Firm Expectations, Innovation and Growth*

Georg Duernecker, Marek Ignaszak, Jisu Jeun, and Leo Kaas†

Goethe University Frankfurt

August 5, 2024
Work in Progress

Abstract

Using a large and representative panel survey of German firms, we document sizable forecast errors in employment growth which decline with firm age and which are related to investment and R&D activity. Motivated by this evidence, we build an endogenous growth model with heterogeneous firms which learn their productivity from noisy signals, decide about innovation activity, employment, and exit. Aggregate productivity growth responds to a selection channel via firm entry and exit and to an innovation channel via R&D investments of heterogeneous firms. We calibrate the model to replicate the realized and expected firm growth rates over the firms' lifecycle in our data. We use the calibrated model to quantify the role of information frictions in the selection and innovation channels behind aggregate productivity growth.

Keywords: Imperfect information; Firm dynamics; Learning; Innovation; Growth
JEL classification: D83, D84, E23, L11, O31, O40

*Georg Duernecker and Leo Kaas thank the German Research Foundation (project number 501107822) for financial support.

†E-mail: duernecker@econ.uni-frankfurt.de, ignaszak@uni-frankfurt.de, jeun@em.uni-frankfurt.de, kaas@wiwi.uni-frankfurt.de

1 Introduction

Modern theories of aggregate economic growth emphasize the important role of firm-level differences in size, age and productivity for R&D and innovation activity (c.f. Klette and Kortum, 2004; Lentz and Mortensen, 2008; Akcigit and Kerr, 2018; Acemoglu et al., 2018). While this literature builds on the full-information rational-expectations (FIRE) paradigm, recent research documents sizable deviations from this benchmark, highlighting systematic differences between firm managers' forecasts and the actual realizations of employment, production or sales (Altig et al., 2022; Bachmann et al., 2021). Moreover, firms' subjective expectations matter for current decisions about production, pricing and investment and hence have macroeconomic implications, e.g. Gennaioli et al. (2016), Bachmann and Zorn (2020), Enders et al. (2022).

This paper analyzes how firms' subjective expectations are linked to their innovation activity and how imperfect information and firm learning matter for aggregate economic growth. We build an endogenous growth model featuring heterogeneous firms which are imperfectly informed about their actual productivity and decide about entry, exit, employment and R&D investment. We calibrate the model using panel data with information about R&D activity and firms' expectations about future employment and use the calibrated model to quantify the role of imperfect information for aggregate productivity growth via selection and innovation channels.

We begin in Section 2 by presenting selected empirical findings about employment forecast errors and their relationship with firm characteristics and R&D investments. We use the IAB Establishment Panel which is a representative annual survey of German firms that covers the period 1996–2019 with about 15,000 observations per year. Next to a set of questions about various business characteristics and activities, the firm reports a forecast about employment one-year ahead, next to information about the current stock of employment. Given the panel structure of the data, we calculate expected employment growth rates and define the employment growth forecast error as the difference between the (ex-post) realized employment growth rate and the (ex-ante) expected employment growth rate.

In line with the literature on subjective firm expectations cited above, expected employment growth rates are substantially more compressed than actual employment growth rates. Moreover, we document that firms are on average overly pessimistic, underpredicting their employment growth by about 2.4 percentage points on average. This is a sizable magnitude, as it is more than twice as large as the standard deviation of aggregate annual employment growth in Germany during the same period. Since many firms do not adjust their employment stock from one year to the next, firm-level employment growth exhibits a spike at zero. This spike is even larger for expected employment growth. Consequently, firms underpredict both job creation and job destruction rates, whereby the job creation error exceeds the job destruction error by over two percentage points. These averages mask a considerable degree of heterogeneity. For instance, we find that forecast errors decline in firm age while they are mildly increasing in size. Firms located in East German states are significantly more pessimistic than their West German counterparts, conditional on other firm characteristics. Importantly, we find that firms involved in R&D activities are considerably more optimistic regarding their employment growth which is reflected by a decline in the job creation error while the job destruction error is similar between firms that conduct R&D and those that do not.

In Section 3, we build a growth model with heterogeneous firms which decide about R&D investment while being imperfectly informed about their productivity. The goal of the model is to rationalize our empirical findings and to quantify the role of information frictions for innovation activity and aggregate economic growth. Unobserved firm productivity follows a stochastic process which is driven by exogenous persistent productivity shocks and stochastic innovations that respond to the firm's R&D investment. While a firm perfectly observes whether or not its R&D investment is successful, it receives noisy signals about true productivity which are used to update productivity beliefs over time. Firms produce differentiated products and decide each period about employment and R&D investment on the basis of its productivity belief. As firms also incur fixed operating costs each period, firms with low perceived productivity decide to exit. Every period an endogenous mass of new firms enter the economy after paying a fixed costs in which case they draw their initial productivity and a productivity signal.

We solve the model for a balanced growth path with a stationary firm distribution. Similar to Lentz and Mortensen (2008), Akcigit and Kerr (2018) and Acemoglu et al. (2018), we can decompose aggregate productivity growth into two main channels. The *innovation channel*, which reflects productivity growth due to innovation activity among incumbent firms, and the *selection channel*, which stands for the productivity growth due to the replacement of exiting (unproductive) firms by startups. The innovation channel can be further decomposed into an *average innovation channel* which is the productivity-weighted average of productivity growth of incumbents and a *between-firm innovation channel* which reflects differences in innovation activity by heterogeneous firms. All channels respond to the presence of information frictions in our model: Productive firms which receive a sequence of detrimental signals may decide to exit or to cut their R&D investments which reduces productivity growth through either the selection or the innovation channel.

We consider a preliminary calibration of our model in Section 4 which is consistent with aggregate growth dynamics and a plausible parameterization of the stochastic process governing firm dynamics. The calibrated model is consistent with the main observations in Section 2: Firms strongly underpredict employment growth, with considerably larger expectation errors for younger firms. Moreover, firms with initally optimistic beliefs invest more in R&D. The calibrated model further reveals interesting selection patterns that are governed by initial beliefs, productivity signals and innovation activity which responds to firm learning over time. For example, firms with initially pessimistic beliefs tend to experience positive productivity surprises which induce these firms to revise their beliefs upwards and to invest more into R&D. These firms are eventually more likely to survive than those firms which begin with excessively optimistic beliefs and are then surprised with low productivity realizations. These patterns explain why, on average, rational yet imperfectly informed firms underpredict their employment growth.

Related Literature

By considering firm learning and selection, we build on the seminal contribution of Jovanovic

(1982) and relate to a recent literature studying the quantitative implications of imperfect information and learning for firm growth. Arkolakis et al. (2018) consider a model where young firms learn about their product appeal over time. Their estimated model replicates size-growth and age-growth relationships from the data, revealing sizable distortions at the entry and exit margins. Albornoz et al. (2012) and Eaton et al. (2021) consider models of firm learning to rationalize the patterns of export dynamics. David et al. (2016) analyze how information distortions impact factor misallocation and aggregate productivity. Sterk et al. (2021) distinguish ex-ante heterogeneity and ex-post shocks as distinct driving forces of firm selection and firm growth, but they do so in an environment with perfect information, discussing the consequences of information frictions only briefly. Our work differs from these contributions in that we consider an endogenous innovation choice of heterogeneous firms that allows us to shed light on the implications of firm learning for aggregate economic growth via selection and innovation margins.

The data analysis in our paper relates to a literature on firm-level subjective expectations that utilizes firm surveys with information on forecasts about investment, production, sales or employment (e.g. Guiso and Parigi, 1999; Bachmann et al., 2013, 2021; Altig et al., 2022) where forecasting horizons are typically a year or a quarter. Using ex-post realizations of the same outcome variables in subsequent waves of the survey, forecast errors can be calculated as the difference between realizations and forecasts. In line with our findings, a common result of many studies is that forecasts are much less dispersed than realizations, so that the dispersion of forecast errors is large.

Several studies document systematic biases in firms' expectations. Consistent with our results, Altig et al. (2022) observe negative mean forecast errors for sales growth (-0.9 percentage points) and employment growth (-1.4 percentage points), suggesting that firms are, on average, over-pessimistic.[1] Gennaioli et al. (2016) use earnings growth expectations of CFOs in large U.S. firms and find systematic deviations from rational expectations: managers are excessively optimistic (pessimistic) when past earnings growth is high (low), indicating the presence of extrapolation errors. Similar extrapolation errors, although no evidence for overoptimism (or overpessimism), are documented by Barrero (2022) who uses the survey data developed by Altig et al. (2022). Bachmann and Elstner (2015) use German firm survey data and document systematic expectation errors for firms' production growth such that smaller and more leveraged firms tend to be over-optimistic. Both Barrero (2022) and Bachmann and Elstner (2015) also build dynamic general equilibrium models with heterogeneous firms to analyze the welfare consequences of biased firm expectations. Chen et al. (2023) calibrate a firm dynamics model with learning which replicates the empirical relationship between sales forecast errors and firm age obtained from Japanese panel data. None of these papers considers the relationship between forecast errors, R&D investment and productivity growth which is the focus of our paper.

[1] Altig et al. (2022) and Bloom et al. (2020) elicit quantitative forecast distributions of firm-level outcome variables, such that the mean and standard deviation of these distributions provide firm-level mean forecast and uncertainty measures. The survey data used in our paper only contain a point forecast about the one-year ahead employment stock.

2 Empirical analysis

2.1 Data

The data used in our empirical analysis is taken from the IAB Establishment Panel (IAB-BP). The IAB-BP is an annual panel survey of German establishments which is conducted by the Institute for Employment Research of the Federal German Employment Agency (IAB). The sample of establishments in the IAB-BP is a representative random sample drawn from the universe of German establishments with at least one employee covered by social insurance. Each year, approximately 15,500 establishments are included in the survey. Establishments are asked in each wave to report on a range of establishment-level outcomes, including sales, intermediate inputs, work force dynamics, investment, R&D activities, total wages and average hours worked. Of interest to us are the questions pertaining to changes in the workforce. These questions assess the total number of employees as of June 30 of the current year, as well as the expected number of employees by June 30 of the following year. Concretely, establishments are asked to indicate whether they expect the number of employees 'to remain constant', 'to increase', or 'to decrease'. In addition, if an establishment indicates that it expects the number of employees to increase or decrease, it is also asked for the expected number of employees. Due to the panel structure of the survey, we can compare the expected and actual changes in employment by establishment and over time.

In our analysis, we use an unbalanced panel of establishments covering the period from 1997 to 2019. We exclude establishments with less than five employees and those that have undergone significant structural changes (including merger, closure, outsourcing, or spin-offs). Additionally, we eliminate observations lacking information on actual and expected employment. After sample selection and cleaning of the raw data, 33,797 number of establishments remain yielding a total of 164,287 observations.

2.2 Expected and actual employment growth

We define actual annual employment growth for a given establishment as the log difference

$$g_{t,t+1} = \log n_{t+1} - \log n_t \ ,$$

where n_t and n_{t+1} represent the establishment's employment level in year t and $t+1$, respectively.[2] Similarly, expected employment growth is defined as

$$g^e_{t,t+1} = \log E_t(n_{t+1}) - \log n_t \ ,$$

where $E_t(n_{t+1})$ denotes the expectation in year t of the employment level in year $t+1$. Lastly, we define the forecast error of employment growth as the difference between actual and expected

[2] All our empirical results are unchanged when we define employment growth rates as $g_{t,t+1} = 2\frac{n_{t+1}-n_t}{n_{t+1}+n_t}$, as is standard in the firm-dynamics literature. Log growth rates have simple closed-form expressions in our model which is why we use them in our empirical analysis.

growth,

$$\varepsilon_{t+1} = g_{t,t+1} - g^e_{t,t+1} \ . \tag{1}$$

A higher than expected growth rate implies a positive forecast error. We refer to this case as reflecting pessimistic expectations by establishments, while a negative forecast error is referred to as reflecting optimistic expectations.

The forecast error defined in Equation (1) can be decomposed into the error pertaining to job creation and job destruction, respectively. Following the existing literature, job creation is defined as $jc_{t,t+1} = \max\{g_{t,t+1}, 0\}$, representing employment gains, while job destruction is defined as $jd_{t,t+1} = \max\{-g_{t,t+1}, 0\}$, representing employment losses. Similarly, expected job creation and job destruction are derived from the respective expected growth rates. Consequently, the forecast error can be decomposed as

$$\varepsilon_{t+1} = \underbrace{(jc_{t,t+1} - jc^e_{t,t+1})}_{\varepsilon^{jc}_{t+1}} - \underbrace{(jd_{t,t+1} - jd^e_{t,t+1})}_{\varepsilon^{jd}_{t+1}} \ ,$$

where ε^{jc}_{t+1} and ε^{jd}_{t+1} refer to the forecast errors in job creation and job destruction, respectively.

In the first step of our analysis, we pool the observations of all survey waves and compute expected and actual employment growth, as well as the implied expectation error. We use the cross-sectional survey weights provided by the IAB-BP to make the reported statistics representative of the population of German establishments.[3] The results are presented in Table 1. According to these results, employment at the establishment level grows at an annual rate of 3.1% on average. However, establishments substantially under-predict employment growth. We find that the expected growth rate is only 0.7%, which implies a forecast error of 2.4 percentage points. In other words, German establishments tend to have pessimistic expectations, on average, about the growth of their workforce. These mean values mask a substantial amount of heterogeneity across establishments in expected and actual employment growth rates, as indicated by sizable standard deviations shown in Table 1. We will analyze this heterogeneity in more detail below.

Interestingly, when decomposing the total forecast error ϵ we obtain a positive value for the job creation error ϵ^{jc} and for the job destruction error ϵ^{jd}. These values indicate that establishments tend to under-predict the magnitude of employment expansions and employment contractions at the same time.

Considering the evolution of expected and actual growth rates over time, panel (a) of Figure 1 shows that establishments have consistently underestimated employment growth (on average)

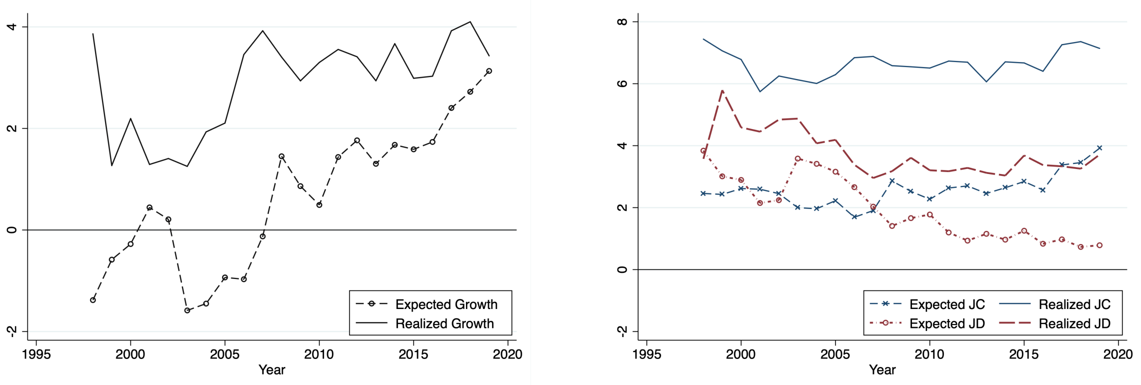
[3] Due to the survey design, the IAB-BP sample overrepresents larger establishments, smaller federal states, and specific sectors. To correct for this bias, the data offers two weighting factors: cross-sectional and longitudinal. While both aim to ensure representativeness within the German establishment population, longitudinal weights are confined to establishments consistently observed in every wave since the initial appearance, making them well-suited for panel analyses but at the expense of sample size and period length. Hence, we mainly apply cross-sectional weights to achieve extensive coverage across the entire sample period. From a robustness check, we find that results using longitudinal weights are consistent with our main findings.

Table 1: Actual and expected employment growth

Employment Growth (in %)		**Forecast Error** (pp)		
Realized (g)	**Expected** (g^e)	**Growth** (ε)	**JC** (ε^{jc})	**JD** (ε^{jd})
3.1 (21.4)	0.7 (15.5)	2.4 (23.2)	4.2 (16.6)	1.8 (14.2)

NOTES: Pooled sample 1997-2019. Standard deviations in parentheses.

during the entire period covered by our data. Growth expectations have become less pessimistic over time, leading to a gradual reduction in the forecast error from 2.5 percentage points in 2005 to 1.7 percentage points in 2015.



