PRODUCTIVITY, DEMAND AND GROWTH*

Marek Ignaszak
Goethe University Frankfurt

Petr Sedláček
University of Oxford & CEPR

July 3, 2021

Abstract

Recent empirical evidence suggests that firm survival and growth are largely demand-driven. We integrate these findings into an endogenous growth model in which heterogeneous firms survive and innovate based on not only productivity, but also demand. We show analytically that firm-level demand variation impacts aggregate growth by changing firms' incentives to innovate. Using U.S. Census firm data, we estimate that 20% of aggregate growth is demand-driven. Moreover, accounting for firm-level demand variation substantially alters the economy's responses to growth policies. Finally, our model mechanism finds empirical support in firm-level data.

Keywords: Demand, Firm Heterogeneity, Growth, Innovation, R&D

JEL Codes: D21, E24, L1, O31, O33

*First version January 2021. Earlier versions were circulated under the title "Demand-Driven Growth". We thank Steve Bond, Vasco Carvalho, Sergio de Ferra, Hugo Hopenhayn, Maarten De Ridder, Georg Duernecker, Giammario Impullitti, Leo Kaas, Benjamin Moll, Lukasz Rachel, Federica Romei, Moritz Schularick, Emily Swift, Francesco Zanetti and seminar participants at Cambridge University, European Commission and ECFIN, Goethe University Frankfurt, London Business School, University of Copenhagen and the University of Oxford for useful comments. Ignaszak gratefully acknowledges research support by Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy –EXC 2126/1– 390838866 and the Research Training Group 2281. Sedláček gratefully acknowledges the financial support of the European Commission: this project has received funding from the European Research Council [grant number 802145]. Ignaszak: ignaszak@uni-frankfurt.de; Sedláček: petr.sedlacek@economics.ox.ac.uk.

1 Introduction

On average, more than half of U.S. business startups shut down within the first 5 years of existence, while those that survive almost double in size.[1] These "up-or-out" dynamics, and associated job turnover, have been linked to productivity-enhancing reallocation of resources (see e.g. Haltiwanger, 2012). Modern growth theory embraces such patterns and explicitly models the entry, growth and exit of individual firms which endogenously innovate on their heterogeneous productivity levels (see e.g. Acemoglu et al., 2018). The resulting productivity-driven process of firm survival and growth is the key factor behind advances in aggregate productivity.

However, an increasing body of empirical evidence suggests that firm-level survival and growth are also strongly influenced by demand-side factors, rather than productivity alone (see e.g. Foster et al., 2008, 2016; Hottman et al., 2016; Kehrig and Vincent, 2021).[2] These findings challenge our current understanding of the sources of aggregate economic growth. To what extent is aggregate economic growth also demand-driven as a consequence of this firm-level evidence? And does the effectiveness of growth policies change, once demand-side factors at the firm-level are taken into account? In this paper, we address these and other questions.

To do so, we develop a novel model of endogenous growth by heterogeneous firms which builds on existing endogenous growth theory (see e.g. Aghion and Howitt, 1992; Grossman and Helpman, 1991; Klette and Kortum, 2004; Acemoglu et al., 2018), but extends it by allowing firm-level outcomes to be affected by variation in demand (see e.g. Arkolakis, 2016; Sedláček and Sterk, 2017; Sterk ⓡ al., 2021). As we explain below, the interplay between demand and supply factors at the firm level is critical for understanding the sources of aggregate growth and its sensitivity to policy interventions.

In our model, businesses produce differentiated consumption goods and invest into research and development (R&D). Successful innovations lead to improvements in firm-level productivity, allowing businesses to produce at lower prices and, in turn, to attract more demand for their goods. At the same time, firms also face stochastic idiosyncratic changes in demand (firm-level market size) which are unrelated to production efficiency.[3]

A key advantage of our framework is that we can analytically characterize several model outcomes. Our theoretical analysis delivers two main takeaway messages. First, demand growth at the firm-level increases incentives to conduct R&D. This is because

[1] These values are computed using information job creation, destruction and employment from the Business Dynamic Statistics of the U.S. Census Bureau. Haltiwanger et al. (2016) document that in fact the majority of young businesses do not grow and only a small share of high-growth firms, "gazelles", is responsible for the observed average growth of young firms. Moreover, Decker et al. (2017) document that the average pace of business dynamism has slowed in recent decades.

