sim N(0, \sigma^2)$

The household is endowed with one unit of labor that it supplies inelastically every period at wage w_t. They receive dividends D_t through ownership of firms in the economy.

We define the price index as follows

$$P_t = \left[\sum_{n=1}^{N} \left(\frac{p_{n,t}}{\epsilon_{n,t}}\right)^{\frac{\rho}{\rho-1}}\right]^{\frac{\rho-1}{\rho}}. \tag{5}$$

Thus, the budget constraint that consumers face is

$$P_t C_t \leq w_t + D_t. \tag{6}$$

3.2 Firms and Technology

Firms are ex-ante identical, and are described by their productivity parameter s. Production requires only one factor, labor. The production technology for the firms that have productivity s, supply to market n, with aggregate shock z_t is given by

$$q_s(t,n) = z_t s \ell_s^{\alpha}(t,n), \tag{7}$$

where $\ell_s(t,n)$ is labor employed in the production of goods from a firm with productivity (s). We assume that firm-level productivity takes values on the set $S = \{\underline{s}, \ldots, \overline{s}\}$, is drawn from distribution $G(s)$, and is constant over the lifespan of the firm. The aggregate technology shock z_t follows a Markov process with the transition function $\Gamma(z', z)$.

Firms can reach and sell to consumers in location/market n by incurring sales, marketing, and other intangible expenses. We assume that these expenses are measured in units of labor and are increasing in the number of markets that the firms serve.[13] The total cost paid by a firm that serves m markets is $w\Phi(m) = w_t \frac{\psi}{z_t}(m-1)^{1+\nu}$. As we will show below, firms have incentives to access other locations in order to increase its customer base and diversify location-specific risk. The assumption that marketing and sales expenses are increasing in the number of markets that the firm serves reflects that the complexity in management is tied to some resource that is in fixed supply.

Firms maximize the expected sum of discounted dividends, acting as price takers in each location where they participate. They discount the future at rate $\beta \in (0,1)$.

3.3 Timing

The timing within a period is as follows:

[13] An interpretation of this is geographical distance or differences in products. Another interpretation is an increasing cost that arises from the complexity of serving many markets.

1. z_t is realized.

2. Firms choose the number of locations in which to operate.

3. Taste shocks $\epsilon_{n,t}$ are realized.

4. Taking prices as given, firms choose labor and produce.

5. Households consume.

3.4 Equilibrium

In this section, we present a Markov Perfect Competitive Equilibrium. We use the standard recursive notation.

3.4.1 Consumer's Problem

Households' optimal demand for consumption good in location n in period t is:

$$c_{n,t} = \epsilon_{n,t}^{\frac{\rho}{1-\rho}} \left(\frac{p_{n,t}}{P_t^{\rho}} \right)^{\frac{1}{\rho-1}} [w_t + D_t] \tag{8}$$

3.4.2 Firm's Problem

Firms are perfect competitors in each market in which they participate. After the shocks z_t and $\epsilon_{n,t}$ are revealed, the firm optimizes over the amount of labor to demand in each market they have perviously chosen to serve.

The profit function for a firm in market n is given by:

$$\pi_s(t,n) = \max_{l_s(t,n)} \{ p_{n,t} q_s(t,n) - w_t l_s(t,n) \} \tag{9}$$

s.t.

$$q_s(t,n) = z_t s \ell_s^{\alpha}(t,n), \tag{10}$$

Which delivers a standard labor demand in market n,

$$l_s(t,n) = \Big(\frac{w_t}{p_{n,t} s z_t \alpha}\Big)^{\frac{1}{\alpha-1}} \tag{11}$$

This implies that profits for a firm with productivity s, in market n, are the following

$$\pi_s(t,n) = (p_{n,t} s z_t)^{\frac{1}{1-\alpha}} w_t^{\frac{\alpha}{\alpha-1}} (\alpha^{\frac{\alpha}{1-\alpha}} - \alpha^{\frac{1}{1-\alpha}}) \tag{12}$$

3.4.3 n^{th} Market equilibrium

The price $p_{n,t}$ is such that it clears the n^{th} market. In other words, it is such that

$$\sum_{s=\underline{s}}^{\overline{s}} \lambda_s(n,t) q_s(n,t) = c_{n,t} \tag{13}$$

where $\lambda_s(n,t)$ refers to the mass of firms with productivity s, participating in market n, in period t. Note that $\lambda_s(n,t) = \frac{\mu_s(t) m_s^*(t)}{N}$, where $\mu_s(n,t)$ refers to the distribution of firms across productivity and $m_s^*(t)$ refers to the number of markets in which each firm, of productivity s, chooses to participate. The equilibrium price in market n is then

$$p_{n,t} = \epsilon_{n,t}^{\frac{\rho(1-\alpha)}{1-\alpha\rho}} A(t) \tag{14}$$

where

$$A(t) = \Big[P_t^{\rho} \Big(\frac{(w_t + D_t) w_t^{\frac{\alpha}{1-\alpha}}}{\widetilde{s}_t z_t^{\frac{1}{1-\alpha}} \alpha^{\frac{\alpha}{1-\alpha}}}\Big)^{1-\rho}\Big]^{\frac{1-\alpha}{1-\alpha\rho}} \tag{15}$$

and

$$\widetilde{s}_t = \frac{1}{N}\sum_{s=\underline{s}}^{\overline{s}} \mu_s(t) m_s^*(t) s^{\frac{1}{1-\alpha}} \tag{16}$$