(a) Growth rate

(b) Job creation and destruction rate

Figure 1: Growth expectation and the realization

However, Panel (b) shows that the decline in the forecast error is primarily due to an increase in the job destruction error over time. That is, establishments increasingly underestimate the magnitude of employment contractions. In contrast, the job creation errors has remained relatively constant over time.

Finally, we examine how forecast errors are related to the characteristics of the establishment. To this end, we estimate Ordinary Least Squares (OLS) regressions

$$\varepsilon_{i,t+1} = \beta_0 + \beta_1 X_{it} + \beta_2 R\&D_{it} + \gamma_t + \nu_{it} \ , \tag{2}$$

where $\varepsilon_{i,t+1}$ is firm i's error in predicting either employment growth, job creation, or job destruction for year $t+1$, X_{it} is a vector containing firm age, firm size and industry categories and a binary indicator for location in East German states, $R\&D_{it}$ is a category for the firm's engagement in research and development activities, and γ_t are year fixed effects.

The results in Table 2 indicate how forecast errors are associated with firm characteristics. First, the employment growth forecast error decreases with age, suggesting that younger establishments tend to be more pessimistic. For example, establishments less than two years old exhibit forecast errors that are 4.4 percentage points higher than firms older than 20 years. This

Table 2: Relationship between forecast errors and firm characteristics

	(1)	(2)	(3)
Dependent variable	**Growth error**	**JC error**	**JD error**
Age (0 to 2)			
3 to 5	-0.028**	-0.020**	0.009
6 to 10	-0.036***	-0.030***	0.006
11 to 15	-0.045***	-0.037***	0.008
16 to 20	-0.039***	-0.038***	0.000
21+	-0.041***	-0.040***	0.001
Left-Censored	-0.052***	-0.045***	0.007
Size (5 to 19)			
20 to 49	0.015***	0.010***	-0.005**
50 to 199	0.010***	-0.003	-0.006***
200+	0.013***	0.001	-0.012***
East Germany	0.017***	0.010***	-0.008***
R&D (No)			
Up to 1%	-0.007*	-0.006*	0.001
1 to 5%	-0.008**	-0.009***	-0.001
5 to 10%	-0.018**	-0.013**	0.006
More than 10%	-0.007	-0.012**	-0.005
Constant	0.071***	0.072***	-0.000
Industry FE	✓	✓	✓
Year FE	✓	✓	✓
N	129,442	129,442	129,442
R-sq	0.012	0.014	0.008

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

relationship is due to the age relationship of the job creation error, as shown in column (2), while expectation errors regarding job destruction do not vary with age. Put differently, younger firms are much more pessimistic regarding their job creation compared to older firms, while firms of all ages underpredict job destruction to a similar degree.

Regarding establishment size, there is a modest difference between very small establishments (5 to 19 employees) and larger establishments whose employment growth forecast errors are 1.0 to 1.5 percentage points larger. Here, forecast errors in job destruction primarily contribute to the size differences, as larger firms are consistently less optimistic regarding their job losses compared to very small firms. Job creation forecast errors do not show meaningful relationships with firm size.

Establishments located in East German states are more pessimistic than their West German counterparts, with 1.8 percentage points higher forecast errors in employment growth. The differences arises since East German establishments are both more pessimistic regarding their job creation and less optimistic regarding their job destruction, with differences of forcast errors relative to West German establishments standing at 1.0 and 0.8 percentage points, respectively.

Lastly, establishments engaged in R&D activities are more optimistic. Involvement in R&D is associated with a reduction of the employment growth forecast error by one percentage point, suggesting less pessimistic expectations. This difference is driven by a more optimistic forecast on job creation, reflected in a 1.1 percentage points lower forecast error. On the other hand, R&D activity is not related with the job destruction forecast error.

Overall, our results suggest that firms underpredict their employment growth on average. These forecast errors diminish as firms age, and they are lower when firms engage in R&D activities. To explain these findings and to quantify the contribution of imperfect information for innovation and aggregate growth, we construct in the next section a growth model which incorporates entry and exit dynamics, learning over the firm's life cycle and choices about R&D investment.

3 Model

Time in the model is discrete and indexed with t. The model economy is populated by an infinitely-lived representative household and a continuous measure of heterogeneous firms.

3.1 Household

The infinitely-lived representative household has preferences over streams of consumption as described by the utility function

$$\sum_{t=0}^{\infty} \beta^t \log(C_t) \tag{3}$$

with discount factor $\beta \in (0, 1)$. C_t is a composite of different varieties of consumption goods and is defined as

$$C_t = \left(\int_0^{M_t} c_{jt}^{\frac{\eta-1}{\eta}} \, dj \right)^{\frac{\eta}{\eta-1}} ,$$

where c_{jt} is the consumption of variety j, M_t is the mass of available varieties at time t, and $\eta > 1$ is the elasticity of substitution between varieties. The household is endowed with one unit of labor which it supplies inelastically, earning a wage W_t. The household holds a balanced portfolio of all firms in the economy. The value of this portfolio is denoted by K_t and the interest rate is denoted i_t.

When p_{jt} is the price of variety j, we can express aggregate consumption expenditures as $P_t C_t = \int_0^{M_t} p_{jt} c_{jt} dj$ with price index

$$P_t \equiv \left(\int_0^{M_t} p_{jt}^{1-\eta} dj \right)^{\frac{1}{1-\eta}} ,$$

and the household's demand for variety j

$$c_{jt} = \left(\frac{p_{jt}}{P_t} \right)^{-\eta} C_t \ . \tag{4}$$

The household's budget constraint is

$$P_t C_t + K_{t+1} = W_t + (1 + i_t) K_t \ , \tag{5}$$

so that the familiar consumption Euler equation is obtained as

$$\frac{P_{t+1} C_{t+1}}{P_t C_t} = \beta (1 + i_{t+1}) \ . \tag{6}$$

See Section A.1 for further details.

3.2 Firms

3.2.1 Productivity and Learning

Each consumption variety j is produced by one firm j. Simplifying notation, we drop index j from all firm-specific variables in the following. Each firm is characterized by a persistent component of idiosyncratic productivity, denoted by z_t, which evolves stochastically over time. The law of motion of z_t is defined below in Equation (8). We assume that firms cannot directly observe z_t. Rather, they infer it from the realizations of actual productivity in period t which is denoted s_t and given by

$$s_t = z_t + \varepsilon_t \ ,$$

where ε_t is a transitory productivity component which is normally distributed with mean zero and variance σ_ϵ^2 and i.i.d. across firms and over time. Each firm produces output using the linear production function $y_t = a_t n_t$ with labor n_t as the only input. Labor productivity is defined by $a_t \equiv e^{\left(s_t - \frac{1}{2}\sigma_\epsilon^2\right)} > 0$.[4]

Firms strive to improve their productivity by investing in R&D. R&D investment is in units

[4]The variance term is subtracted from s_t to make sure that the variance of the transitory productivity shock variance does not affect the productivity level.

of labor. Concretely, if a firm employs r_t units of labor, it generates a successful innovation with probability x_t, assuming the following convex cost function for innovation:

$$r_t = \rho \frac{1}{\psi} \bar{r}_t x_t^{\psi} \quad , \quad 0 \leq x_t \leq 1 \ . \tag{7}$$

Here $\rho > 0$ and $\psi > 1$ are scale and elasticity parameters, and $\bar{r}_t \equiv (a_t/A_t)^{\varphi}$ is a firm-specific scaling factor. The ratio of firm's productivity a_t to aggregate productivity A_t (defined below in Equation (15)) captures that already productive firms may face higher costs of innovating and becoming even more productive. The parameter φ governs how much harder it gets for a productive firm to become even more productive. After a successfully innovation, the firm's labor productivity increases by the factor $1+\lambda > 1$, where λ is identical for all firms and constant over time. Importantly, we assume that firms observe if the innovation was successful or not.[5] The firm's persistent component of productivity z_t evolves stochastically according to

$$z_{t+1} = \begin{cases} z_t + \log(1+\lambda) + \zeta_{t+1} & \text{prob. } x_t \text{ (innovation)} \ , \\ z_t + \zeta_{t+1} & \text{prob. } 1 - x_t \text{ (no innovation)} \ . \end{cases} \tag{8}$$

where ζ_{t+1} is normally distributed with mean zero and variance σ_ζ^2, and i.i.d. across firms and over time. Taken together, the firm's actual and observed productivity s_t changes over time due to intentional innovation efforts, as well as purely transitory disturbances captured by ε_t and persistent shocks ζ_t. The firm's initial level of the persistent productivity component z_1 is drawn randomly at firm entry, as specified below in Section 3.3.

As firms cannot observe z_t, they infer it from actual productivity realizations s_t. Concretely, the underlying state z_t can be estimated using the a Kalman filter recursion. Let $\hat{z}_t \equiv \mathbb{E}[z_t|s_1, s_2, \ldots s_t]$ denote the firm's current posterior belief about the underlying state z_t after having observed the sequence of past productivity draws $\{s_1, s_2, ..., s_t\}$. The firm enters a given period with prior belief $\hat{z}_{t-1}$, and after observing the realized productivity s_t and learning if the innovation was successful, it updates its belief according to:

$$\hat{z}_t = \begin{cases} \hat{z}_{t-1} + \log(1+\lambda) + \frac{k_t}{k_t+\sigma_\varepsilon^2}(s_t - \hat{z}_{t-1} - \log(1+\lambda)) & \text{prob. } x_{t-1} \text{ (innovation)}, \\ \hat{z}_{t-1} + \frac{k_t}{k_t+\sigma_\varepsilon^2}(s_t - \hat{z}_{t-1}) & \text{prob. } 1 - x_{t-1} \text{ (no innovation)} \ . \end{cases} \tag{9}$$

where k_t is the variance of the one-step predictor, i.e., $k_t \equiv \text{var}(z_t|s_1, s_2, \ldots s_t)$.[6]

The state of the firm in period t is given by the vector $\mathcal{S}_t = (s_t, \hat{z}_t, k_t)$ consisting of observed current productivity, s_t, the value of posterior beliefs $\hat{z}_t$ and the variance of the predictor, k_t.

[5] In our view, this is a natural assumption as otherwise we would imply that firms do not realize if they have discovered new products or technologies, or if they have installed new machines.

[6] The variance evolves according to $k_{t+1} = \frac{k_t \sigma_\varepsilon^2}{k_t+\sigma_\varepsilon^2} + \sigma_\zeta^2$ and approaches in the limit the value $k_\infty = \frac{1}{2}[\sigma_\zeta \sqrt{4\sigma_\varepsilon^2 + \sigma_\zeta^2} + \sigma_\zeta^2]$.

3.2.2 Firms' Choices and Profits

At the beginning of a period, a firm observes its state $\mathcal{S}_t$ and first decides whether to stay in the market and produce, or to exit. The value of exit is set to zero. Conditional on staying, the firm hires production labor n_t and R&D labor r_t (which determines the innovation probability x_t), and sets the price p_t for its variety. The flow profit of the firm is given by

$$\pi_t(p_t, n_t, x_t; \mathcal{S}_t) = p_t a_t n_t - W_t n_t - W_t r_t - W_t \gamma_f \ ,$$

where $\gamma_f \geq 0$ is a fixed operating cost (in units of labor). Notice that profits depend only on firm's observed productivity s_t (via a_t), but not on its beliefs, $\hat{z}_t$ or k_t. A firm exits when the net present value of the firm is negative. In addition, the firm faces the possibility of exogenous destruction which occurs with probability δ at the end of a period. The value function of the firm is given by

$$V_t(\mathcal{S}_t) = \max \left\{ 0, \max_{p_t, n_t, x_t} \left[\pi_t(p_t, n_t, x_t; \mathcal{S}_t) + \frac{1-\delta}{1+i_t} \mathbb{E}_{\mathcal{S}_t, x_t} V_{t+1}(\mathcal{S}_{t+1}) \right] \right\} \tag{10}$$

where $\mathbb{E}_{\mathcal{S}_t, x_t}$ is the firm's subjective expectations in period t, conditional on the current state $\mathcal{S}_t$ and the chosen innovation probability x_t. The outer "max" operator in (10) expresses the firm's exit choice, whereas the inner "max" operator represents the optimal labor choices and the pricing decision.

The firm's next period's values of $(s_{t+1}, \hat{z}_{t+1})$ depend on the realization of productivity shocks $(\varepsilon_{t+1}, \zeta_{t+1})$ and on the outcome of the innovation process in period t. With probability x_t, there is an innovation in which case the firm's productivity increases by a factor $1+\lambda$. Hence, we can write the expected continuation value in Equation (10) as

$$\mathbb{E}_{\mathcal{S}_t, x_t} V_{t+1}(\mathcal{S}_{t+1}) = x_t \mathbb{E}_{\mathcal{S}_t, \Lambda_t = 1} V_{t+1}(\mathcal{S}_{t+1}) + (1 - x_t) \mathbb{E}_{\mathcal{S}_t, \Lambda_t = 0} V_{t+1}(\mathcal{S}_{t+1}) \ ,$$

where $\Lambda_t = 1$ $(\Lambda_t = 0)$ indicates the case of an innovation (no innovation) in period t.

As each firm is the sole producer of a given variety, we substitute the demand function in Equation (4) into the firm's profit function, thereby eliminating n_t as a choice variable in the firm's optimization problem. As a result, the remaining choice variables of the firm are the price of its variety, p_t, and the innovation probability, x_t. Notice that the firm's pricing decision is a static choice as today's price does not affect the firm's future value. The first-order condition with respect to p_t yields the familiar pricing rule whereby the firm's optimal price is set as a constant markup over marginal costs:

$$p_t = \frac{\eta}{\eta - 1} \frac{W_t}{a_t} \ . \tag{11}$$

Since the wage W_t is the same across firms, more productive firms (with higher a_t) choose to set a lower price and to produce more. We substitute the optimal price into the consumer's demand function to obtain the firm's labor demand:

$$n_t = \left(\frac{\eta - 1}{\eta}\right)^{\eta} a_t^{\eta-1} \left(\frac{W_t}{P_t}\right)^{-\eta} C_t \ . \tag{12}$$

This expression implies that more productive firms are larger in terms of employment. Next, we solve for the firm's optimal choice of the innovation probability x_t. We replace R&D labor r_t in the firm's profit function by the cost function, defined in Equation (7), and derive the first-order condition with respect to the innovation probability x_t. The condition for an interior solution is given by:

$$W_t \rho \bar{r}_t x_t^{\psi-1} = \frac{1-\delta}{1+i_t} \left[\mathbb{E}_{\mathcal{S}_t, \Lambda_t=1} V_{t+1}\big(\mathcal{S}_{t+1}\big) - \mathbb{E}_{\mathcal{S}_t, \Lambda_t=0} V_{t+1}\big(\mathcal{S}_{t+1}\big)\right] \tag{13}$$

In the optimum, the marginal cost (left-hand side) of an incremental increase in the innovation probability is equal to the marginal gain. For $\varphi > 0$ the marginal cost increases in a_t (via $\bar{r}_t$), implying that more productive firms face a higher marginal cost of innovating. The marginal gain depends on the change in the firm value implied by a successful innovation. As we will see later, firms with a higher productivity belief $\hat{z}_t$ perceive a higher marginal benefit of R&D. At the same time, the marginal gain of R&D is strictly positive and independent of x_t; therefore, it is never optimal for the firm to choose $x_t = 0$. Moreover, since the innovation probability is bounded above by unity, some firms may choose a corner solution with $x_t = 1$. This case occurs when the marginal gain exceeds the marginal cost at $x_t = 1$.

3.3 Entry

Each period, there is a continuum of potential entrants. To enter the economy, entrants must incur a one-time fixed cost $\gamma_e \geq 0$ in units of labor. Upon entering in period t, firms draw their initial value of the persistent component of productivity z_1 from the distribution G_t^e.[7] While the realization of z_1 is unknown to firms, they observe actual productivity in the entry period $s_1 = z_1 + \epsilon_1$. Entering firms have (correct) prior beliefs $\hat{z}_0 = \mathbb{E}_t(z) = \int z_1 dG_t^e(z_1)$. Given the prior $\hat{z}_0$ and observed productivity s_1, the firm forms posterior beliefs $\hat{z}_1$ according to the second line in the updating Equation (9). We assume that firms enter with the perceived variance of z equal to $k = k_\infty$ as defined in Equation 6, so that k is constant over time and hence can be dropped from the state vector. Conditional on the initial state $\mathcal{S}_1 = (s_1, \hat{z}_1)$, the firm decides, prior to production, whether to stay in the market or to exit.