[2] Similar results have also been found in other countries (see e.g. Bachmann and Zorn, 2020; Eslava and Haltiwanger, 2021; Bernard et al., forthcoming).

[3] Appendix D extends the baseline model to allow for endogenous demand accumulation and shows that our baseline results become slightly stronger.

firms are able to reap larger benefits from successful innovations if the demand for their product expands. This mechanism is akin to the "aggregate market size effect" identified in earlier models of endogenous growth (see e.g. Jones, 1995). However, in our framework this effect occurs at the firm-level, consistent with a range of empirical studies (see e.g. Acemoglu and Linn, 2004; Jaravel, 2019; Aghion et al., 2020).

Second, we show analytically that higher demand growth at the firm level raises *aggregate* economic growth. Note that in our economy firm-level demand growth is a zero-sum game with aggregate demand being fixed. Consequently, higher demand growth for some firms necessarily comes at the expense of declining demand growth among other businesses or even firm exit. Therefore, changes in firm-level demand impact aggregate growth only indirectly through their effect on innovation incentives.

Next, we take our model to the data. Towards this end, we first extend our framework to include endogenous firm entry and exit, allowing for rich up-or-out business dynamics. We then discipline our model by making it match a range of empirical moments based on aggregate information and firm data from the U.S. Census Bureau.

A key ingredient in our parametrization strategy is the autocovariance structure of firm-level log-employment.[4] In particular, we show analytically that a large class of endogenous growth models – those which imply firm-level employment growth to follow a random walk with drift – are inconsistent with the observed autocovariance structure.[5] In contrast, our model – in which both productivity and demand are allowed to vary at the firm-level – can match the autocovariance structure of log-employment well. Given that firm dynamics lie at the heart of modern models of endogenous growth, this is key for the quantitative analysis of aggregate growth.

In addition, our framework has at least four advantages.[6] First, it is firmly grounded in the current literature as it combines features of existing endogenous growth models with those of demand-driven firm dynamics (see e.g. Klette and Kortum, 2004; Sterk ⓡ al., 2021). Second, it conforms with research on the determinants of firm-level R&D, which suggests an important role for market power and frictions in expanding market size (see e.g. Acemoglu and Linn, 2004; De Ridder, 2020). Third, it is consistent with recent empirical evidence on the importance of demand-side factors for firm selection and growth (see e.g. Foster et al., 2008; Hottman et al., 2016). Finally, replicating the analysis in Foster et al. (2008), who estimate productivity and demand shocks using detailed firm-level information on quantities and prices, we show that the productivity and demand

[4] Sterk ⓡ al. (2021) argue that the autocovariance structure of firm-level employment is crucial for disciplining firm-level heterogeneity.

[5] A random walk process with drift implies a zero covariances between the level and the growth rate of employment, i.e. $\text{cov}(\ln n_a, \ln n_a - \ln n_{a-h}) = 0$ for $h \in (0, a]$. This is counterfactual to the data, which features strongly decreasing covariances, i.e. $\text{cov}(\ln n_a, \ln n_a - \ln n_{a-h}) < 0$ for $h \in (0, a]$.

[6] Other extensions may, in principle, also reconcile the model with the data. For instance, Akcigit and Kerr (2018) stress the departure from Gibrat's law as one of their central results and we discuss such and other extensions in Section 3.2.

dynamics implied by our model are very close to those estimated in the data.

We use our quantitative model to address three sets of questions. First, we analyze the role of demand variation for firm-level outcomes. Towards this end, we consider a counterfactual economy which is identical to our baseline model, but in which firms assume their idiosyncratic demand is fixed. Comparing the outcomes in our baseline model to those of the counterfactual economy allows us to isolate the impact of idiosyncratic demand variation on firms' choices. We find that demand variation plays a dominant role in determining up-or-out dynamics, consistent with evidence in e.g. Foster et al. (2008); Hottman et al. (2016); Eslava and Haltiwanger (2021); Bernard et al. (forthcoming). In addition, expected demand growth is also crucial for firm-level R&D decisions, consistent with existing empirical studies (see e.g. Acemoglu and Linn, 2004; Jaravel, 2019; Aghion et al., 2020).