Firms enter the n^{th} market as long as

$$E(\pi_s(n,t)) \geq w_t\Big(\Phi(n) - \Phi(n-1)\Big). \tag{17}$$

Using the equation of the market clearing price, it becomes

$$s^{\frac{1}{1-\alpha}} B(t) \geq w_t\Big(\Phi(n) - \Phi(n-1)\Big) \tag{18}$$

where

$$B(t) = e^{\frac{\sigma^2}{2}} z_t^{\frac{1}{1-\alpha}} w_t^{\frac{\alpha}{\alpha-1}}\Big(\alpha^{\frac{\alpha}{1-\alpha}} - \alpha^{\frac{1}{1-\alpha}}\Big) A(t)^{\frac{1}{1-\alpha}}. \tag{19}$$

3.5 Aggregates

Aggregate dividends are defined by the following equation

$$D_t = \Pi_t - w\sum_{s=\underline{s}}^{\overline{s}} \mu_s(t)\Phi(m_s^*(t)), \tag{20}$$

where Π_t denotes the sum of profits across markets and is given by

$$\Pi_t = \sum_{s=\underline{s}}^{\overline{s}}\sum_{n=1}^{N} \lambda_s(n,t)\pi_s(n,t). \tag{21}$$

The equilibrium condition in the labor market is

$$\sum_{s=\underline{s}}^{\overline{s}}\sum_{n=1}^{N} \lambda_s(n,t) l_s(n,t) + \sum_{s=\underline{s}}^{\overline{s}} \mu_s(t)\Phi(m_s^*(t)) = 1. \tag{22}$$

This implies

$$w_t + D_t = \frac{w_t}{\alpha}\left[1 - \sum_{s=\underline{s}}^{\overline{s}} \mu_s(t)\Phi(m_s^*(t))\right] \tag{23}$$

$$\Rightarrow \frac{\Pi_t}{w_t} = \left[1 - \sum_{s=\underline{s}}^{\overline{s}} \mu_s(t)\Phi(m_s^*(t))\right]\Big(\frac{1}{\alpha} - 1\Big). \tag{24}$$

The price index P_t, becomes in equilibrium

$$P_t = \left[\frac{1 - \sum_{s=\underline{s}}^{\overline{s}} \mu_s(t)\Phi(m_s^*(t))}{\widetilde{s}_t}\right]^{1-\alpha} \frac{w_t}{\alpha z_t}\left(Ne^{\frac{\sigma^2}{2}}\right)^{\frac{\alpha\rho-1}{\rho}}. \tag{25}$$

The system is now one of three equations and three unknowns $(\sum_{s=\underline{s}}^{\overline{s}} \mu_s(t)\Phi(m_s^*(t)), w_t, \widetilde{s}_t)$.

3.6 Firm Level Volatility

Conditional on the aggregate shock z_t, the model predicts a relationship between the firm's idiosyncratic productivity s and its volatility. The coefficient of variation of the shock to which the firm is exposed, conditional on serving m markets, is

$$CV_s(t) = \frac{\sqrt{Var(\sum_{i=1}^{m} sz_t p_{it})}}{E(\sum_{i=1}^{m} sz_t p_{it})}. \tag{26}$$

From (18), the firm will participate in an increasing number of markets as a function of its productivity s. Therefore, the coefficient of variation is a function of the firm's productivity, through its effect on the optimal number of markets that the firm will serve. Then, the coefficient of variation for a firm with productivity s can be written as

$$CV_s(t) = \sqrt{\frac{e^{(1-\alpha)^2\sigma^2} - 1}{m_s^*(t)}}. \tag{27}$$

This result is based on the assumption that the shocks $\epsilon_{n,t}$ are iid. However, as long as the shocks are not perfectly correlated (which would make them, in fact, one unique shock), the coefficient of variation decreases as the firm is exposed to an increasing number of shocks. This can be seen by analyzing the variance covariance matrix of the shocks $\epsilon_{n,t}$. Given that they have the same variance, the variance covariance matrix can be rewritten in terms of the correlation coefficient between two shocks multiplied by the common variance term. The coefficient of variation is then given by the following expression for the case of a firm that serves $m_s^*(t)$ markets

$$CV(s) = \frac{\sqrt{(e^{(1-\alpha)^2\sigma^2} - 1)\sum_{u=1}^{n}\sum_{v=1}^{n}\rho_{uv}}}{m_s^*(t)}, \tag{28}$$

where ρ_{uv} is the correlation coefficient between the shocks u and v. In the case of iid shocks, the double sum equals the number of shocks and in the case of perfectly correlated shocks it equals the square of the number of shocks. Anything in between means that the coefficient of variation drops as the number of varieties increases.