In equilibrium, the expected return of entering must be equal to the cost of entry. The implied free-entry condition is given by

$$W_t \gamma_e = \int V_t(\mathcal{S}_1) d\Psi(\mathcal{S}_1) \tag{14}$$

where $\Psi(\mathcal{S}_1)$ is the distribution of $\mathcal{S}_1$ conditional on the firm's (correct and common) prior belief $\hat{z}_0$.

[7] The mean of the distribution G_t^e grows over time together with aggregate productivity, as specified below.

3.4 Equilibrium

Choosing C_t as the numeraire, the aggregate price index P_t is normalized to one in every period.

3.4.1 Balanced Growth Path and Stationary Representation

We restrict attention to balanced growth path equilibria. Along a balanced growth path all aggregate variables grow at the (constant) growth rate of aggregate productivity. Let A_t denote the level of aggregate productivity at time t defined as

$$A_t = \left(\int_0^{M_t} a_{j,t}^{\eta-1} dj \right)^{\frac{1}{\eta-1}} \tag{15}$$

and denote its growth factor by $1 + g \equiv \frac{A_{t+1}}{A_t}$. To render the model economy stationary, we divide all growing variables by A_t. We denote the stationary variables by their symbols without time subscripts. For example, $C \equiv \frac{C_t}{A_t}$, $W \equiv \frac{W_t}{A_t}$. Regarding idiosyncratic firm productivity, we define $\tilde{a}_{j,t} = \frac{a_{j,t}}{A_t}$, $\tilde{s}_{j,t} = s_{j,t} - \log A_t$, $\tilde{z}_{j,t} = z_{j,t} - \log A_t$ and $\tilde{\hat{z}}_{j,t} = \hat{z}_{j,t} - \log A_t$. We again drop index j for notational convenience. With abuse of notation, we also drop the tilde from these variables and read a_t, s_t, z_t and $\hat{z}_t$ as the stationary representations of these (time-varying idiosyncratic) variables. Likewise, the firm's state $\mathcal{S}_t = (s_t, \hat{z}_t)$ is rewritten in terms of the detrended variables. The stationary value of initial persistent firm productivity is drawn from time-invariant distribution G^e which implies that the distribution of initial firm productivity (before detrending) shifts together with aggregate productivity growth according to $G^e_t(z_1) = G^e_{t-1}(z_1 - \log(1 + g))$.

The stationary representation of the household's budget constraint in Equation (5) is obtained by dividing both sides by A_t and using the fact that household's assets, K_t, grow at rate g. Together with the normalization of the aggregate price level $P = 1$ the budget constraint becomes

$$C = (i - g)K + W \ . \tag{16}$$

Moreover, since $C_{t+1}/C_t = 1 + g$, the Euler equation in (6) implies that

$$\beta = \frac{(1 + g)}{(1 + i)} \ .$$

To render the firm's value function stationary, we divide both sides of Equation (10) by A_t. This transformation changes the firm's effective discount factor by adding the growth factor $1 + g$. The stationary representation of the value function (i.e. the original value function divided by A_t) is given by:

$$V(\mathcal{S}) = \max \left\{ 0, \max_{p,n,r} \left[\pi(p, n, x; \mathcal{S}) + \frac{(1 + g)(1 - \delta)}{1 + i} \mathbb{E}_{\mathcal{S},x} V(\mathcal{S}') \right] \right\} \ , \tag{17}$$

where $\pi(.)$ is the stationary representation of the firm's profit (i.e. π_t divided by A_t), $\mathcal{S}'$ represents the firm's state next period and $\mathbb{E}_{\mathcal{S},x}$ is the firm's subjective expectation conditional on the current information $\mathcal{S}$ and the innovation probability x. The first-order condition for R&D becomes

$$W\rho\bar{r}x^{\psi-1} \leq \frac{(1+g)(1-\delta)}{1+i}\left[\mathbb{E}_{\mathcal{S},\Lambda=1}V(\mathcal{S}') - \mathbb{E}_{\mathcal{S},\Lambda=0}V(\mathcal{S}')\right] \quad , \quad x \leq 1 \ , \tag{18}$$

with complementary slackness.

In the stationary model representation, the law of motion of the persistent component of firm productivity z is given by:

$$z' = \begin{cases} z + \log\left(\frac{1+\lambda}{1+g}\right) + \zeta' & \text{prob. } x \text{ (innovation)} \ , \\ z + \log\left(\frac{1}{1+g}\right) + \zeta' & \text{prob. } 1 - x_t \text{ (no innovation)} \ . \end{cases} \tag{19}$$

This equation illustrates the tension between the firm's productivity growth due to innovation, represented by the term $(1+\lambda)$, and the growth of aggregate productivity, represented by $(1+g)$. Due to constant growth in aggregate productivity, the firm's productivity level declines relative to aggregate productivity. Thus, in the stationary model economy, the firm's relative productivity decreases each period by the factor $1+g$. This decline is counteracted by the firm's innovation effort which increases productivity by the factor $1+\lambda$. This tension is also reflected in the stationary representation of the updating equation for the firm's belief, which is given by:

$$\hat{z}' = \begin{cases} \hat{z} + \log\left(\frac{1+\lambda}{1+g}\right) + \frac{k}{k+\sigma_\epsilon^2}\left(s' - \hat{z} - \log\left(\frac{1+\lambda}{1+g}\right)\right) & \text{prob. } x \text{ (innovation)} \ , \\ \hat{z} + \log\left(\frac{1}{1+g}\right) + \frac{k}{k+\sigma_\epsilon^2}\left(s' - \hat{z} - \log\left(\frac{1}{1+g}\right)\right) & \text{prob. } 1 - x_t \text{ (no innovation)} \ , \end{cases}$$

where $s' = z' + \varepsilon'$ is productivity in the next period.

3.4.2 Stationary Firm Distribution

Let $\mu_t(z, \hat{z})$ denote the period-t measure of active firms over the persistent productivity component z and productivity belief $\hat{z}$. Note that we do not include the firm's actual (observed) productivity s as a separate variable in μ_t. This is because s is obtained from z and the transitory productivity shock which is i.i.d. across firms and over time. We can characterize the evolution of μ_t by taking into account the transitions of firms across the different states. Concretely, the evolution of μ_t between two periods is governed by the following events: (1) Exogenous firm exit; (2) Entry of new firms; (3) Transition of firms between states including endogenous exit. The law of motion of μ_t can be written as:

$$d\mu_{t+1}\big(z_{t+1}, \hat{z}_{t+1}\big) = (1-\delta)\int F(z_{t+1}, \hat{z}_{t+1}|z_t, \hat{z}_t)\ d\mu_t\big(z_t, \hat{z}_t\big) + F^e(z_{t+1}, \hat{z}_{t+1}) \times M_{t+1}^e \ . \tag{20}$$

Here, $F(z_{t+1}, \hat{z}_{t+1}|z_t, \hat{z}_t)$ denotes the probability measure that a firm has persistent productivity component z_{t+1}, productivity belief $\hat{z}_{t+1}$ and stays active in $t+1$, conditional on being in state $(z_t, \hat{z}_t)$ in period t. Likewise $F^e(z_{t+1}, \hat{z}_{t+1})$ is the probability measure that an entrant begins with persistent productivity component z_{t+1}, belief $\hat{z}_{t+1}$ and decides to stay. M_{t+1}^e denote the mass of entering firms in period $t+1$. Exogenous firm exit occurs with probability δ and affects only incumbent firms. Hence, a fraction $(1-\delta)$ of these firms survives until the next period. See Appendix A.2 for details on calculating $F(z_{t+1}, \hat{z}_{t+1}|z_t, \hat{z}_t)$ and $F^e(z_{t+1}, \hat{z}_{t+1})$.

Along the balanced growth path, the measure function is stationary meaning that $\mu_{t+1}(z, \hat{z}) = \mu_t(z, \hat{z}) \equiv \mu(z, \hat{z})$. The mass of entering firms, M^e, is an equilibrium object that is determined by the free entry condition. Let M denote the mass of active firms in a stationary equilibrium, i.e. $M = \int d\mu(z, \hat{z})$.

3.4.3 Labor, Goods, and Capital Market Clearing

Here we describe the market clearing conditions in stationary equilibrium. Further details about the balanced growth path are contained in Appendix A.3.

In the labor market, the household supplies one unit of labor inelastically. Total labor demand is composed of the firms' labor demand for production n, innovation r, and to cover the fixed costs of production γ_f. Moreover, entering firms demand labor to cover the fixed costs of entry γ_e.[8]

In stationary equilibrium there is a constant mass M of active firms in every period and M^e firms that pay the entry cost. As specified above, the stationary measure of firms over the persistent productivity component z and the belief $\hat{z}$ is denoted $\mu(z, \hat{z})$. Note again that the firm's actual productivity in a given period is $s = z + \varepsilon$ where ε is drawn from a normal distribution with pdf $\Phi(.; 0, \sigma_\varepsilon^2)$. The labor market clears if

$$1 = \int \int n(z+\varepsilon) + r(z+\varepsilon, \hat{z}) d\mu(z, \hat{z}) d\Phi(\varepsilon; 0, \sigma_\varepsilon^2) + M\gamma_f + M^e\gamma_e \; . \tag{21}$$

Here, $n(s)$ and $r(s, \hat{z})$ is the demand for production and R&D workers of a firm in state $(s, \hat{z})$. The former is obtained from Equation (12), which becomes in stationary equilibrium

$$n(s) = \left(\frac{\eta-1}{\eta}\right)^\eta W^{-\eta} C e^{(\eta-1)\left(s-\frac{1}{2}\sigma_\epsilon^2\right)} \; . \tag{22}$$

Demand for R&D workers is obtained from the first-order condition (18) and the innovation production function (7).

The goods market clears because all goods produced by firms are sold to the representative household. By implication, aggregate expenditures for consumption goods $PC = C$ are identical to aggregate revenues of firms which are $\int p_j y_j dj$.

The capital market clears when the assets held by the representative household are identical to the aggregate value of all firms in the economy. This identity follows from Walras's law. To see this, write F_t for the value of a mutual fund that holds all the firms in the economy at the beginning of a period. The recursive formulation of F_t is

$$F_t = C_t - W_t + \frac{1}{1+i_t} F_{t+1} \; ,$$

[8] Notice that the entry cost is sunk. That is, all entering firms pay γ_e, also those firms which exit (prior to production) after having observed initial productivity s_1.

because $C_t - W_t$ are aggregate profits, i.e. revenues minus labor costs for production, R&D, operating costs, and entry. Divide F_t by aggregate productivity A_t and impose a stationary equilibrium so that $\tilde{F} = F_t/A_t$ is constant, satisfying the steady-state condition

$$(i-g)\tilde{F} = (1+i)(C-W) \ .$$

From the household budget constraint (16) follows that $\tilde{F} = (1+i)K$, so that the value of the mutual fund is identical to the value of household assets at the beginning of a period (cum-dividend).

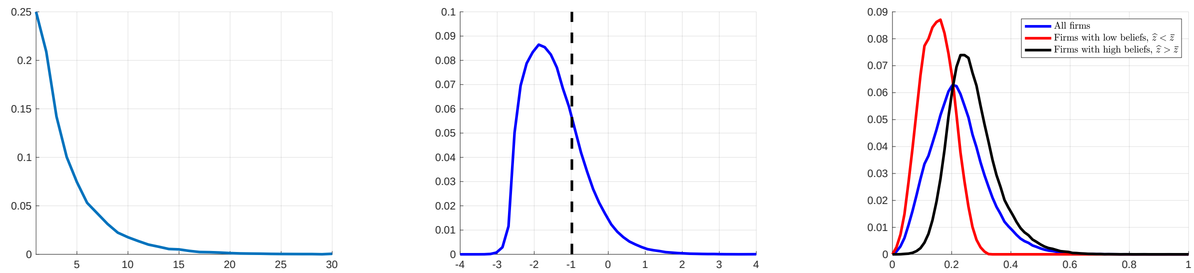
4 Quantitative Model Evaluation

The purpose of this section is to quantitatively examine the equilibrium properties of the model economy and, in particular, to study the effect of imperfect information on firms' decisions and aggregate growth. We consider this analysis to be a first exploratory step rather than a fully-fledged calibration, which we plan to implement in future work. Table 3 shows the parameter values used in the analysis. The model period is one year. The equilibrium of the model is solved numerically. See Appendix C for a description of the computational algorithm.

Table 3: Parameter values

Household		
β	0.983	Personal discount factor
η	2.5	Elasticity of substitution between c_i
Firm shocks		
δ	0.20	Probability of exogenous firm destruction
σ^2_ϵ	0.10	Variance of temporary productivity shock ϵ
σ^2_ζ	0.05	Variance of persistent productivity shock ζ
Distribution of initial productivity z_1		
μ_e	-0.5	Mean
σ^2_e	0.2	Variance
Fixed costs		
γ_f	0.1	Fixed cost of production
γ_e	0.2	Fixed cost of entry
Innovation and R&D technology $r = \frac{\rho}{\psi} a^\varphi x^\psi$		
λ	0.10	Productivity increase after innovation
ρ	2.0	Scaling factor
φ	1.0	Elasticity wrt firm productivity
ψ	2.5	Elasticity wrt innovation probability

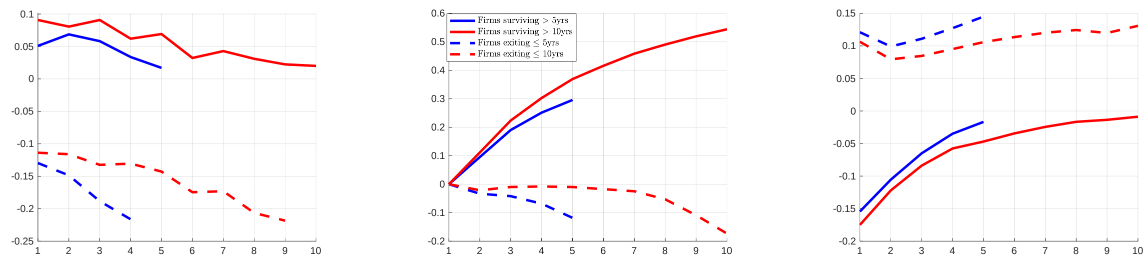
For our calibration, the model implies a real interest rate of 4% and an annual aggregate growth rate of the economy of 2.2%. As can been seen in Panels (a) and (b) of Figure 2, the firm size and age distributions are positively skewed implying that the majority of firms in the economy are young and small. Moreover, Panel (c) of the same figure shows that all active firms' engage in R&D and choose an interior value of the innovation probability x.



(a.) Age (yrs) (b.) Log(Employment) (c.) Innovation probability x

Figure 2: Firm distribution across age, size and innovation

After entry into the economy, firms perform very differently in terms of survival and employment growth; see Panel (a) in Figure 3. While some firms expand rapidly after entry and grow into mature firms, other firms decline and eventually exit. This process is governed by two types of selection effects. The first one is the classical selection effect, which is typical for the class of models following Hopenhayn (1992). Accordingly, after entry, some firms experience a sequence of positive productivity shocks that lead to employment growth at the firm level, while other firms experience negative productivity shocks that lead to the decline and eventual demise of these firms.



(a.) Employment growth (b.) Change in beliefs $\hat{z} - \hat{z}_1$ (c.) Forecast Error $\hat{z} - z$

Figure 3: Employment growth, beliefs and forecast errors after entry

The second selection effect is due to the presence of imperfect information at the firm level. In particular, we find that firms which initially underestimate (overestimate) their true productivity z tend to experience positive (negative) employment growth after entry and are more likely to survive (exit). The intuition is as follows. Consider firms with pessimistic initial beliefs, $\hat{z}_1 < z_1$. Over time, these firms tend to experience a sequence of positive productivity surprises. As a consequence of these surprises, firms gradually revise their productivity beliefs upwards, as illustrated by the solid lines in Panel (b) of Figure 3. As beliefs improve, firms invest more in R&D, since $\frac{\partial x(s,\hat{z})}{\partial \hat{z}} > 0$. This leads to higher productivity at the firm level. More productive firms demand more labor, as implied by $\frac{\partial n(s)}{\partial s} > 0$. Hence, we obtain the result that initially pessimistic firms experience on average positive productivity and employment growth after entry. This effect diminishes over time as firms gradually learn their true productivity; see

Panel (c) in Figure 3.