Second, we turn our attention to the impact of firm-level demand variation for aggregate economic growth. Using our counterfactual economy, we quantify that 20 percent of aggregate economic growth is in fact demand-driven. Since our economy does not feature market size effects at the aggregate level, the contribution of demand to aggregate growth is entirely driven by it's influence on the incentives of individual businesses to conduct R&D.

Third, we show that ignoring firm-level demand variation can fundamentally change the aggregate impact of growth policies predicted by the model. To do so, we compare results from our baseline economy to a restricted version of our model which assumes firm-level demand to be common and constant across all businesses, as is common in existing models of endogenous growth.[7] We consider two examples of growth policies: subsidies to the costs of operation and to R&D. We find that these policies have quantitatively very different effects in the two versions of the model.

The key difference between the baseline and the restricted economy lies in the firm selection process. In the restricted model, by assumption, firm selection occurs only on productivity. Therefore, policies which impact firm selection necessarily affect firm-level, and in turn, aggregate productivity. In contrast, firms shut down based on profitability (i.e. a combination of productivity and demand) in the baseline economy. Moreover, in the baseline model firm selection occurs to a large extent on demand-side factors, consistent with empirical evidence (see e.g. Foster et al., 2008; Eslava and Haltiwanger, 2021). Therefore, the quantitative impact of growth policies in our baseline framework will differ from that in the restricted model depending on the precise effect such policies have on the *joint distribution* of productivity and demand at the firm-level.

These policy experiments highlight that explicitly accounting for both productivity

[7] We parametrize the restricted model to match the same targets as the baseline economy, with the exception of the autocovariance structure which we show that it cannot match. In both cases, we consider long-run general equilibrium changes, ignoring the respective transition paths.

and demand variation at the firm-level is crucial for quantitative analysis of economic growth. Moreover, our framework also opens up the possibility that short-run macroeconomic demand shocks, but also monetary or fiscal stabilization policies, can have long-run effects on productivity. This could potentially link transitory fluctuations to the medium- to long-run and help explain the persistent effects of business cycles (see e.g. Anzoategui et al., 2019; Beraja and Wolf, 2021; Fornaro and Wolf, 2021; Jorda et al., 2021).

As a final step in our analysis, we provide empirical evidence for our key channel – a link between expected demand growth and productivity growth at the firm-level. We first review existing studies documenting that firm-level market-size expansions influence innovation decisions (see e.g. Acemoglu and Linn, 2004; Aghion et al., 2020). Next, we turn to firm-level data from Compustat and estimate this relationship directly. Specifically, we follow the procedure in Foster et al. (2008) and methodology of Levinsohn and Petrin (2003) to estimate firm-level (revenue-based) TFP and demand shocks. We then regress realized productivity growth on expected demand growth at the firm-level, both in the data and our baseline model.[8] The estimates suggest that our key model mechanism is not only present in the data, but also quantitatively very similar to that predicted by the model. These results, therefore, further validate our model and the associated quantitative analysis.

Literature overview. Our paper is related to a number of different strands of research. First, we build on the literature of firm dynamics, which highlights the importance of firm-level heterogeneity for understanding aggregate patterns (see e.g. Jovanovic, 1982; Hopenhayn and Rogerson, 1993; Lee and Mukoyama, 2015; Clementi and Palazzo, 2016) and in particular to those papers which focus on demand-side factors at the firm-level (see e.g. Gourio and Rudanko, 2014; Arkolakis, 2016; Sedláček and Sterk, 2017; Perla, 2019; Sterk ⓡ al., 2021; Kaas and Kimasa, forthcoming). We add to these papers the context of aggregate growth.

In doing so, we bridge the literature on firm dynamics with general equilibrium models of endogenous innovation by heterogeneous firms which, however, ignore demand side factors at the firm level (see e.g. Klette and Kortum, 2004; Lentz and Mortensen, 2008a; Akcigit and Kerr, 2018; Acemoglu et al., 2018; Mukoyama and Osotimehin, 2019; De Ridder, 2020). Exceptions are Cavenaile and Roldan-Blanco (2021) and Rachel (2021), where the former extends the model of Akcigit and Kerr (2018) by assuming that firms can raise (perceived) product quality either through innovation or by static advertising decisions.[9] Rachel (2021) studies how firms build brand equity by providing free leisure-enhancing technologies and how this interacts with innovation and productivity growth.