4 Calibration

The parameters to be calibrated are the following: $\rho = .83$ to deliver a markup of 20%; $\sigma = 3.04$ as calibrated from the KFS under the assumption that these firms are exposed to only one market and using their coefficient of variation for the firm level idiosyncratic productivity; β from the risk-free bond; and the parameters for the distribution of productivity across firms and the market expansion function. Assuming that productivity is distributed following a lognormal distribution across firm, this leaves us with four parameters: μ_s, σ_s, ν, and ψ, where μ_s and σ_s refer to the mean and standard deviation of the lognormal distribution of firm-level productivity. The four parameters are jointly calibrated to match the firm size distribution in terms of the number of employees.

Table 3 describes the main parameters of the model.

Table 3: Model Parameters

Parameter	Value
ρ	0.83
α	0.64
σ	3.04
ν	0.56
ψ	0.46
N	100
μ_s	ln(1.7)
σ_s	0.4

The match between the distribution of employees across firms is shown in Table 4.

Table 4: Firms size distribution - Number of employees

Employment size	Data	Model
Firms with 1 to 4 employees	0.610	0.601
Firms with 5 to 9 employees	0.176	0.209
Firms with 10 to 19 employees	0.107	0.100
Firms with 20 to 99 employees	0.089	0.064
Firms with 100 to 499 employees	0.015	0.021
Firms with 500 employees or more	0.003	0.005

The data is obtained from the Statistics of U.S Businesses from the U.S. Census Bureau for 2008. It covers all employer firms (around 6 million) that ,in total, employ around 121 million workers. Given that in the model we have a measure of firms equal to one, the model distribution reported was adjusted for the difference in mean. Also note that for 2008 we observe 5186 firms in our Compustat sample, less than 0.1% of the total number of firms.

5 Effect of TFP changes

In this section an aggregate TFP shock is introduced. In particular, z_t is lowered from 1 to 0.9 and its effect on the distribution of firms and on volatility is analyzed. Figure 6 shows the effect of the aggregate shock on the endogenous number of markets that each firm serves. Note that as z increases, there is a additional effect on measured TFP coming from the fact that the more productive firms are the ones expanding more rapidly. This endogenous amplifying effect on TFP is non-negligible, and amounts to an additional 13% increase in measured TFP. Changes in TFP also generate an endogenous change in firm-level volatility. The average firm-level volatility decreases by 0.023%, and the average firm-level volatility for the top 10% and 1% of firms falls by 1.4% and 3.9%, respectively. This uneven change in volatility comes as a direct consequence of the asymmetric impact that aggregate TFP has on firm-level decisions. The more productive firms are the ones that engage in expansions in response to the aggregate shock, whereas the less productive firms do not change their decision by much.

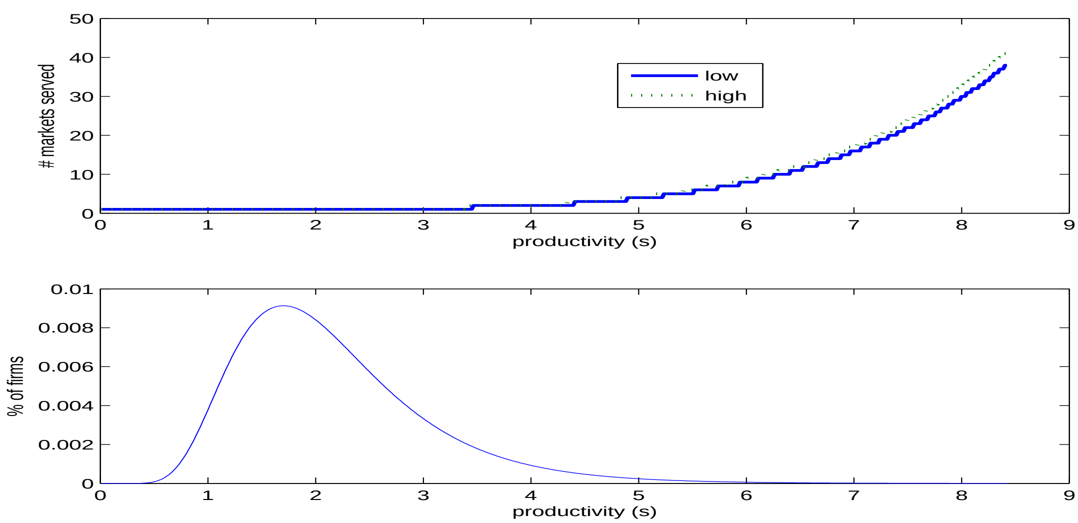


Figure 6: Markets served by productivity level

The labor force that is employed in the production of intangibles accounts for 17.8% of the total labor force during boom years, and is reduced by of 2.65% in the low-TFP years.