The opposite effect occurs for firms with optimistic initial beliefs. These firms tend to experience negative productivity surprises over time which leads to sustained downward revisions of beliefs. With deteriorating beliefs, firms invest less in R&D. As a result, productivity gradually declines, which in turn leads to negative employment growth at the firm level and eventually to firm exit.

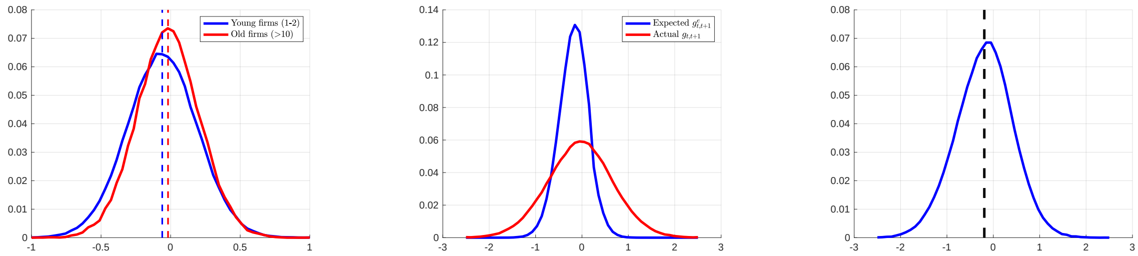
Two remarks are in order. First, the two selection effects described previously operate independently of one another. The classical selection effect à la Hopenhayn (1992) does not require imperfect information, whereas the second effect operates also in the absence of shocks to firms' productivity, that is when $\sigma_\zeta^2 = \sigma_\epsilon^2 = 0$. However, the second effect is reinforced by the first one. Consider a positive shock to firm productivity z. This shock leads to an upward-revision of beliefs and thereby stimulates R&D which, in turn, further raises firm's productivity.

Second, notice that there is a tension between the firms' speed of learning and their innovation activities. To understand this tension, consider firms with optimistic initial beliefs, as reflected by $\widehat{z}_1 > z_1$. As described previously, these firms tend to experience negative productivity surprises, which ultimately lead to a decline in firm productivity z, due to reduced innovation efforts as a result of deteriorating beliefs. However, this effect is counteracted by the fact that the optimistic initial beliefs, per se, induce firms to invest strongly in R&D after entry. This can be seen in Panel (c) of Figure 2. There, the black line shows that firms with above-average beliefs $\widehat{z} > \bar{z}$, choose a higher innovation probability than firms with below-average beliefs (red line). Therefore, if firms' speed of learning is sufficiently slow, then the optimistic initial belief becomes self-fulfilling as the firms' sustained R&D investment leads to an improvement in actual productivity z.

We find that for our calibration, the average of the forecast error $\widehat{z} - z$ across all firms in the economy is negative and equal to -0.043. This means that, on average, firms in the economy are pessimistic because they underestimate their true productivity. This result is a consequence of the (second) selection effect described above. Accordingly, the optimistic firms, i.e. those with $\widehat{z} > z$, are more likely to exit, while the pessimistic firms tend to grow and survive. In addition, the distribution of productivity forecast errors for young firms is more dispersed than the distribution for firms older than 10 years. This is illustrated in Panel (a) of Figure 4. Moreover, the mean also differs, suggesting that young firms are on average more pessimistic than older firms, with an average forecast error of -0.061, compared to -0.018 for older firms.

In the data, we do not observe firms' beliefs about productivity, but their expectations of firm-level employment growth and the corresponding realizations, as described above in Section 2. Thus, in the next step, we examine the model's predictions of firms' expected and actual employment growth and compare these predictions with the empirical results in Section 2.

Let $g_{t,t+1}$ and $g^e_{t,t+1}$ denote the actual and expected log employment growth rates at the firm level between periods t and $t+1$, respectively. Using the firm's optimal labor choice in



(a.) Forecast Error $\widehat{z} - z$ (b.) Employment growth (c.) $g^e_{t,t+1} - g_{t,t+1}$

Figure 4: Forecast error, expected and actual employment growth

Equation (22), these growth rates can be written as[9]

$$g_{t,t+1} = \log n_{t+1} - \log n_t = (\eta - 1)(s_{t+1} - s_t)$$

$$g^e_{t,t+1} = \mathbb{E}(\log n_{t+1} | s_t, \hat{z}_t) - \log n_t = (\eta - 1)\left[\hat{z}_t - s_t - \log(1+g) + x_t \log(1+\lambda)\right] .$$

Consistent with the evidence presented in Section 2, the model generates a distribution of expected employment growth which is much less dispersed than the distribution of actual employment growth; see Panel (b) in Figure 4.

In the model, we calculate the forecast error of employment growth as:

$$\begin{aligned} e_{t+1} &= g_{t,t+1} - g^e_{t,t+1} \\ &= (\eta - 1)\left[s_{t+1} - \hat{z}_t + \log(1+g) - x_t \log(1+\lambda)\right] \\ &= (\eta - 1)\begin{cases} z_t - \hat{z}_t + \varepsilon_{t+1} + \zeta_{t+1} + (1 - x_t)\log(1+\lambda) & \text{, prob. } x_t \\ z_t - \hat{z}_t + \varepsilon_{t+1} + \zeta_{t+1} - x_t \log(1+\lambda) & \text{, prob. } 1 - x_t \end{cases} \end{aligned}$$

Forecast errors arise due to four reasons: (i) transitory productivity shocks ε_{t+1}, (ii) persistent productivity shocks ζ_{t+1}, (iii) stochastic innovations, (iv) the "pessimism" term $z_t - \hat{z}_t$. To understand the latter, if $\hat{z}_t < z_t$ the firm underpredicts true productivity, hence it underpredicts future employment growth and has a systematically positive forecast error. The reverse is the case for an optimistic firm where $\hat{z}_t > z_t$.

In the model, the forecast error e_{t+1} is positive, on average, implying that firms systematically underestimate employment growth. This result is in line with the evidence presented in Section 2. Note that the gap between productivity belief $\hat{z}_t$ and actual productivity z_t is the only source of a systematic forecast error because of

$$\mathbb{E}_t e_{t+1} = (\eta - 1)(\hat{z}_t - z_t) .$$

As mentioned before, young firms in the model have larger forecast errors $\hat{z}_t - z_t$ than older firms. Consequently, the model predicts – in line with the data – that young firms are more pessimistic, on average, about employment growth than older firms; see Figure 5.

[9]See Appendix B.1 for details on the calculation.

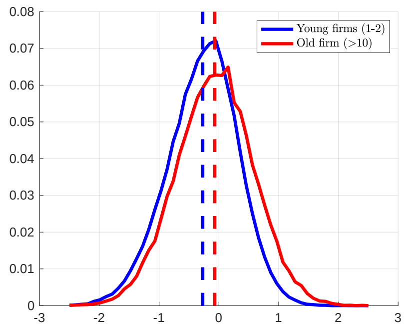


Figure 5: Forecast error of employment growth

5 Conclusion

To be written.

References

Acemoglu, D., U. Akcigit, H. Alp, N. Bloom, and W. Kerr (2018). Innovation, reallocation, and growth. *American Economic Review 108*(11), 3450–3491.

Akcigit, U. and W. R. Kerr (2018). Growth through heterogeneous innovations. *Journal of Political Economy 126*(4), 1374–1443.

Albornoz, F., H. F. C. Pardo, G. Corcos, and E. Ornelas (2012). Sequential exporting. *Journal of International Economics 88*(1), 17–31.

Altig, D., J. M. Barrero, N. Bloom, S. J. Davis, B. Meyer, and N. Parker (2022). Surveying business uncertainty. *Journal of Econometrics 231*(1), 282–303.

Arkolakis, C., T. Papageorgiou, and O. A. Timoshenko (2018). Firm learning and growth. *Review of Economic Dynamics 27*, 146–168.

Bachmann, R., K. Carstensen, S. Lautenbacher, and M. Schneider (2021). Uncertainty and change: survey evidence of firms' subjective beliefs. NBER Working Paper No. 29430.

Bachmann, R. and S. Elstner (2015). Firm optimism and pessimism. *European Economic Review 79*, 297–325.

Bachmann, R., S. Elstner, and E. R. Sims (2013). Uncertainty and economic activity: Evidence from business survey data. *American Economic Journal: Macroeconomics 5*(2), 217–249.

Bachmann, R. and P. Zorn (2020). What drives aggregate investment? evidence from german survey data. *Journal of Economic Dynamics and Control 115*, 103873.

Barrero, J. M. (2022). The micro and macro of managerial beliefs. *Journal of Financial Economics 143*(2), 640–667.

Bloom, N., S. J. Davis, L. Foster, B. Lucking, S. Ohlmacher, and I. Saporta-Eksten (2020). Business-level expectations and uncertainty. NBER Working Paper No. 28259.

Chen, C., T. Senga, C. Sun, and H. Zhang (2023). Uncertainty, imperfect information, and expectation formation over the firm's life cycle. *Journal of Monetary Economics 140*, 60–77.

David, J. M., H. A. Hopenhayn, and V. Venkateswaran (2016). Information, misallocation, and aggregate productivity. *Quarterly Journal of Economics 131*(2), 943–1005.

Eaton, J., M. Eslava, D. Jinkins, and J. R. Tybout (2021). A search and learning model of export dynamics. NBER Working Paper No. 29100.

Enders, Z., F. Hünnekes, and G. Müller (2022). Firm expectations and economic activity. *Journal of the European Economic Association 20*(6), 2396–2439.

Gennaioli, N., Y. Ma, and A. Shleifer (2016). Expectations and investment. *NBER Macroeconomics Annual 30*(1), 379–431.

Guiso, L. and G. Parigi (1999). Investment and demand uncertainty. *The Quarterly Journal of Economics 114*(1), 185–227.

Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica: Journal of the Econometric Society*, 1127–1150.

Jovanovic, B. (1982). Selection and the evolution of industry. *Econometrica*, 649–670.

Klette, T. J. and S. Kortum (2004). Innovating firms and aggregate innovation. *Journal of Political Economy 112*(5), 986–1018.

Lentz, R. and D. T. Mortensen (2008). An empirical model of growth through product innovation. *Econometrica 76*(6), 1317–1373.

Sterk, V., P. Sedláček, and B. Pugsley (2021). The nature of firm growth. *American Economic Review 111*(2), 547–579.

Appendix

A Further Model Details

A.1 Household Problem

We solve the household's optimization problem in two steps. In the first step, we solve the intra-period choice of consumption varieties c_{jt}, for a given level of C. The optimization problem is written as cost minimization problem:

$$\min_{c_{jt}} \int_0^{M_t} p_{jt} c_{jt} dj \qquad \text{subject to} \qquad C_t = \left(\int_0^{M_t} c_{jt}^{\frac{\eta-1}{\eta}} dj \right)^{\frac{\eta}{\eta-1}},$$

where p_{jt} is the price of variety j. The first-order condition with respect to a generic variety j is given by

$$p_{jt} = \gamma_t C_t^{\frac{1}{\eta}} c_{jt}^{-\frac{1}{\eta}}, \tag{23}$$

where γ_t is the Lagrange multiplier associated with the constraint. We use the first-order condition to derive the consumer's demand function as well as the aggregate price index. To this end, we multiply both sides of (23) with c_{jt} and integrate over j to obtain:

$$\int_0^{M_t} p_{jt} c_{jt} dj = \gamma_t C_t \tag{24}$$

Next, we rearrange the first-order condition to express $c_{jt} = \gamma_t^{\eta} C_t p_{jt}^{-\eta}$ and plug this expression into the definition of C_t to obtain:

$$\gamma_t = \left(\int_0^{M_t} p_{jt}^{1-\eta} dj \right)^{\frac{1}{1-\eta}} \equiv P_t$$

We use the definition of the aggregate price index P_t to rewrite the household's demand function for variety j as in equation (4).

In the second step, we solve the household's inter-temporal optimization problem:

$$\max_{C_t, K_{t+1}} \sum_{t=0}^{\infty} \beta^t \log(C_t) \qquad \text{subject to} \qquad P_t C_t + K_{t+1} = W_t + (1+i_t) K_t \ , \tag{25}$$

where K_t is the household's assets and i_t is the interest rate. The first-order conditions with respect to C_t and K_{t+1} are

$$\begin{aligned} C_t : &\qquad \beta^t \tfrac{1}{C_t} - \beta^t \mu_t P_t = 0 &&\Rightarrow \quad \mu_t = \tfrac{1}{P_t C_t} \\ K_{t+1} : &\qquad -\beta^t \mu_t + \beta^{t+1} \mu_{t+1} (1 + i_{t+1}) &&\Rightarrow \quad \mu_t = \beta \mu_{t+1} (1 + i_{t+1}) \end{aligned}$$

where μ_t is the Lagrange multiplier. Combining the two condition yields the familiar consumption Euler equation in equation (6).

A.2 Stationary Firm Distribution

Consider entrant firms in period $t+1$ first. These firms draw the persistent component of productivity z_{t+1} from a normal distribution with mean μ_e and variance σ_e^2, and they observe actual productivity $s_{t+1} = z_{t+1} + \varepsilon_{t+1}$ so that they form the productivity belief

$$\hat{z}_{t+1} = \mu_e + \frac{k}{k+\sigma_\varepsilon^2}(z_{t+1} - \mu_e) + \frac{k}{k+\sigma_\varepsilon^2}\varepsilon_{t+1} \ .$$

As a result, the joint probability to draw z_{t+1} and to form belief $\hat{z}_{t+1}$ is

$$P^e(z_{t+1}, \hat{z}_{t+1}) \equiv \phi(z_{t+1}; \mu_e, \sigma_e^2) \times \phi\Big(\hat{z}_{t+1}; \mu_e + \frac{k}{k+\sigma_\varepsilon^2}(z_{t+1} - \mu_e), \Big(\frac{k}{k+\sigma_\varepsilon^2}\Big)^2 \sigma_\varepsilon^2\Big) \ ,$$

where $\phi(.; \mu, \sigma^2)$ is the pdf of a normal distribution with mean μ and variance σ^2. Conditional on the belief $\hat{z}_{t+1}$, the productivity is

$$s_{t+1} = \mu_e + \frac{k+\sigma_\varepsilon^2}{k}(\hat{z}_{t+1} - \mu_e)$$

and the firm exits endogenously when $\Theta(s_{t+1}, \hat{z}_{t+1}) = 1$ as defined in Equation (10). Consequently, the probability that an entrant begins with $(z_{t+1}, \hat{z}_{t+1})$ and decides to stay in the market is

$$F^e(z_{t+1}, \hat{z}_{t+1}) = \left(1 - \Theta\left(\mu_e + \frac{k+\sigma_\varepsilon^2}{k}(\hat{z}_{t+1} - \mu_e), \hat{z}_{t+1}\right)\right) \times P^e(z_{t+1}, \hat{z}_{t+1}) \ .$$

Next consider continuing firms for which we split the probability that a firm with $(z_t, \hat{z}_t)$ in period t enters period $t+1$ with $(z_{t+1}, \hat{z}_{t+1})$ and decides to stay into two terms reflecting the outcome with and without a successful innovation:

$$\begin{aligned} F(z_{t+1}, \hat{z}_{t+1} | z_t, \hat{z}_t) &= F(z_{t+1}, \hat{z}_{t+1} | \Lambda_t = 1, z_t, \hat{z}_t) \times \mathrm{Prob}(\Lambda_t = 1 | z_t, \hat{z}_t) \\ &\quad + F(z_{t+1}, \hat{z}_{t+1} | \Lambda_t = 0, z_t, \hat{z}_t) \times (1 - \mathrm{Prob}(\Lambda_t = 1 | z_t, \hat{z}_t)) \ . \end{aligned}$$

In Equation (13) we establish that the innovation probability is a function of the firm's state, $x(s_t, \hat{z}_t)$. Using $s_t = z_t + \epsilon_t$, we calculate the innovation probability for firms with state $(z_t, \hat{z}_t)$ as

$$\mathrm{Prob}(\Lambda_t = 1 | z_t, \hat{z}_t) = \frac{\int x\big(z_t + \varepsilon, \hat{z}_t\big)\big(1 - \Theta(z_t + \varepsilon, \hat{z}_t)\big)\phi\big(\varepsilon; 0, \sigma_\varepsilon^2\big) d\varepsilon}{\int \big(1 - \Theta(z_t + \varepsilon, \hat{z}_t)\big)\phi\big(\varepsilon; 0, \sigma_\varepsilon^2\big) d\varepsilon} \ .$$

The innovation probability takes into account that the current draw of ε (thus current actual productivity) is high enough so that the firm decides to stay in the market, hence it is conditional on firm $(z_t, \hat{z}_t)$ being active.

With a successful innovation, the persistent productivity component and the productivity belief evolve according to

$$z_{t+1} = z_t + \log\left(\tfrac{1+\lambda}{1+g}\right) + \zeta_{t+1} \ ,$$
$$\hat{z}_{t+1} = \hat{z}_t + \log\left(\tfrac{1+\lambda}{1+g}\right) + \frac{k}{k+\sigma_\varepsilon^2}(z_{t+1} - \hat{z}_t - \log\left(\frac{1+\lambda}{1+g}\right)) + \frac{k}{k+\sigma_\varepsilon^2}\varepsilon_{t+1} \ ,$$

so that the probability of the joint draw $(z_{t+1}, \hat{z}_{t+1})$ is

$$P(z_{t+1}, \hat{z}_{t+1})|\Lambda_t = 1, z_t, \hat{z}_t) \equiv$$
$$\phi\left(z_{t+1}; z_t + \log\left(\tfrac{1+\lambda}{1+g}\right), \sigma_\zeta^2\right) \times \phi\Big(\hat{z}_{t+1}; \hat{z}_t + \log\left(\tfrac{1+\lambda}{1+g}\right) + \frac{k}{k+\sigma_\varepsilon^2}(z_{t+1} - \hat{z}_t - \log\left(\frac{1+\lambda}{1+g}\right)), \left(\frac{k}{k+\sigma_\varepsilon^2}\right)^2 \sigma_\varepsilon^2\Big) \ .$$

Conditional on the belief $\hat{z}_{t+1}$, the productivity signal is

$$s_{t+1} = \hat{z}_t + \log\left(\frac{1+\lambda}{1+g}\right) + \frac{k+\sigma_\varepsilon^2}{k}\left(\hat{z}_{t+1} - \hat{z}_t - \log\left(\tfrac{1+\lambda}{1+g}\right)\right) \ ,$$

and the firm exits endogenously with probability $\Theta(s_{t+1}, \hat{z}_{t+1})$. Therefore, the probability that a firm in state $(z_t, \hat{z}_t)$ and with a successful innovation enters next period with $(z_{t+1}, \hat{z}_{t+1})$ and decides to stay is

$$F(z_{t+1}, \hat{z}_{t+1}|\Lambda_t = 1, z_t, \hat{z}_t) =$$
$$\left(1 - \Theta\left(\hat{z}_t + \log\left(\tfrac{1+\lambda}{1+g}\right) + \tfrac{k+\sigma_\varepsilon^2}{k}\left(\hat{z}_{t+1} - \hat{z}_t - \log\left(\tfrac{1+\lambda}{1+g}\right)\right), \hat{z}_{t+1}\right)\right) \times P(z_{t+1}, \hat{z}_{t+1})|\Lambda_t = 1, z_t, \hat{z}_t) \ .$$

For those firms without a successful innovation, we obtain very similar expressions where the terms $1+\lambda$ are simply replaced by 1. Therefore,

$$F(z_{t+1}, \hat{z}_{t+1}|\Lambda_t = 0, z_t, \hat{z}_t) =$$
$$\left(1 - \Theta\left(\hat{z}_t + \log\left(\tfrac{1}{1+g}\right) + \tfrac{k+\sigma_\varepsilon^2}{k}\left(\hat{z}_{t+1} - \hat{z}_t - \log\left(\tfrac{1}{1+g}\right)\right), \hat{z}_{t+1}\right)\right) \times P(z_{t+1}, \hat{z}_{t+1})|\Lambda_t = 0, z_t, \hat{z}_t) \ ,$$

with

$$P(z_{t+1}, \hat{z}_{t+1})|\Lambda_t = 0, z_t, \hat{z}_t) \equiv$$
$$\phi\left(z_{t+1}; z_t + \log\left(\tfrac{1}{1+g}\right), \sigma_\zeta^2\right) \times \phi\Big(\hat{z}_{t+1}; \hat{z}_t + \log\left(\tfrac{1}{1+g}\right) + \frac{k}{k+\sigma_\varepsilon^2}(z_{t+1} - \hat{z}_t - \log\left(\frac{1}{1+g}\right)), \left(\frac{k}{k+\sigma_\varepsilon^2}\right)^2 \sigma_\varepsilon^2\Big) \ .$$

A.3 Balanced Growth Path

Note that aggregate consumption is given by

$$C_t = \left(\int_0^{M_t} c_{jt}^{\frac{\eta-1}{\eta}}\, dj\right)^{\frac{\eta}{\eta-1}} \ , \tag{26}$$

We consider a balanced growth path in which the number of firms is constant. Note that the household optimization implies $c_{jt} = \left(\frac{p_{jt}}{P_t}\right)^{-\eta} C_t$. Together with firms' pricing decisions, we can express the consumption of each variety in an equilibrium as

$$c_{jt} = P_t^{\eta} \left(\frac{\eta}{\eta - 1} \frac{W_t}{A_t} \right)^{-\eta} C_t \tag{27}$$

Let us normalize the aggregate price index to unity. The aggregate consumption is then given by

$$\begin{aligned} C_t &= \left(\int_0^{M_t} \left[\left(\frac{\eta}{\eta - 1} \frac{W_t}{A_{j,t}} \right)^{-\eta} C_t \right]^{\frac{\eta-1}{\eta}} dj \right)^{\frac{\eta}{\eta-1}}, \\ &= C_t W_t^{-\eta} \left(\frac{\eta}{\eta - 1} \right)^{-\eta} \left(\int_0^{M_t} A_{j,t}^{\eta-1} dj \right)^{\frac{\eta}{\eta-1}}, \end{aligned} \tag{28}$$

Define the aggregate productivity index as

$$A_t = \left(\int_0^{M_t} A_{j,t}^{\eta-1} dj \right)^{\frac{1}{\eta-1}} , \tag{29}$$

which implies that

$$W_t = \frac{\eta - 1}{\eta} \left(\int_0^{M_t} A_{j,t}^{\eta-1} dj \right)^{\frac{1}{\eta-1}} = \frac{\eta - 1}{\eta} A_t. \tag{30}$$

Hence the wage is proportional to the aggregate productivity index, and on the balanced growth path the normalized wage is equal to

$$W = \frac{\eta - 1}{\eta} .$$

From labor market clearing we get that

$$N = 1 = \int_0^{M_t} (n_{j,t} + r_{j,t}) dj + M_t \gamma_f + M^e \gamma_e \tag{31}$$

$$= \left(\frac{\eta - 1}{\eta} \right)^{\eta} W_t^{-\eta} C_t \int_0^{M_t} A_{j,t}^{\eta-1} dj + \int_0^{M_t} r_{j,t} dj + M_t \gamma_f + M^e \gamma_e \tag{32}$$

The goods market clearing (28) implies that $W_t^{-\eta} = \left(\frac{\eta}{\eta-1} \right)^{\eta} \left(\int_0^{M_t} A_{j,t}^{\eta-1} dj \right)^{\frac{-\eta}{\eta-1}}$ so that (31) becomes

$$1 = \left(\int_0^{M_t} A_{j,t}^{\eta-1} dj \right)^{\frac{-\eta}{\eta-1}} C_t \int_0^{M_t} A_{j,t}^{\eta-1} dj + \int_0^{M_t} r_{j,t} dj + M_t \gamma_f + M^e \gamma_e \tag{33}$$

$$1 = A_t^{-1} C_t + X + M_t \gamma_f + M^e \gamma_e . \tag{34}$$

On the balanced growth path we have

$$1 = C + X + M \gamma_f + M^e \gamma_e ,$$

where

$$X = \frac{\rho}{\psi} \int \int \left(e^{z+\varepsilon-\frac{1}{2}\sigma_\varepsilon^2} \right)^\varphi x(z+\varepsilon, \hat{z})^\psi d\mu(z, \hat{z}) d\phi(\varepsilon; 0\sigma_\varepsilon^2)$$

is aggregate R&D labor input.

B Employment growth

B.1 Employment growth and expectation errors (without exit)

Employment of a firm in period t is

$$n_t = \underbrace{\left(\frac{\eta-1}{\eta} \right)^\eta W^{-\eta} C e^{-(\eta-1)\frac{1}{2}\sigma_\varepsilon^2}}_{\equiv n^*} \times e^{(\eta-1)s_t} \ .$$

Taking logs, we have

$$\log n_t = \log n^* + (\eta - 1)s_t \ ,$$

so that the actual log growth rate of employment is

$$g_{t,t+1} = \log n_{t+1} - \log n_t = (\eta-1)(s_{t+1} - s_t) \ .$$

Regarding expected employment, note first that

$$\begin{aligned} \mathbb{E}(z_{t+1}|s_t, \hat{z}_t) &= x_t \mathbb{E}\left(z_t + \log(1+\lambda) - \log(1+g) | s_t, \hat{z}_t \right) + (1-x_t)\mathbb{E}\left(z_t - \log(1+g) | s_t, \hat{z}_t \right) \\ &= \hat{z}_t - \log(1+g) + x_t \log(1+\lambda) \ , \end{aligned}$$

where $x_t = x(s_t, \hat{z}_t)$ is the innovation probability in period t. Because of $\mathbb{E}(s_{t+1}|s_t, \hat{z}_t) = \mathbb{E}(z_{t+1}|s_t, \hat{z}_t)$, expected log employment is

$$\mathbb{E}(\log n_{t+1}|s_t, \hat{z}_t) = \log n^* + (\eta-1)[\hat{z}_t - \log(1+g) + x_t \log(1+\lambda)] \ .$$

Expected log employment growth is

$$\begin{aligned} g^e_{t,t+1} &= \mathbb{E}(\log n_{t+1}|s_t, \hat{z}_t) - \log n_t \\ &= (\eta-1)\left[\hat{z}_t - s_t - \log(1+g) + x_t \log(1+\lambda) \right] \ . \end{aligned}$$

The forecast error is

$$\begin{aligned} e_{t+1} &= g_{t,t+1} - g^e_{t,t+1} \\ &= (\eta-1)\left[s_{t+1} - \hat{z}_t + \log(1+g) - x_t \log(1+\lambda) \right] \\ &= (\eta-1) \begin{cases} z_t - \hat{z}_t + \varepsilon_{t+1} + \zeta_{t+1} + (1-x_t)\log(1+\lambda) & , \text{ prob. } x_t \\ z_t - \hat{z}_t + \varepsilon_{t+1} + \zeta_{t+1} - x_t \log(1+\lambda) & , \text{ prob. } 1 - x_t \end{cases} \end{aligned}$$

Therefore, forecast errors arise due to four reasons: (i) noisy signals ε_{t+1}, (ii) productivity shocks ζ_{t+1}, (iii) stochastic innovations, (iv) the "pessimism" term $z_t - \hat{z}_t$. To understand the latter, if $z_t - \hat{z}_t > 0$ the firm underpredicts true productivity, hence it underpredicts future employment growth and has a systematically positive forecast error. The reverse is the case for an optimistic firm where $z_t - \hat{z}_t < 0$.

Note that the gap between $\hat{z}_t$ and z_t is the only source of a systematic forecast error because of

$$\mathbb{E}_t e_{t+1} = (\eta - 1)(z_t - \hat{z}_t) \ .$$

B.2 Average (actual) employment growth across surviving firms

The goal is to compute the average employment growth rate across all active firms which survive into the next period.

We start by considering an active firm with current state $(z, \hat{z}, \epsilon)$. Suppose that this firm has a successful innovation (denoted by $\Lambda = 1$). The state of the firm next period is given by $(s', \hat{z}')$, where

$$\begin{aligned} s' &= z' + \epsilon' \\ &= z + \log(1+\lambda) - \log(1+g) + \nu' \ , \\ \hat{z}' &= \hat{z} + \log(1+\lambda) - \log(1+g) + \tfrac{k}{k+\sigma_\epsilon^2}\Big(s' - \hat{z} - \log(1+\lambda) + \log(1+g)\Big) \\ &= \hat{z} + \log(1+\lambda) - \log(1+g) + \tfrac{k}{k+\sigma_\epsilon^2}\Big(z - \hat{z} + \nu'\Big) \ , \end{aligned} \tag{35}$$

where $\nu' = \zeta' + \epsilon'$. Considering these expressions, we can express the firm's next period state in terms of $(z, \hat{z}, \nu')$. The firm remains active in period $t+1$ when $S^1(z, \hat{z}, \nu') = 1$ where S^1 is the survival probability of an innovating firm as defined in Appendix C. Thus, the growth rate of employment for an active firm which survives and innovates is given by:

$$\begin{aligned} g\big(\epsilon, \nu' | \Lambda' = 1, S^1 = 1\big) &= \log(n') - \log(n) \\ &= (\eta - 1)(s' - s) \\ &= (\eta - 1)\Big(\log(1+\lambda) - \log(1+g) + \nu' - \epsilon\Big) \ . \end{aligned} \tag{36}$$

Analogously, the growth rate of employment for an active firm which survives and does not innovate is given by:

$$g\big(\epsilon, \nu' | \Lambda' = 0, S^0 = 1\big) = (\eta - 1)\Big(- \log(1+g) + \nu' - \epsilon\Big) \tag{37}$$

where $S^0(z, \hat{z}, \nu')$ is the survival probability of a firm without an innovation. Notice that conditional on a firm's innovation outcome, the growth rate of employment for a surviving firm is independent of the firm's productivity (s, z) and beliefs $\hat{z}$. Moreover, in the absence of innovation and shocks, the firm's employment gradually declines at rate $\log(1+g)$.

The average employment growth of active firms with state $(z, \hat{z}, \epsilon)$ – conditional on surviving – can be written as:

$$
\begin{aligned}
g_{t+1|t}(z,\hat{z},\epsilon) &= x(z+\varepsilon,\hat{z})\frac{\displaystyle\int S^1(z,\hat{z},\nu')g\big(\epsilon,\nu'|\Lambda'=1,S^1=1\big)d\Phi(\nu')}{\displaystyle\int S^1(z,\hat{z},\nu')d\Phi(\nu')} \\
&+\big(1-x(z+\varepsilon,\hat{z})\big)\frac{\displaystyle\int S^0(z,\hat{z},\nu')g\big(\epsilon,\nu'|\Lambda'=0,S^0=1\big)d\Phi(\nu')}{\displaystyle\int S^0(z,\hat{z},\nu')d\Phi(\nu')} \ .
\end{aligned}
\tag{38}
$$

Let by $S(z, \hat{z}, \epsilon)$ denote the share of firms with state $(z, \hat{z}, \epsilon)$ which survive into the next period. We can write this share as:

$$
S(z,\hat{z},\epsilon) = x(z+\varepsilon,\hat{z})\int S^1(z,\hat{z},\nu')d\Phi(\nu') + \big(1-x(z+\varepsilon,\hat{z})\big)\int S^0(z,\hat{z},\nu')d\Phi(\nu') \ .
$$

Let by $\mu(z, \hat{z}, \epsilon)$ denote the measure of firms with state $(z, \hat{z}, \epsilon)$. The following has to hold:

$$
\begin{aligned}
\mu(z,\hat{z},\epsilon) = 0 \quad &\text{if} \quad \Theta(z+\epsilon,\hat{z}) = 1 \ , \\
\mu(z,\hat{z},\epsilon) > 0 \quad &\text{if} \quad \Theta(z+\epsilon,\hat{z}) = 0 \ .
\end{aligned}
$$

That is, the measure of firms is positive when the state is such that firms remain active (a firm remains active when $\Theta(z + \epsilon, \hat{z}) = 0$). Conversely, the measure of firms is zero when the state is such that firms exit ($\Theta(z + \epsilon, \hat{z}) = 1$). $\mu(z, \hat{z}, \epsilon)$ is given by:

$$
\mu(z,\hat{z},\epsilon) = \mu(z,\hat{z})\frac{(1-\Theta(z,\hat{z},\epsilon))\phi(\epsilon)}{\displaystyle\int (1-\Theta(z,\hat{z},\epsilon))d\Phi(\epsilon)} \tag{39}
$$

where $\mu(z, \hat{z})$ is the stationary firm measure, and $\phi(\epsilon)$ is the pdf. Moreover, the following has to hold: $\mu(z, \hat{z}) = \int \mu(z, \hat{z}, \epsilon)d\epsilon$.