6 Simulation of the model

Once the model is calibrated, we perform a simulation of the model in order to see how it compares to the empirical results. In order to perform the simulation, we draw 100,000 firms from the model following the productivity distribution and simulate the model 20 times for 50 periods in each simulation, with an of 80% probability of booms and a 20% probability of recessions (the TFP parameter z_t is set to 1 in booms and 0.9 in recessions).

With the information from the firms in the model, we perform the same exercise as in the empirical section of the paper. That is, we compute the log difference in sales from period t to $t+1$ and regress it against a firm's fixed effect, size (in terms of number of employees), and a time dummy capturing aggregate conditions (booms or recessions). We obtain the errors from that first equation, and then regress the log of the errors squared against a firm fixed effect and the log of the intangible expenses in period t.

The first equation is the following:

$$\Delta \ln(sales_{it}) = \mu s_i + \delta z_t + \beta \ln(size)_{it} + \epsilon_{it+1}, \tag{29}$$

and the second is:

$$\ln(\epsilon^2_{it+1}) = \gamma s_i + \alpha \ln(expense_{it}) + u_{ijt}. \tag{30}$$

The parameter α is equal to -0.1807, significant at a 95% level. This value is right in the middle of the estimates from the KFS and Compustat, reported in Table 2. Note that in the model we sample from the whole distribution of firms and therefore this result should be expected, given that in the KFS we observe only young entrepreneurial firms and in Compustat we observe firms that are public and relatively large.

The cyclical characteristics generated by the model are very much in line with the evidence. The correlation between median intangible expenses for the top 5% of firms (when sorted by productivity) and GDP is 1. Also, the correlation between the sales' weighted cross sectional standard deviation of idiosyncratic risk and GDP equals -0.0605 (significant at the 5% level) for the top 10% of firms (when sorted by productivity).[14] The closest comparison with our data is the right tail of the firm's distribution, given that our data for the cyclical properties of risk and intangible expenditures comes from Compustat (i.e. only public firms). In the data, the correlation between expenditures and GDP is 0.336 and between the sales' weighted cross sectional standard deviation of risk and GDP is -0.315 with a 95% interval equal to $[-0.546, -0.041]$, so the model value of -0.06 is well within the data range.

7 Conclusion

The determinants of firm-level risk are documented in the data. Looking at both ends of the firm distribution, a channel from intangible expenses to volatility is uncovered. In terms of cyclical properties, risk and intangible investment are countercyclical and pro-cyclical, respectively.

These findings are then explained through the use of a model based on location-specific demand shocks and intangible expenditures as a means of expanding the firm to a larger number of markets. The result is that high-productivity firms expand to a larger set of markets, making them less volatile than their low-productivity counterparts. Also, low-productivity firms do not react to the cycle, whereas larger firms do, explaining the cyclicality in risk at the upper end of the firm size distribution.

[14] Having fewer firms (conditioning at the top 5% or 1%) gives us small sample issues. By looking at the top 10% of firms we have 10,000 firms per period, which is the same order or magnitude as Compustat for the late part of the sample.

8 References

Arellano, Cristina, Bai, Yan and Kehoe Patrick (2011) “Financial Markets and Fluctuations in Uncertainty”, Federal Reserve Bank of Minneapolis, Research Department Staff Report.

Arkolakis, Costas (2010) “Market Penetration Costs and the New Consumers Margin in International Trade”, Journal of Political Economy, 118 (6), 1151-1199.

Bachman, Rudiger and Moscarini, Giuseppe (2011) “Business Cycles and Endogenous Uncertainty” University of Michigan mimeo

Bernard, Andrew, Redding, Stephen and Schott, Peter (2010) “Multiple-Product Firms and Product Switching”, American Economic Review 100 (1): 70-97.

Bloom, Nicholas (2009) “The Impact of Uncertainty Shocks”, Econometrica 77 (3): 623-685.

Bloom, Nicholas, Eifert, Benn, Mahajan, Aprajit, McKenzie, David and Roberts, John (2010) “Does Management Matter? Evidence from India” Stanford University mimeo.

Castro, Rui, Gian Luca Clementi, and MacDonald (2009) “Legal Institutions, Sectoral Heterogeneity, and Economic Development”, Review of Economic Studies, 76, 529-561.