Consequently, $\mu(z, \hat{z}, \epsilon)S(z, \hat{z}, \epsilon)$ is the measure of firms with current state $(z, \hat{z}, \epsilon)$ which survive into the next period. Let by $\mu^S(z, \hat{z})$ denote the measure of firms with current state $(z, \hat{z})$ which survive into the next period. We can write it as

$$
\mu^S(z,\hat{z}) = \int \mu(z,\hat{z},\epsilon)S(z,\hat{z},\epsilon)d\epsilon \ .
$$

Using these expressions, we can write the average employment growth rate across firms with state $(z_t, \hat{z}_t)$, conditional on being active and surviving into the next period as:

$$
g_{t+1|t}(z,\hat{z}) = \int \underbrace{\frac{\mu(z,\hat{z},\epsilon)S(z,\hat{z},\epsilon)}{\mu^S(z,\hat{z})}}_{\text{Weighting factor}} g_{t+1|t}(z,\hat{z},\epsilon)d\epsilon \ . \tag{40}
$$

Lastly, let by $\mu^S = \int d\mu^S(z, \hat{z})$ denote the total mass of surviving firms and write the average employment growth rate across active firms which survive into the next period as

$$g_{t+1,t} = \frac{1}{\mu^S} \int g(z,\hat{z}) d\mu^S(z,\hat{z}) \ . \tag{41}$$

B.3 Expected employment growth of surviving firms

The goal is to compute average expected employment growth across firms which survive into the next period. The first step is to compute expected next period's employment. In the calculation, we take into account that for some future states firms will exit. Concretely, we consider only those future states for which firms will not exit. That is, we compute the expected employment, conditional on surviving. [Note for us: The underlying survey question to the firm could be something like: "If you are still in business next year, what do you think will the number of employees be".]

Consider an active firm with state $(s,\hat{z})$. The firm's next period's employment n' is given by:

$$n' = n^* e^{(\eta-1)s'}$$

where n^* is a positive constant defined above. Let by $S^1(s',\hat{z}') \in \{0,1\}$ denote the firms' survival probability conditional on having a successful innovation ($\Lambda' = 1$). Analogously $S^0(s',\hat{z}') \in \{0,1\}$ is the survival probability in case of no innovation ($\Lambda' = 0$). Realizations of s' are normally distributed. Let by $\Phi^1_f\big(s'|\hat{z},\Lambda' = 1\big)$ denote the firm's subjective probability distribution of s' conditional on beliefs $\hat{z}$ and having a successful innovation. The mean of this (subjective) probability distribution is given by

$$\begin{aligned} E\big(s'|\hat{z},\Lambda'=1\big) &= E\Big(z' + \epsilon'|\hat{z},\Lambda'=1\Big) \\ &= E\Big(z + \log(1+\lambda) - \log(1+g) + \zeta' + \epsilon'|\hat{z}\Big) \\ &= E\big(z|\hat{z}\big) + \log(1+\lambda) - \log(1+g) \\ &= \hat{z} + \log(1+\lambda) - \log(1+g) \end{aligned}$$

The variance of the distribution is

$$\begin{aligned} Var\big(s'|\hat{z},\Lambda'\big) &= E\Big(s' - E(s')\Big)^2 \\ &= E\Big(z - \hat{z} + \epsilon' + \zeta'\Big)^2 \\ &= E\big(z-\hat{z}\big)^2 + E\big(\epsilon^2\big) + E\big(\zeta^2\big) \\ &= k + \sigma^2_\epsilon + \sigma^2_\zeta \end{aligned}$$

where $E\big(z-\hat{z}\big)^2 = k$ is the variance of the forecast error. Similarly, $\Phi^0_f\big(s'|\hat{z},\Lambda'=0\big)$ denotes the firm's subjective probability distribution of s' conditional on beliefs $\hat{z}$ and having no innovation.

To summarize, from the viewpoint of the firm, next period's productivity s' is distributed according to

$$
\begin{aligned}
s' &\sim N\Big(\hat{z} + \log(1+\lambda) \quad - \log(1+g),\ \sigma_\epsilon^2 + \sigma_\zeta^2 + k\Big) \quad \text{(innovation)}\\
s' &\sim N\Big(\hat{z} \qquad\qquad\qquad\quad - \log(1+g),\ \sigma_\epsilon^2 + \sigma_\zeta^2 + k\Big) \quad \text{(no innovation)}
\end{aligned}
$$

Using the previous expressions, we can write expected next period's employment for active firm with state $(s, \hat{z})$ as

$$
E\big(n'|s,\hat{z}\big) = x\big(s,\hat{z}\big)\frac{\displaystyle\int S^1(s',\hat{z}')n^* e^{(\eta-1)s'} d\Phi_f^1(s'|\hat{z})}{\displaystyle\int S^1(s',\hat{z}')d\Phi_f^1(s'|\hat{z})} + \big(1 - x(s,\hat{z})\big)\frac{\displaystyle\int S^0(s',\hat{z}')n^* e^{(\eta-1)s'} d\Phi_f^0(s'|\hat{z})}{\displaystyle\int S^0(s',\hat{z}')d\Phi_f^0(s'|\hat{z})}
$$

where E_t^s indicates the firm's subjective expected value.

Note: Next period's productivity s' for a firm with state $(z, \epsilon, \hat{z})$, is distributed according to

$$
\begin{aligned}
s' &\sim N\Big(z + \log(1+\lambda) \quad - \log(1+g),\ \sigma_\epsilon^2 + \sigma_\zeta^2\Big) \quad \text{(innovation)}\\
s' &\sim N\Big(z \qquad\qquad\qquad\quad - \log(1+g),\ \sigma_\epsilon^2 + \sigma_\zeta^2\Big) \quad \text{(no innovation)}
\end{aligned}
$$

Let by $\Phi^1\big(s'|z, \Lambda' = 1\big)$ (and $\Phi^0\big(s'|z, \Lambda' = 1\big)$) denote the corresponding probability distribution of s' conditional on state z and having a successful innovation (no innovation). Notice that, Φ^1 is different from Φ_f^1, since the latter is the firm's subjective probability distribution. As a result, we can write next period's employment, averaged across surviving firms as

$$
E\big(n'|z,\epsilon,\hat{z}\big) = x\big(s,\hat{z}\big)\frac{\displaystyle\int S^1(s',\hat{z}')n^* e^{(\eta-1)s'} d\Phi^1(s'|z)}{\displaystyle\int S^1(s',\hat{z}')d\Phi^1(s'|z)} + \big(1 - x(s,\hat{z})\big)\frac{\displaystyle\int S^0(s',\hat{z}')n^* e^{(\eta-1)s'} d\Phi_f^0(s'|z)}{\displaystyle\int S^0(s',\hat{z}')d\Phi_f^0(s'|z)}
$$

where $s = z + \epsilon$

C Numerical implementation

We solve the general equilibrium of the model numerically by fixed-point iteration. For this purpose, we develop a numerical algorithm consisting of two nested loops. The outer loop in this algorithm simultaneously iterates over the aggregate growth rate g, aggregate consumption C, and the mass of entering firms M^e. We show in the next section that the detrended wage on a balanced-growth path is given by $W = (\eta - 1)/\eta$. For given values of (g, C, M^e), the inner loop performs the following three steps. In the first step, the value function of the firm $V(s, \hat{z})$ is computed by means of value function iteration. The value function is represented on a discrete grid of 100×100 equally spaced grid points for the firm's state $(s, \hat{z})$. The firm's demand for production labor $n(s)$ follows from (22), while R&D labor $r(s, \hat{z})$ is computed directly from the optimality condition (18). The exit policy function $\Theta(s, \hat{z})$ is obtained from the outer maximization in the Bellman equation (17).

In the second step, the stationary measure $\mu(z, \hat{z})$ is computed (up to a scalar which is the number of entrant firms M^e) by iteration of the law of motion in Equation (20). The state space $(z, \hat{z})$ is represented by a discrete grid with 100×100 equally spaced grid points. When we set $M^e = 1$, we write the stationary measure μ_0 and the mass of firms $M_0 = \int d\mu_0$. For an arbitrary number of entrant firms, the stationary measure μ and the mass of firms M scale linearly according to $\mu = M^e \times \mu_0$ and $M = M^e \times M_0$.

In the third step of the algorithm, the values of (g, C, M^e) are updated. Concretely, the labor market clearing condition in Equation (21) can be written

$$1 = M^e \left[\int \int n(z+\varepsilon) + r(z+\varepsilon, \hat{z}) d\mu_0(z, \hat{z}) d\Phi(\varepsilon; 0, \sigma_\varepsilon^2) + M_0 \gamma_f + \gamma_e \right] .$$

Therefore, we obtain an update for the mass of entrants M^e:

$$M^e = \left(\int \int n(z+\varepsilon) + r(z+\varepsilon, \hat{z}) d\mu_0(z, \hat{z}) d\Phi(\varepsilon; 0, \sigma_\varepsilon^2) + M_0 \gamma_f + \gamma_e \right)^{-1} .$$

The free-entry condition in Equation (14) is used to update the level of aggregate demand C such that consumption is increased (thus raising the value of an entrant firm) if $\int V(\mathcal{S}_1) d\Psi(\mathcal{S}_1) - W\gamma_e$ is negative and consumption is decreased if $\int V(\mathcal{S}_1) d\Psi(\mathcal{S}_1) - W\gamma_e$ is positive.

Lastly the value of the aggregate growth rate g is updated using the definition of the growth factor A in Equation (15). In detrended notation $\tilde{a}_{t,j} = a_{t,j}/A_t$ and dropping the tilde, this condition simply says

$$1 = \int_0^{M_t} a_{j,t}^{\eta-1} dj .$$

In a stationary equilibrium, detrended productivities are $a = e^{z+\varepsilon-\frac{1}{2}\sigma_\varepsilon^2}$ where the persistent productivity component z is drawn from measure $\mu(z, \hat{z})$ and the transitory productivity component ε is normally distributed with pdf $\Phi(.; 0, \sigma_\varepsilon^2)$. Hence,

$$1 = \int e^{(\eta-1)z} \int e^{(\eta-1)(\varepsilon-\frac{1}{2}\sigma_\varepsilon^2)} d\Phi(\varepsilon; 0, \sigma_\varepsilon^2) \; d\mu(z, \hat{z}) = e^{\frac{1}{2}\sigma_\varepsilon^2[(\eta-1)^2-1]} \int e^{(\eta-1)z} d\mu(z, \hat{z}) . \tag{42}$$

To see how the growth rate is determined, consider an active firm $(z, \hat{z})$ in a given period. From the previous section, the firm innovates with probability

$$X(z, \hat{z}) \equiv \frac{\int x\big(z+\varepsilon, \hat{z}\big)\big(1 - \Theta(z+\varepsilon, \hat{z})\big)\phi\big(\varepsilon; 0, \sigma_\varepsilon^2\big) d\varepsilon}{\int \big(1 - \Theta(z+\varepsilon, \hat{z})\big)\phi\big(\varepsilon; 0, \sigma_\varepsilon^2\big) d\varepsilon} .$$

Next period, the firm's persistent productivity component is

$$z' = z - \log(1+g) + \zeta' + \begin{cases} \log(1+\lambda) & \text{prob. } X(z, \hat{z}) \\ 0 & \text{else} \end{cases} ,$$

and the firm's actual productivity is

$$s' = z' + \varepsilon' \ .$$

Therefore labor productivity next period is

$$a' = e^{s' - \frac{1}{2}\sigma_\varepsilon^2} = \frac{1}{1+g} e^{z+\varepsilon'+\zeta'-\frac{1}{2}\sigma_\varepsilon^2} \times \begin{cases} (1+\lambda) & \text{prob. } X(z,\hat{z}) \\ 1 & \text{else} \end{cases} \ .$$

Write the survival probability next period $S^1(z,\hat{z},\varepsilon'+\zeta')$ if the firm innovates and $S^0(z,\hat{z},\varepsilon'+\zeta')$ if the firm does not innovate. We derive S^1 and S^0 below.

Next, consider firms entering next period. These firms draw z' from distribution G^e and transitory productivity ε' so that productivity in the entry period is $s' = z' + \varepsilon'$, and labor productivity is

$$a' = e^{z'+\varepsilon'-\frac{1}{2}\sigma_\varepsilon^2} \ .$$

The entrant firm survives (i.e. does not exit immediately after entering) with probability

$$S^e(z',\varepsilon') \equiv 1 - \Theta(z'+\varepsilon', \mu^e(1-K) + K(z'+\varepsilon')) \ ,$$

where K is the Kalman parameter $K = k/(k+\sigma_\varepsilon^2)$ and μ^e is the mean of G^e.

Now, we utilize the fact that the integral over $(a')^{\eta-1}$ for all firms that are active next period must be identical to one:

$$1 = \int\int (1+g)^{1-\eta} e^{(\eta-1)(z+\nu'-\frac{1}{2}\sigma_\varepsilon^2)} \left[(1+\lambda)^{\eta-1}X(z,\hat{z})S^1(z,\hat{z},\nu') + (1-X(z,\hat{z}))S^0(z,\hat{z},\nu')\right] d\mu(z,\hat{z})d\Phi(\nu';0,\sigma_\nu^2)$$
$$+M^e \int\int e^{(\eta-1)(z'+\varepsilon'-\frac{1}{2}\sigma_\varepsilon^2)} S^e(z',\varepsilon')dG^e(z')d\Phi(\varepsilon';0,\sigma_\varepsilon^2) \qquad (43)$$

Here $\nu' = \varepsilon' + \zeta'$ is normally distributed with variance $\sigma_\nu^2 = \sigma_\varepsilon^2 + \sigma_\zeta^2$.

Define

$$\Gamma(z,\hat{z},\nu') \equiv \left[(1+\lambda)^{\eta-1}X(z,\hat{z})S^1(z,\hat{z},\nu') + (1-X(z,\hat{z}))S^0(z,\hat{z},\nu')\right]$$

to reflect growth of incumbent firms due to their innovation and survival decisions. We further simplify the above integrals by defining joint measures $\mu^I(z,\hat{z},\nu') = \mu(z,\hat{z})\Phi(\nu';0,\sigma_\nu^2)$ and $\mu^E(z,\varepsilon') = M^e G^e(z)\Phi(\varepsilon';0,\sigma_\varepsilon^2)$ for incumbent and entrant firms. Further write $A(z) = e^{(\eta-1)(z-\frac{1}{2}\sigma_\varepsilon^2)}$ to denote the transformation of labor productivity in equation (43). This equation can then be written as a decomposition of aggregate growth into two terms:

$$(1+g)^{\eta-1} = \underbrace{\int A(z+\nu')\Gamma(z,\hat{z},\nu')d\mu^I(z,\hat{z},\nu')}_{=\text{Contribution of incumbent firms}} + \underbrace{\int A(z'+\varepsilon')(1+g)^{\eta-1}S^e(z',\varepsilon')d\mu^E(z',\varepsilon')}_{=\text{Contribution of entrant firms}} \ . \qquad (44)$$

A few remarks:

- We can split the first term into a reallocation term and another "selection" term, similar to Marek's expression below and Lentz and Mortensen. Lentz and Mortensen however have a model with permanent firm types. In our model, firm type z varies over time and responds endogenously to innovation. So I am not sure about the proper interpretation.

- A side remark, isn't a $(1-\delta)$ term missing in the two first terms of Marek's decomposition? The third term actually only reflects entry and not exit. We could write a separate (negative) term which is the productivity loss due to exiting firms (relecting both exogenous and endogenous exit). This could provide a different decomposition.

- The entry term also includes the $(1+g)^{\eta-1}$ factor. The intuition is that productivity of entrants is itself growing with aggregate productivity, reflecting "standing on the shoulder of giants". I guess it is an unavoidable feature.