Castro, Rui, Clementi, Gian Luca, and Lee, Yoonsoo (2010) “Cross-Sectoral Variation in Firm-Level Idiosyncratic Risk”, New York University mimeo.

Christiano, Lawrence, Motto, Roberto and Rostagno, Massimo (2011) “Risk Shocks”, Northwestern University mimeo.

Comin, Diego and Phillipon, Thomas (2005) “The Rise in Firm-Level Volatility: Causes and Consequences” NBER Macroeconomics Annual 2005.

Davis, Stephen, Haltiwanger, John, Jarmin, Ron, and Miranda, Javier (2006) “Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms” NBER Macroeconomics Annual 2006.

Gourio, Francois and Leena, Rudanko (2011) “Customer Capital,” Boston University mimeo.

Hall, Brownyn (1990) “The Manufacturing Sector Master File: 1959-1987” NBER Working Paper 3366.

Haltiwanger, John, Jarmin, Ron and Miranda, Javier (2010) “Who Creates Jobs? Small vs. Large vs. Young” NBER working paper 16300.

Imrohoroglu, Ayse and Tuzel, Selale (2011) “Firm Level Productivity, Risk and Return”, University of Southern California mimeo.

Kauffman Firm Survey (2010) Data downloaded from kaufmann.org in December 2010.

Kehrig, Matthias (2011) “The Cyclicality of Productivity Dispersion” Working Papers 11-15, Center for Economic Studies, U.S. Census Bureau.

Olley, Steve and Pakes Ariel (1996) “The Dynamics of Productivity in the Telecommunications Equipment Industry” Econometrica 64 (6), 1263-98

A1 Appendix

A1.1 Firm Level Volatility based on TFPR

In this section, we present a measure of firm-level idiosyncratic risk derived from the portion of growth in firms' total factor productivity (TFPR) that is not accounted for by industry, firm, or time effects or the firm's size and age. The first step towards obtaining the idiosyncratic risk component is to find the firm-level total factor productivity. Using our data, we estimate the following equation

$$z_{ijt} = y_{ijt} - \alpha^k k_{ijt} - \alpha^n n_{ijt}, \tag{A.1.1}$$

where z_{ijt} is log firm level productivity, y_{ijt} denotes log-real output (or value added), k_{ijt} is log-real capital, and n_{ijt} corresponds to log total workers for firm i, in industry j, in period t. Besides correctly deflating value added and obtaining a measure of a firm's capital, an important aspect of deriving firm-level productivity from equation (A.1.1) is the value of factor shares. We set factor shares equal to $\alpha^k = 0.85 \times (1/3)$ and $\alpha^n = 0.85 \times (2/3)$, a standard value in the firm dynamics literature.[15] Because firms use many inputs in their production, such as raw materials, labor, and energy, we follow Imrohoroglu and Tuzel (2011) and focus on labor and physical capital as the main inputs and define value added as gross output net of expenditures in materials and other expensed items such as advertising, R&D expenditures, and rental expenses. In the KFS sample, we proxy the capital stock using total real equipment, land, and buildings, vehicles, and other production assets. In the Compustat sample, firm level capital is given by gross Plant, Property, and Equipment (PPEGT). In

[15] We are working on estimating factor shares using our own data. Because estimates of factor shares from equation (A.1.1) suffer from endogeneity bias, we use the method proposed by Olley and Pakes (1996) to derive total factor productivity. This is a three-step procedure that uses the information contained in investment to reduce the endogeneity problem (see Sections A1.2 and A1.3 below for details on the construction of variables and the estimation procedure).

both samples, we deflate the capital stock following Hall (1990).

Finally, we estimate the following equation:

$$\Delta z_{ijt} = \mu_i + \delta_{jt} + \beta_{1j}\ln(size_{ijt}) + \beta_{2j}\ln(age_{ijt}) + \epsilon_{ijt}, \quad \text{(A.1.2)}$$

where Δz_{ijt} is the growth of total factor productivity for firm i, in industry j, in period t. As with sales growth, equation (A.1.2) is estimated using the fixed effects panel estimator with robust standard errors.

After obtaining ϵ_{ijt} as the residual from equation (A.1.2), we estimate an equation relating firm-level volatility and intangible expenses (identical to (2)). Table 5 presents the results.

Table 5: Firm level Idiosyncratic Volatility from Productivity (TFPR) Growth

	KFS Sample	Compustat Sample
	Dependent Variable $\ln(\epsilon^2_{ijt+1})$	
$\ln(expense_{ijt})$	-0.153	-0.233
Std Error	0.047***	0.012***
Industry-Year Controls	Yes	Yes
Time trend	Yes	Yes
N obs	1769	146472
R-squared	0.007	0.0587

Note: *** denotes significant at the 1% level, ** at the 5% level and * at the 10% level.