- In the following, I propose another decomposition that aims to understand the role of innovation and survival for productivity growth of incumbent firms and how they lead to reallocation of production across firms.

Define average survival and innovation probabilities of incumbent firms as follows:

$$\begin{aligned}\bar{S}^1 &= \int S^1(z,\hat{z},\nu')d\mu^I(z,\hat{z},\nu')\ ,\\ \bar{S}^0 &= \int S^0(z,\hat{z},\nu')d\mu^I(z,\hat{z},\nu')\ ,\\ \bar{X} &= \int X(z,\hat{z})d\mu^I(z,\hat{z},\nu')\ .\end{aligned}$$

Next, write actual survival and innovation rates as deviations from their average values:

$$\begin{aligned}S^1(z,\hat{z},\nu') &= \bar{S}^1 + \Delta S^1(z,\hat{z},\nu')\ ,\\ S^0(z,\hat{z},\nu') &= \bar{S}^0 + \Delta S^0(z,\hat{z},\nu')\ ,\\ X(z,\hat{z}) &= \bar{X} + \Delta X(z,\hat{z})\ .\end{aligned}$$

Then split the growth factor of incumbent firms into four terms as follows:

$$\begin{aligned}\Gamma(z,\hat{z},\nu') = &\underbrace{(1+\lambda)^{\eta-1}\bar{X}\bar{S}^1 + (1-\bar{X})\bar{S}^0}_{\equiv\bar{\Gamma}}\\ &+\underbrace{(1+\lambda)^{\eta-1}\bar{X}\Delta S^1(z,\hat{z},\nu') + (1-\bar{X})\Delta S^0(z,\hat{z},\nu')}_{\equiv\Gamma^S(z,\hat{z},\nu')}\\ &+\underbrace{(1+\lambda)^{\eta-1}\Delta X(z,\hat{z})\bar{S}^1 - \Delta X(z,\hat{z})\bar{S}^0}_{\equiv\Gamma^X(z,\hat{z})}\\ &+\underbrace{(1+\lambda)^{\eta-1}\Delta X(z,\hat{z})\Delta S^1(z,\hat{z},\nu') - \Delta X(z,\hat{z})\Delta S^0(z,\hat{z},\nu')}_{\equiv\Gamma^I(z,\hat{z},\nu')}\ .\end{aligned}$$

$\bar{\Gamma}$ is average growth of incumbent firms, $\Gamma^S(z,\hat{z},\nu')$ measures deviations from average growth due to variation in survival rates among heterogeneous firms, $\Gamma^X(z,\hat{z})$ measures deviations from average growth due to variation in innovation rates among heterogeneous firms, and $\Gamma^I(z,\hat{z},\nu')$ is an interaction term (which is hard to interpret...).

With this notation, the contribution of incumbent firms to aggregate productivity growth (the first term in equation (44)) is decomposed as follows

$$\begin{aligned}\int A(z+\nu')\Gamma(z,\hat{z},\nu')d\mu^I(z,\hat{z},\nu') = &\underbrace{\int A(z+\nu')\bar{\Gamma}d\mu^I(z,\hat{z},\nu')}_{=\text{Average growth}} \\ &+\underbrace{\int A(z+\nu')\Gamma^S(z,\hat{z},\nu')d\mu^I(z,\hat{z},\nu')}_{=\text{Reallocation via survival}} \\ &+\underbrace{\int A(z+\nu')\Gamma^X(z,\hat{z})d\mu^I(z,\hat{z},\nu')}_{=\text{Reallocation via innovation}} \\ &+\underbrace{\int A(z+\nu')\Gamma^I(z,\hat{z},\nu')d\mu^I(z,\hat{z},\nu')}_{=\text{Interaction term}} .\end{aligned}$$

The first term simply stands for the average productivity growth of incumbent firms which itself reflects average innovation and survival rates. The other three terms can be positive or negative. The second term is larger when more productive firms have a higher survival probability, in which case employment is reallocated from less productive to more productive firms due to differential decisions whether or not to exit the market. The third term is larger when more productive firms innovate more, in which case employment is reallocated from less productive to more productive firms due to differences in innovation activity. The fourth term has no straightforward interpretation.

It remains to specify the survival probabilities S^1 and S^0. Consider an innovating firm first. This firm enters next period with productivity

$$s' = z' + \varepsilon' = z + \varepsilon' + \zeta' + \log\left(\frac{1+\lambda}{1+g}\right) ,$$

and with the belief

$$\hat{z}' = \hat{z}(1-K) + zK + (\varepsilon'+\zeta')K + \log\left(\frac{1+\lambda}{1+g}\right)K ,$$

with Kalman parameter K as defined above. The firm survives with probability

$$\begin{aligned}S^1(z,\hat{z},\varepsilon'+\zeta') &\equiv (1-\delta)(1-\Theta(s',\hat{z}')) \\ &= (1-\delta)\left(1-\Theta\left(z+\varepsilon'+\zeta'+\log\left(\frac{1+\lambda}{1+g}\right), \hat{z}(1-K)+zK+(\varepsilon'+\zeta')K+\log\left(\frac{1+\lambda}{1+g}\right)K\right)\right)\end{aligned}$$

For a non-innovating firm the expression is analogously derived:

$$
\begin{aligned}
S^0(z, \hat{z}, \varepsilon' + \zeta') &\equiv (1-\delta)\left(1 - \Theta(s', \hat{z}')\right) \\
&= (1-\delta)\left(1 - \Theta\left(z + \varepsilon' + \zeta' + \log\left(\frac{1}{1+g}\right), \hat{z}(1-K) + zK + (\varepsilon' + \zeta')K + \log\left(\frac{1}{1+g}\right)K\right)\right)
\end{aligned}
$$

Note: Compared to an earlier version I include exogenous exit into these expressions so that δ does not show up in the growth equation above.

Marek's derivation from here on

Let us express the aggregate productivity in terms of firm measure $\mu(s, \hat{z})$,

$$
A_t = \left(\int\int a_t(s)^{\eta-1} M\mu(s, \hat{z})\, \mathrm{d}s\, \mathrm{d}\hat{z}\right)^{\frac{1}{\eta-1}} \tag{45}
$$

Let us define the normalized level of productivity of a firm of type s by $\widehat{a}(s) \equiv \frac{a(s)}{A}$. Consider first the productivity level at an individual firm. The firm is characterized by an initial state $(s, \hat{z})$ and productivity $a_t(s)$ in period t. It transitions to $a_t(s+\epsilon+\zeta)(1+\lambda)$ with probability $x \times P(\epsilon) \times P(\zeta)$, where $P(\epsilon)$ and $P(\zeta)$ are probabilities of drawing the particular realization of ϵ and ζ, respectively. The firm transitions to $a_t(s+\epsilon+\zeta)$ with probability $(1-x(s,\hat{z})) \times P(\epsilon) \times P(\zeta)$. In deriving the joint probabilities, we used the fact that productivity shocks and innovation outcomes are are jointly independent. Finally, after receiving realization of all three shocks (innovation, transitory, persistent) the firm decides whether to pay fixed cost and continue operating or to exit. The exit choice if capture by an indicator variable $\Theta(s, \hat{z})$

Let us now explicitly write down the value of aggregate productivity index in $t+1$ relative to productivity index in t

$$
\frac{A_{t+1}}{A_t} = \left(\frac{1}{A_t^{\eta-1}}\int\int a_{t+1}(s)^{\eta-1} M\mu(s, \hat{z})\, \mathrm{d}s\, \mathrm{d}\hat{z}\right)^{\frac{1}{\eta-1}} = \left(\int\int \left(\frac{a_{t+1}(s)}{A_t}\right)^{\eta-1} M\mu(s, \hat{z})\, \mathrm{d}s\, \mathrm{d}\hat{z}\right)^{\frac{1}{\eta-1}}. \tag{46}
$$

Next, we characterize the evolution of $a_t(s)$. We need to introduce some intermediate notation to make the expressions manageable. Let $\hat{z}^+ \equiv \hat{z}^+(s, \hat{z}, \epsilon, \zeta)$ denote the updated beliefs after a successful innovation and conditional on drawing particular values ϵ and ζ of transitory and persistent shocks, respectively and conditional on current beliefs $\hat{z}$. Let $\hat{z}^-$ denote the beliefs after an unsuccessful innovation. Let Θ^+ and Θ^- correspond to exit decisions in each scenario. Again note that both depend on $s, \hat{z}_{t+1}, \epsilon, \zeta$.

$$
\mathbb{E}_t[a_{t+1} \mid a_t = a(s), \hat{z}] = \int\int [(1-\Theta^+)a(s+\epsilon+\zeta)(x(1+\lambda) + (1-\Theta^-)(1-x)]P(\epsilon)P(\zeta)\, \mathrm{d}\epsilon\, \mathrm{d}\zeta
$$

where $x \equiv x(s, \hat{z})$.

Taken together, we can write the growth of the aggregate productivity index in $t+1$ as

$$A_{t+1} = \Bigg(\int\int \Bigg[\int\int a_t(s+\epsilon+\zeta)^{\eta-1}[(1-\Theta^+)(x(1+\lambda)^{\eta-1} + (1-\Theta^-)(1-x))] P(\epsilon)P(\zeta)\,\mathrm{d}\epsilon\,\mathrm{d}\zeta \Bigg] M\mu(s,\hat{z})\,\mathrm{d}s\,\mathrm{d}\hat{z}$$
$$+ M^e \int\int\int (1-\Theta(s+\epsilon+\zeta, \hat{z}(\epsilon+\zeta)))(a_{t+1}(s)A_{t+1})^{\eta-1} P(\epsilon)P(\zeta)\,\mathrm{d}\epsilon\,\mathrm{d}\zeta G^e(s)\,\mathrm{d}s \Bigg)^{\frac{1}{\eta-1}}.$$

Using the same logic as in (46), we can express the growth rate as

$$\frac{A_{t+1}}{A_t} = \Bigg(\int\int \Bigg[\int\int \hat{a}_t(s+\epsilon+\zeta)^{\eta-1}[(1-\Theta^+)(x(1+\lambda)^{\eta-1} + (1-\Theta^-)(1-x))] P(\epsilon)P(\zeta)\,\mathrm{d}\epsilon\,\mathrm{d}\zeta \Bigg] M\mu(s,\hat{z})\,\mathrm{d}s\,\mathrm{d}\hat{z}$$
$$+ M^e \int\int\int (1-\Theta(s+\epsilon+\zeta, \hat{z}_1(\epsilon+\zeta)))(\hat{a}_{t+1}(s)(1+g))^{\eta-1} P(\epsilon)P(\zeta)\,\mathrm{d}\epsilon\,\mathrm{d}\zeta G^e(s)\,\mathrm{d}s \Bigg)^{\frac{1}{\eta-1}}. \tag{47}$$

Note that we now use normalized productivity $\hat{a}_t(s) \equiv \frac{a_t(s)}{A_t}$.

In the above, M is the equilibrium firm mass, μ is the stationary density function and where

$$M^e = M \left[\delta + (1-\delta) \int \Theta P(\epsilon)P(\zeta)\mu(s,\hat{z})\,\mathrm{d}(s,\hat{z},\epsilon,\zeta) \right] \tag{48}$$

is the relative mass of entrants equal to the relative mass of exitors.

Next, we derive the decomposition of the aggregate growth into several channels. For the sake of notation, let $\Pi(s,\hat{z}) = \int\int [(1-\Theta^+)(x(1+\lambda)^{\eta-1} + (1-\Theta^-)(1-x)) P(\epsilon)P(\zeta)\,\mathrm{d}\epsilon\,\mathrm{d}\zeta$. We can decompose the aggregate growth into

$$\begin{aligned}(1+g)^{\eta-1} = {} & M \int\int \Pi(s,\hat{z})\hat{a}_t(s)^{\eta-1}\Big(\mu(s,\hat{z}) - G^e(s,\hat{z})\Big)\,\mathrm{d}s\,\mathrm{d}\hat{z} \\ & + M \int\int \Pi(s,\hat{z})\hat{a}_t(s)^{\eta-1} G^e(s,\hat{z})\,\mathrm{d}s\,\mathrm{d}\hat{z} \\ & + M \Bigg[\underbrace{\delta + (1-\delta)\int_{s,z} \Theta(s,\hat{z})\mu(s,\hat{z})\,\mathrm{d}s\,\mathrm{d}\hat{z}}_{=M^e/M} \Bigg] \int (1-\Theta(s,\hat{z}_0))\hat{a}_{t+1}(s)^{\eta-1}(1+g)^{\eta-1} G^e(s)\,\mathrm{d}s.\end{aligned}$$

The first term accounts for distributional impact of differential firm growth rates after entry. We dub it a "reallocation channel", since it is shaped by reallocation of resources over firms' lifecycle. The second term is the contribution to growth of continuing firms under the counterfactual assumption that the share of products supplied by continuing firms of each type is the same as at entry. This is a "selection" channel, because the stationary and entry distributions differ due to the endogenous exit and innovation. The third term is the net contribution of entry and exit.

The reallocation channel can be further decomposed into an *average innovation channel* which is the average of productivity growth of incumbents and a *between-firm innovation channel* which reflects differences in innovation activity by heterogeneous firms. Formally, we can decompose

$$\mathbb{E}_m\left[P(s,\hat{z})a(s)^{\eta-1}\right] = \mathbb{E}_m\left[P(s,\hat{z})\right]\mathbb{E}_m\left[a(s)^{\eta-1}\right] + \mathrm{cov}_m\left[P(s,\hat{z}), a(s)^{\eta-1}\right], \tag{49}$$

where the expectations are calculated with respect to an arbitrary measure m. For instance, in the reallocation channel $m = \mu - G^e$. The covariance term captures whether the effective innovation probability, i.e. taking into account exit probability, is higher among more productive firms. Note that any measure m can be decomposed into impact of optimists, m^+ and pessimists m^-.

Each component can be decomposed further into the contribution of optimists and pessimists. We can decompose the firm measure into a sum $\mu(s, \hat{z}) = \mu^+(s, \hat{z}) + \mu^-(s, \hat{z})$ where $\mu^-(s, \hat{z}) = \max\{0, z(s) - \hat{z}\}$ and $\mu^+(s, \hat{z}) = \max\{0, \hat{z} - z(s)\}$. In words, μ^- is the measure of pessimists , i.e. those firms for which $z > \hat{z}$, and μ^+ is the measure of optimists.

D Illustrative example

In this section, we characterize analytically the balanced growth path equilibrium in a special case. To make the model analytically tractable, we assume that firms perfectly observe their productivity and that there is a deterministic drift in the productivity process, i.e. $\mathbb{E}[\zeta_t] = \log(1 + g_e)$ and $Var[\zeta_t] = 0$ for all t. This allows us to shed some light on the implication of firms' belief about future productivity on current R&D activity.

Formally, we assume that

Assumption 1. *There is perfect information and $\sigma_\varepsilon = 0$ and $\sigma_\zeta = 0$. The productivity evolves as $z_{j,t+1} = \log(1 + g_e) + z_{j,t} + \Lambda_{j,t} \log(1 + \lambda)$ for all firms j.*

To keep the model analytically tractable, we introduce several parametric assumptions. We let the R&D cost to be quadratic and we set the markup to 2. Finally, we consider a model with only exogenous exit at the rate δ. The assumptions are summarized in Assumption 2 below.

Assumption 2. *Assume that $\psi = 2$, $\eta = 2$, $\rho = 1$, $V^x = 0$, $\bar{r}_t = a_t$. There is only exogenous exit.*

The key assumption for analytical characterization is the assumption that $\bar{r}_t = a_t$, i.e., the cost of R&D scales linearly with the current productivity. Together with $\psi = 2$, these restrictions deliver the firm growth rate that is independent of firm size, so-called Gibrat's law. This facilitates a simple aggregation of innovation in the economy and enables a closed-form solution for the aggregate growth rate.

The balanced growth path equilibrium is summarized in Proposition 1. The aggregate growth rate of the economy is given by $g = \lambda x + g_e$. The endogenous component of aggregate growth is simply the innovation step size times the probability of a successful innovation in each firm. As indicated above, the innovation rate is constant over time and the same for all firms.

The innovation rate, however, is proportional to the expected drift in the productivity process $1 + g_e$. The expected market size is a key determinant of the gains from R&D - this is known as the *market size effect* in the economic growth literature. The higher is the expected future productivity compared to today's, the higher is the marginal benefit of investing in R&D today. The higher is the expected market size, the more resources invested in R&D, and the higher the aggregate growth rate.

In the quantitative model developed in the current paper, the expected future productivity is of course endogenous and heterogeneous across firms. However, the economic mechanism behind their impact on current innovation is captured well in this simple example.

Proposition 1. *Consider a perfect-information economy. Let Assumptions 1 and 2 hold. Then,*

1. *The optimal R&D intensity is constant and the same in all firms. It reads*

$$x = (1+g_e)\lambda \frac{\beta\,(1-\delta)}{w(1+g)}\mathcal{V},$$

where $\mathcal{V} > 0$ is implicitly defined as the positive real solution to

$$\mathcal{V}\,(1+\beta(1-\delta)\,(1+x)) - \frac{1}{4} - \frac{w}{2}\left(\lambda(1+g_e)\frac{\beta\,(1-\delta)}{w(1+g)}\mathcal{V}\right)^2 = 0$$

2. *The aggregate growth rate in the economy is*

$$g = \lambda x + g_e.$$

Proof. The profit reads $\pi = \left(\frac{\eta}{\eta-1}\right)^{1-\eta} W^{1-\eta}\exp\left((\eta-1)\,Z\right) - Wr$. With our parametric assumption on ρ and ψ, first-order condition for R&D reduces to

$$w\exp\left(Z_t\right)x_t = (1-\delta)\,\beta\left[V\left(Z_t+\ln\left(1+\lambda\right)\right) - V\left(Z_t\right)\right](1+g_e)$$

Guess that the *average* firm value function is linear in persistent productivity, i.e., $\mathbb{E}_\varepsilon V(a_t) \equiv \mathcal{V}\exp\left(Z_t\right)$ for some $\mathcal{V} > 0$. Since all shocks are mean-zero and firm is risk neutral, the stochastic components to not affect decisions. Note that this allows us to express the stationarized firm value as (lower-case symbols denote stationarized variables)

$$V(Z)/A = \mathcal{V}e^Z/A = \pi/A + \frac{\beta(1-\delta)}{1+g}(1+g_e)\left[x\mathcal{V}z(1+\lambda) + (1-x)\mathcal{V}z\right],$$

where $z_t \equiv \exp\left(Z_t\right)/A_t$ and $\exp\left(Z_{t+1}\right)/A_t = \frac{1}{1+g}z_{t+1}$. Stationarized profits are

$$\pi/A = \left(\frac{\eta}{\eta-1}\right)^{1-\eta} w^{1-\eta}z^{\eta-1}.$$

Note that the effective discount factor in a stationarized economy is $\frac{\beta(1-\delta)}{1+g}$.

Next, we verify our guess. Assuming linearity delivers

$$\mathbb{E}_t\left[V(z_{t+1})\right] = x_t\mathcal{V}z_t(1+\lambda)(1+g_e) + (1-x_t)\mathcal{V}z_t(1+g_e).$$

The FOC for R&D implies

$$x_t = \lambda(1+g_e)\frac{(1-\delta)\,\beta}{w(1+g)}\mathcal{V},$$

which implies that R&D intensity is time invariant and the same for all firms.

The stationarized firm value then becomes (imposing $\eta = 2$)

$$\mathcal{V}e^{z} = \frac{1}{4}w^{-1}e^{z}c - w\frac{1}{4}x^{2} + \beta(1-\delta)\mathcal{V}z\left(1+\lambda x\right) \tag{50}$$

This verifies our guess, because x is independent of z.

In the last step, we use the labor and goods markets clearing to solve for the aggregate wage

$$1 = \left[\int c^{1/2}\left(\mathcal{S}\right)\mu\left(\mathcal{S}\right)\right]^{2} = \frac{1}{4}W^{-2}A^{2}$$

Which implies that $W = A/2$. The labor market clearing is

$$1 = \int n(\mathcal{S})\mu(\mathcal{S})\,\mathrm{d}\mathcal{S} = \frac{1}{4}W_t^{-2}C_t\exp\left(z_t\right)\mu(z)\,\mathrm{d}z = \frac{1}{4}\frac{C}{W^2}A$$

Note further that $c = p^{-2}C$ where $p = 2\frac{W}{a}$ and hence $C = A = 2W$. This means that $w \equiv W/A = 1/2$.

E Kalman filter

E.1 Updating equation

We start with the law of motion of persistent productivity z_t:

$$z_t = \begin{cases} z_{t-1} + \log(1+\lambda) + \zeta_t & \text{if } \Lambda_t = 1 \quad \text{(innovation)}, \\ z_{t-1} + \zeta_t & \text{if } \Lambda_t = 0 \quad \text{(no innovation)}. \end{cases}$$

First, we detrend the law of motion. To this end, we subtract from both sides $\log(A_t)$, where A_t is aggregate productivity.

$$z_t - \log(A_t) = \begin{cases} z_{t-1} - \log(A_t) + \log(1+\lambda) + \zeta_t & \text{if } \Lambda_t = 1 \quad \text{(innovation)}, \\ z_{t-1} - \log(A_t) + \zeta_t & \text{if } \Lambda_t = 0 \quad \text{(no innovation)}. \end{cases}$$

Next, we add and subtract $\log(A_{t-1})$ from the right-hand side and use the definition $\frac{A_t}{A_{t-1}} = 1+g$ to re-write the right-hand side as

$$z_t - \log(A_t) = \begin{cases} z_{t-1} - \log(A_{t-1}) - \log(1+g) + \log(1+\lambda) + \zeta_t & \text{if } \Lambda_t = 1 \quad \text{(innovation)}, \\ z_{t-1} - \log(A_{t-1}) - \log(1+g) + \zeta_t & \text{if } \Lambda_t = 0 \quad \text{(no innovation)}. \end{cases}$$

We define $\tilde{z}_t = z_t - \log(A_t)$ and obtain the stationary version of the law of motion

$$\tilde{z}_t = \begin{cases} \tilde{z}_{t-1} - \log(1+g) + \log(1+\lambda) + \zeta_t & \text{if } \Lambda_t = 1 \quad \text{(innovation)}, \\ \tilde{z}_{t-1} - \log(1+g) + \zeta_t & \text{if } \Lambda_t = 0 \quad \text{(no innovation)}. \end{cases}$$

We abuse notation and use from now on the variables without tilde to denote the stationarized variables.

To derive the Kalman filter, we first note that the state equation is given by

$$z_t = \begin{cases} z_{t-1} - \log(1+g) + \log(1+\lambda) + \zeta_t & \text{if } \Lambda_t = 1 \quad \text{(innovation)}, \\ z_{t-1} - \log(1+g) + \zeta_t & \text{if } \Lambda_t = 0 \quad \text{(no innovation)}. \end{cases}$$

and the observation equation is given by

$$s_t = z_t + \varepsilon_t \ .$$

Let by $\mathbb{E}_{t-1}(z_t)$ denote the prediction of the unobserved state z_t in period $t-1$. Given the state equation, we can write $\mathbb{E}_{t-1}(z_t)$ as

$$\mathbb{E}_{t-1}(z_t) = \begin{cases} \mathbb{E}_{t-1}(z_{t-1}) - \log(1+g) + \log(1+\lambda) & \text{if } \Lambda_t = 1 \quad \text{(innovation)}, \\ \mathbb{E}_{t-1}(z_{t-1}) - \log(1+g) & \text{if } \Lambda_t = 0 \quad \text{(no innovation)}. \end{cases}$$

Let by $\mathbb{E}_t(z_t)$ denote the update of the prediction in period t, i.e. when new information in the form of a new observation of s_t has arrived. The generic updating equation is given by

$$\mathbb{E}_t(z_t) = \mathbb{E}_{t-1}(z_t) + \mathcal{K}_t\Big(s_t - \mathbb{E}_{t-1}(z_t)\Big)$$

where $\mathcal{K}_t$ is the Kalman gain. Using the previous expression for $\mathbb{E}_{t-1}(z_t)$ we can write the updating equation as

$$\mathbb{E}_t(z_t) = \begin{cases} \mathbb{E}_{t-1}(z_{t-1}) - \log(1+g) + \log(1+\lambda) + \mathcal{K}_t\Big(s_t - \mathbb{E}_{t-1}(z_{t-1}) + \log(1+g) - \log(1+\lambda)\Big) & \text{if } \Lambda_t = 1 \ , \\ \mathbb{E}_{t-1}(z_{t-1}) - \log(1+g) + \mathcal{K}_t\Big(s_t - \mathbb{E}_{t-1}(z_{t-1}) + \log(1+g)\Big) & \text{if } \Lambda_t = 0 \ . \end{cases}$$

Using our notation from the main text $\mathbb{E}_t(z_t) = \hat{z}_t$ and $\mathcal{K}_t = \frac{k_t}{k_t+\sigma_\epsilon^2}$, this expression reads

$$\hat{z}_t = \begin{cases} \hat{z}_{t-1} - \log(1+g) + \log(1+\lambda) + \frac{k_t}{k_t+\sigma_\epsilon^2}\Big(s_t - \hat{z}_{t-1} + \log(1+g) - \log(1+\lambda)\Big) & \text{if } \Lambda_t = 1 \ , \\ \hat{z}_{t-1} - \log(1+g) + \frac{k_t}{k_t+\sigma_\epsilon^2}\Big(s_t - \hat{z}_{t-1} + \log(1+g)\Big) & \text{if } \Lambda_t = 0. \end{cases}$$

E.2 Conditional variance $k_{t|t}$

Let $k_t = var(z_t \mid s_1, s_2, ..., s_{t-1})$ be the variance of z_t, conditional on information available until period $t-1$. Let $k_{t|t} = var(z_t \mid s_1, s_2, ..., s_t)$ denote the variance of z_t, conditional on information available until period t. The conditional variance $k_{t|t}$ is given by

$$k_{t|t} = \frac{k_t\sigma_\epsilon^2}{k_t + \sigma_\epsilon^2}$$

and the update of the variance k_t

$$k_{t+1} = k_{t|t} + \sigma_\zeta^2$$

We can express the above equation in terms of Kalman gain, $\mathcal{K}_t = \frac{k_t}{k_t+\sigma_\epsilon^2} = \frac{var(z_t|s_{t-1})}{var(z_t+\epsilon_t-\hat{z}_t|s_{t-1})}$. Using the expression for the Kalman gain in the previous expression, we obtain

$$\mathbb{E}[z_{t+1} \mid s_t] = \mathbb{E}[z_t \mid s_t] = \mathbb{E}[z_t \mid s_{t-1}] + \mathcal{K}_t(s_t - \mathbb{E}[z_t \mid s_{t-1}]) \tag{51}$$

$$k_{t|t} = k_t(1 - \mathcal{K}_t) \tag{52}$$

$$k_{t+1} = k_t(1 - \mathcal{K}_t) + \sigma_\zeta^2 \tag{53}$$

F Perfect Information Model

To analyze the role of imperfect information, we contrast our model with an economy where firms are perfectly informed about their persistent productivity process. We derive the main model equations for a stationary equilibrium (balanced growth path). The equilibrium conditions on the household side and the market-clearing conditions are the same. Therefore, we only describe the conditions characterizing firm decisions and firm dynamics and the stationary firm distribution.

F.1 Firms' Problem

The relevant state vector of a firm in the perfect-information model is $\mathcal{S} = (s, z)$ where z is the (observed) persistent productivity component and $s = z + \varepsilon$ is current productivity. Both are expressed in the stationary representation of the model, so that z evolves as in equation (19). The firm's decisions about exiting, pricing, employment in production and R&D labor, take the same form as in the benchmark model, hence are described by Bellman equation (17) whose first-order condition for the choice of the innovation probability is given by equation (18). The only difference to the benchmark model is that the expectation operators are formed over the realization of the future state variables (s, z) rather than $(s, \hat{z})$ where z evolves according to equation equation (19). Likewise, the firm's exit policy function is denoted $\Theta(s, z) \in \{0, 1\}$, and the firm's choice of production labor is given by equation (22).

As in the benchmark model, entrant firms draw initial productivity z_1 from distribution G^e and transitory shock ε_1 so that initial productivity is $s_1 = z_1 + \varepsilon_1$ and the firm decides to stay if $\Theta(s_1, z_1) = 0$. The free-entry condition is again given by (14).

F.2 Stationary Firm Distribution

In the perfect-information model, the measure of active firms is denoted $\mu(z)$ with single state variable z. The law of motion for μ is similar to equation (20) written as

$$d\mu_{t+1}(z_{t+1}) = (1 - \delta) \int F(z_{t+1}|z_t)\ d\mu_t(z_t) + F^e(z_{t+1}) \times M^e_{t+1}\ ,$$

where $F(z_{t+1}|z_t)$ is probability measure that a firm has persistent productivity component z_{t+1} and stays active in $t + 1$, conditional on being in state z_t in period t, and $F^e(z_{t+1})$ is the probability measure that an entrant begins with persistent productivity component z_{t+1} and decides to stay. These measures are derived as follows.

Consider first entrant firms in period $t+1$ which draw the persistent component of productivity z_{t+1} from a normal distribution with mean μ_e and variance σ_e^2, as well as the initial transitory productivity shock ε_{t+1}. As a result, the joint probability to draw z_{t+1} and to stay in the market is

$$F^e(z_{t+1}) = \int 1 - \Theta\left(z_{t+1} + \varepsilon_{t+1}, z_{t+1}\right) d\Phi(\varepsilon_{t+1}; 0, \sigma_\varepsilon^2) \times \phi(z_{t+1}; \mu_e, \sigma_e^2) \ .$$

Next consider continuing firms for which we split the probability that a firm with persistent productivty z_t in period t enters period $t+1$ with z_{t+1} and decides to stay into two terms reflecting the outcome with and without a successful innovation:

$$F(z_{t+1}|z_t) = F(z_{t+1}|\Lambda_t = 1, z_t) \times \text{Prob}(\Lambda_t = 1|z_t) + F(z_{t+1}|\Lambda_t = 0, z_t,) \times (1 - \text{Prob}(\Lambda_t = 1|z_t)) \ .$$

The innovation probability is a function of the firm's state, $x(z_t + \varepsilon_t, z_t)$. Taking into account that innovation is conditional on firms being active in period t, we calculate the innovation probability for firms with state z_t as

$$\text{Prob}(\Lambda_t = 1|z_t) = \frac{\int x\big(z_t + \varepsilon, z_t\big)\big(1 - \Theta(z_t + \varepsilon, z_t)\big)\phi\big(\varepsilon; 0, \sigma_\varepsilon^2\big)d\varepsilon}{\int \big(1 - \Theta(z_t + \varepsilon, z_t)\big)\phi\big(\varepsilon; 0, \sigma_\varepsilon^2\big)d\varepsilon} \ .$$

With a successful innovation, the persistent productivity component evolves according to

$$z_{t+1} = z_t + \log\left(\tfrac{1+\lambda}{1+g}\right) + \zeta_{t+1} \ ,$$

so that the probability of the realization z_{t+1} is

$$P(z_{t+1}|\Lambda_t = 1, z_t) \equiv \phi\left(z_{t+1}; z_t + \log\left(\tfrac{1+\lambda}{1+g}\right), \sigma_\zeta^2\right) \ .$$

Therefore, the probability that a firm in state z_t and with a successful innovation enters next period with z_{t+1} and decides to stay is

$$F(z_{t+1}|\Lambda_t = 1, z_t) = \int 1 - \Theta(z_{t+1} + \varepsilon, z_{t+1})\Phi(\varepsilon; 0, \sigma_\varepsilon^2)d\varepsilon \times P(z_{t+1}|\Lambda_t = 1, z_t) \ .$$

For those firms without a successful innovation, we obtain the similar expressions with terms $1 + \lambda$ replaced by 1:

$$P(z_{t+1}|\Lambda_t = 0, z_t) \equiv \phi\left(z_{t+1}; z_t + \log\left(\tfrac{1}{1+g}\right), \sigma_\zeta^2\right) \ ,$$

$$F(z_{t+1}|\Lambda_t = 0, z_t) = \int 1 - \Theta(z_{t+1} + \varepsilon, z_{t+1})\Phi(\varepsilon; 0, \sigma_\varepsilon^2)d\varepsilon \times P(z_{t+1}|\Lambda_t = 0, z_t) \ .$$