Table 5 shows that in both samples there is a significant and negative relationship between firm-level idiosyncratic risk, when estimated from TFPR growth, and intangible expenses. This is found even after controlling for industry-time fixed effects and a time trend. The elasticity of volatility with respect to expenses lies between -0.15 to -0.23.

A1.2 Kauffman Firm Survey Sample

The Kauffman Firm Survey (KFS) is a large panel of "young" businesses. Firms in the sample were founded in 2004 and have been tracked annually.[16] This panel was created using a random sample from Dun and Bradstreet's database list of new businesses. The target population consisted of all new business that were started in the 2004 calendar year in the United States, and excludes any branch or subsidiary owned by an existing business or a business inherited from someone else. The sample for the first survey consisted of 4,928 businesses.

The KFS provides us with the unique opportunity to study a panel of new businesses from startup, using available data on their revenues and expenses, the number of workers, the products, services, and innovations that these business possess and develop in their early years of existence as well as the extent to which these business are involved in innovative activities. One drawback of the publicly available KFS data is that some variables such as assets (and its components) and sales are only reported within certain ranges.[17] We recognize that this creates some measurement problems and will try to address them in the future. As a first pass to the data, we set the value of the corresponding variables to the middle value of the reported range.[18]

Our unit of observation is the firm, as defined by the KFS. The change in sales is constructed from total revenues from sales of goods, services, or intellectual properties. Size, as is standard in the literature, is defined as the number of employees. Because firms use many inputs in production, such as raw materials, labor, and energy, we follow Imrohoroglu and Tuzel (2011) and focus on labor and physical capital as the main inputs and define value

[16] Current available data expands through year 2008. Firms will continue to be tracked through 2011. See http://www.kauffman.org/kfs/ for a detailed description of the data and the public-use microdata itself.

[17] For example, ranges for revenues are 0, \$1-1000, \$1001-5000, \$5001-10000, \$10001-25000, \$25001-100000, and \$100000 or more.

[18] The set of variables that we use which present this problem are: revenue from sales of goods, services, or intellectual properties, expenses, wages, and assets (and its components).

added as gross output net of expenditures in materials, as well as net of other expensed items such as advertising, R&D expenditures, and rental expenses. The capital stock is proxied by total equipment, land and buildings, vehicles, and other production assets. We use the two digits NAICS codes to control for industry effects. All variables are deflated using the GDP deflator except for capital, which is deflated using the price deflator for investment following Hall (1990). Expenses is defined as expenses that do not correspond to production inputs. It is constructed as total expenses in intangibles that include expenses in, for example, design of new products, brand development, advertising, marketing, organizational development, or management consulting. For firm/year observations with missing values of expenses in intangibles, we compute the average ratio of intangible expenses to total expenses and input intangible expenses from this ratio and total expenses. When computing idiosyncratic uncertainty from changes in TFP, the size of the KFS does not allow us to conduct the first stage using the procedure proposed by Olley and Pakes (1996). We use equation (A.1.1) and set factor shares to $\alpha^k = 0.85 \times \frac{1}{3}$ and $\alpha^n = 0.85 \times \frac{2}{3}$, which are standard values in the literature and are consistent with the values in our model. In our final regression we control for financial frictions using the leverage ratio, defined as total debt divided by total assets.

Table 6 presents the distribution of real sales, real intangible expenses, and real capital for newborn firms (i.e. the distribution of firms in 2004) and for firms that survive until the end of our sample (2008).

Table 6: Distribution of Sales, Int. Expenses and Assets (%)

	Year 2004			Year 2008		
	Sales	Int. Expenses	Capital	Sales	Int. Expenses	Capital
\$ 1 − 3,000	14.52	55.09	22.49	6.80	37.87	16.94
\$ 3,001 − 10,000	14.39	26.40	22.49	8.71	28.02	16.59
\$ 10,001 − 50,000	14.59	0.00	20.77	8.81	13.83	19.02
\$ 50,001 − 100,000	28.58	16.64	16.27	22.27	15.21	17.91
\$ > 100,000	27.92	1.87	17.97	53.40	5.07	29.54
# Firms	3,037	4,382	3,650	1,940	1,381	1,977

Note: Data from KFS. Sales and Exp. Intangibles are deflated using the GDP deflator. Capital is deflated using the price deflator for investment following Hall (1990).

We observe that many firms are relatively small, with sales, intangibles expenses, and capital below \$10,000. This is still the case even after four years of existence. However, a non-trivial number of new firms have sales, intangible expenses, and capital above \$100,000. The distributions clearly shift upwards as the cohort of firms becomes older, grows and selection takes place.

Table 7 presents the distribution of newly created firms as seen in the KFS, a comparison with the size distribution of new firms from Census data, and also the distribution of firms over employment for our cohort of firms in 2008.[19]

[19] For comparison, we report the distribution conditional on firms having more than one worker. In the KFS, we find that in 2004, 58% of active firms hire zero workers and this value equals 44% in 2008.

Table 7: Distribution of workers (%)

Number of Employees	KFS (2004)	Census (2004)	KFS (2008)
1−4	74.4	76.7	64.8
5−9	15.3	13.0	17.8
10−19	6.6	6.0	9.5
20−99	3.4	3.8	2.9
100−499	0.3	0.4	5.0
500 +	0.0	0.0	0.0

Note: KFS refers to Kauffman Firm Survey. Census corresponds to Office of Advocacy, Small Business Administration, Statistics of U.S. Business, U.S. Census 2004.

Table (7) shows that a large fraction of firms start up with only a few workers. More than 70% of new firms hire between one and four workers. As a comparison, we present the distribution of new firms from Census data and we note that the distributions are very similar. This reassures us that we have a representative sample of new firms, despite some differences in the distribution of new firms across industries and the different methodologies used across sources. Finally, and consistent with the evidence presented in Table (6), when looking at the distribution of active firms in the KFS in 2008, we observe that there is a sizable reduction in the fraction of firms with less than 4 workers, as well as an increase in the fraction of firms with more than 10 workers.

Table (8) displays the distribution of firms across some representative industries and their one year survival rates.

Table 8: Distribution of firms across industries, and survival rates

Industry	Fraction of Firms (%)	One Year Survival Rate (%)
Construction	10.0	91.9
Manufacturing	7.1	92.0
Wholesale	5.4	88.7
Retail	15.9	86.1
Transportation and Warehousing	3.4	84.7
Information	2.7	84.6
Finance and Insurance	4.7	95.8
Administration and Support	9.6	91.7
Accommodation and Food Services	4.3	77.7

Source: Kauffman Firm Survey.

A1.3 Compustat Sample

We use Compustat's fundamental annual data. The sample period ranges from 1960 to 2010. We exclude financial firms with standard industrial classification (SIC) codes between 6000 and 6999, utility firms with SIC codes between 4900 and 4999, and firms with SIC codes greater than 9000 (residual categories). Observations are deleted if they do not have a positive book value of assets, or if gross capital stock or sales are either zero, negative, or missing. The final sample is an unbalanced panel with more than 8400 firms, which has 232,193 firm/year observations.

Our data variables are defined as follows. The change in sales is constructed from the variable SALE. Size, as it is standard in the literature, is defined as the number of employees, using the variable EMP. The construction of variables and the estimation of productivity is derived from firm-level TFP and closely follows Imrohoroglu and Tuzel (2011). Because firms use many inputs in their production, such as raw materials, labor, and energy, we focus on labor and physical capital as the main inputs and define value added as gross output net of expenditures in materials, as well as net of other expensed items such as advertising, R&D expenditures, and rental expenses. Value added is computed as sales minus expenses in materials. Materials is measured as total expenses minor labor expenses. Total expenses is approximated by sales minus operating income before depreciation and amortization (OIBDP). Labor expenses is calculated using the variable XLR. For those units with missing XLR values, we use EMP multiplied by a measure of industry-level wages estimated from our sample. These steps lead to a definition of value added that is equal to OIBDP plus labor expenses. Capital is given by gross plant, property, and equipment (PPEGT). Investment is derived from the variable CPAXV. We use two digits NAICS codes to control for industry effects. Firm age is proxied by the number of years since the firm's first-year observation in Compustat.

All variables except for the capital stock are deflated using the BEAs 2-digit sector-specific price deflator for value added. The capital stock is deflated by the price deflator for investment following Hall (1990). Because investments were made at various times in the past, we need to calculate the average age of the capital in every year for each company. The average age of capital is calculated by dividing accumulated depreciation (Gross PPE - Net PPE) by current depreciation (DP). The resulting capital stock is lagged by one period to measure the available capital stock at the beginning of the period.

The size of our Compustat sample allows us to estimate firm-level TFP using the following two alternatives. First, as in our KFS sample, we use equation (A.1.1) and set factor shares to $\alpha^k = 0.85 \times \frac{1}{3}$ and $\alpha^n = 0.85 \times \frac{2}{3}$, which are standard values in the literature and are consistent with the values in our model. Second, we estimate firm-level TFP following Olley and Pakes (1996)[20]. Specifically, consistent with our model, log-output y_{it} for firm i in period t is given by

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_n n_{it} + s_{it} + e_{it}, \tag{A.1.3}$$

[20] Estimates from this procedure are not available at this moment.

where k_{it} is the capital stock, n_{it} is labor, s_{it} is firm level productivity, and e_{it} is a mean-zero error term unobservable to the econometrician. Firm-level productivity s_{it} is observed before investment and labor decisions are made, so this gives rise to an endogeneity problem that affects the estimates of β_k and β_n.

The firm's optimization problem results in the following investment function

$$i_{it} = i(k_{it}, s_{it}), \tag{A.1.4}$$

where $i(\cdot)$ is increasing in s_{it}. Assuming that this function is invertible and taking the inverse we obtain

$$s_{it} = h(k_{it}, i_{it}), \tag{A.1.5}$$

where $h(\cdot)$ is increasing in i_{it}. Let $\phi_{it} = \beta_0 + \beta_k k_{it} + h(k_{it}, i_{it})$. Then, from equation (A.1.3) we obtain

$$y_{it} = \beta_n n_{it} + \phi_{it} + e_{it}. \tag{A.1.6}$$

Using this equation, we can estimate $\hat{\beta}_n$ consistently. Following Imrohoroglu and Tuzel (2010), the estimation of equation (A.1.6) is performed using industry-specific time dummies, and the function ϕ_{it} is approximated with a second-order polynomial in capital and investment. To estimate β_k, we first iterate forward (A.1.3), subtract $\beta_n n_{it}$, and take expectations (as of period t) to obtain

$$\begin{aligned} E_t[y_{it+1} - \beta_n n_{it+1}] &= \beta_0 + \beta_k k_{it+1} + E_t[s_{it+1} | s_{it}, survival], \\ &= \beta_0 + \beta_k k_{it+1} + g(s_{it}, P_{survival}), \end{aligned} \tag{A.1.7}$$

where $P_{survival}$ denotes the one-period probability of survival. Thus, after estimating $P_{survival}$, we can estimate $\hat{\beta}_k$ from equation (A.1.7) by using our estimate of $\hat{\beta}_n$. If we assume that

$g(s_{it}, P_{survival})$ is linear in both arguments and that s_{it} follows an AR(1), $\hat{\beta}_k$ is estimated using the following non-linear equation:

$$\begin{aligned} y_{it+1} - \hat{\beta}_n n_{it+1} &= \beta_0 + \beta_k k_{it+1} + \rho s_{it} + \delta \hat{P}_{survival} + e_{it}, \\ &= \beta_0 + \beta_k k_{it+1} + \rho(\phi_{it} - \beta_0 - \beta_k k_{it}) + \delta \hat{P}_{survival} + e_{it}. \end{aligned} \quad \text{(A.1.8)}$$

To perform this step, $\hat{P}_{survival}$ is estimated using a probit model in capital, investment, and time-industry fixed effects. Once we find $\hat{\beta}_0$, $\hat{\beta}_k$, and $\hat{\beta}_n$, we can derive our estimate of firm-level productivity from equation (A.1.3).

Table 9 presents the distribution of real sales, real intangible expenses, and real capital for firms in 1980 and 2008.[21]

Table 9: Distribution of Sales, Expenses and Assets (%)

	Year 1980			Year 2008		
In millions of US$	Sales	Int. Expenses	Capital	Sales	Int. Expenses	Capital
$ < 10	15.59	27.73	18.20	12.73	21.70	20.42
$ $10 - 20$	7.84	12.41	7.44	5.42	12.74	8.75
$ $20 - 50$	14.29	16.80	11.18	10.12	16.85	11.84
$ $50 - 100$	11.11	11.98	9.29	10.20	13.14	9.26
$ $100 - 250$	15.03	12.94	13.18	14.08	13.65	12.31
$ > 250	36.15	18.14	40.71	47.45	21.92	37.41
# Firms	4,592	4,150	4,006	5,186	4,741	5,084

Note: Data from Compustat. Sales and Expenses are deflated using GDP deflator. Capital is deflated using the price deflator for investment following Hall (1990).

We note that firms' sales, intangible expenses, and capital are considerably larger than those in the KFS sample.

Table (10) presents the distribution of employment size for 1980 and 2008. To simplify the comparison, the size bins are the same as the ones we used for the KFS sample.

[21]Our data extends to 2010 but we present the year 2008 to allow a comparison with the last year of our KFS sample.

Table 10: Distribution of workers (%)

Number of Employees	1980	2008
$1-4$	1.50	1.39
$5-9$	1.79	1.79
$10-19$	2.72	3.32
$20-99$	11.06	13.21
$100-499$	24.06	21.54
$500+$	58.86	58.75
# Firms	4,592	5,186

Note: Data from Compustat.

Most firms in the Compustat sample have more than 500 workers, whereas in the KFS sample this value is less than 1%. Finally, in Table (11) we present the distribution of firm age (computed as the number of years in the sample).

Table 11: Age Distribution (%)

Firm's Age	1980	2008
$1-5$	22.43	31.72
$5-10$	42.97	19.57
$10-15$	14.12	17.03
$15-20$	5.68	10.07
$20-25$	0.00	6.40
$25+$	0.00	11.21
Top Censored	14.80	4.00
# Firms	4,592	4,123

Note: Data from Compustat. Top Censored corresponds to firms that are in our sample starting in 1960.