[8] Expected demand growth is instrumented using current-period demand growth and a range of observable characteristics.

[9] Cavenaile et al. (2021) further consider how the interplay between intrinsic quality (affected by innovation) and extrinsic quality (affected by advertising) impacts market concentration and welfare.

In contrast to these papers, we focus on a broader definition of demand factors which encompass any force that impacts sales but is unrelated to productivity.[10] Moreover, we stress that firms' decisions crucially depend on dynamic heterogeneous demand growth profiles, grounded in models of firm dynamics and consistent with existing empirical evidence on the role of changes in firm-level market size for innovation incentives discussed above.

Finally, our paper also relates to research studying the link between growth expectations and aggregate demand (see e.g. Blanchard et al., 2017; Benigno and Fornaro, 2018). Our framework differs from these studies in its focus on firm-level demand as a driver of aggregate growth and in the absence of wage rigidities, the zero lower bound or multiple equilibria as forces connecting the two.

Paper outline. The remainder of the paper is structured as follows. The next Section describes our theoretical framework and provides an analytical characterization of some of its key features. Section 3 parametrizes the model and shows how it fits the data. Section 4 provides our quantitative results and the final Section concludes.

2 Theoretical framework

This section builds a novel general equilibrium model of endogenous growth with heterogeneous firms. The key distinction from the rest of the literature is that we allow firms' profitability to depend not only on their firm-specific productivity levels, but also on other factors. In particular, we decide to explicitly model demand for firms' products. This is consistent with recent empirical evidence suggesting that demand variation is key in determining firm performance (see e.g. Foster et al., 2008; Hottman et al., 2016; Eslava and Haltiwanger, 2021). In general, however, similar conclusions could be made about other factors which are independent of productivity but which affect profitability. Such factors could include entrepreneurial learning, the expansion of production or customer networks or the development of long-term business relationships (see e.g. Stein, 1997; Foster et al., 2016; Fujii et al., 2017).

After we present our model, we derive analytical characterizations of model outcomes for both the decentralized economy and for the allocation of a social planner. These provide intuition for why profitability, as opposed to productivity alone, is key for understanding innovation at the firm-level and, in turn, aggregate economic growth.

[10] Comin et al. (2021) consider non-homothetic preferences in a multi-sector growth model and study the role of demand and supply in driving structural change in developing economies.

2.1 The model

This subsection introduces our new structural model and characterizes the stationary balanced growth equilibrium.

Consumers. We assume a representative household which consists of workers who supply labor, consume goods varieties and invest into assets (firms). The utility function of the representative household has the following form:

$$\mathbb{U} = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\ln C_t - \upsilon N_t\right], \tag{1}$$

where $\beta \in (0,1)$ is the discount factor, C_t is consumption and N_t is aggregate labor supply in period t. As in Hansen (1985) and Rogerson (1988), we adopt the indivisible-labor formulation in which N_t represents the fraction of workers who are employed.

Aggregate consumption, C_t, consists of a combination of individual goods varieties, $c_{j,t}$. Preferences over these goods varieties are described as:

$$C_t = \left[\int_{j\in\Omega_t} (b_{j,t})^{\frac{1}{\eta}} c_{j,t}^{\frac{\eta-1}{\eta}} dj\right]^{\frac{\eta}{\eta-1}}, \tag{2}$$

where Ω_t is the mass of producers, η is the elasticity of substitution between goods varieties and $b_{j,t}$ is a utility (demand) weight of good j.[11] While time-varying demand weights can be micro-founded through a process of customer acquisition or product awareness (see e.g. Gourio and Rudanko, 2014; Sedláček and Sterk, 2017; Perla, 2019), here we model them as exogenous for the sake of simplicity. In particular, we assume

$$b_{j,t+1} = \theta_{j,t} b_{j,t}, \tag{3}$$

where $\theta_{j,t} > 0$ is goods-specific and potentially time-varying and stochastic demand growth and where initial values of demand $b_{j,0}$ are given. Note that this specification allows for both increases and decreases in the level of demand for good j. For convenience, we assume that aggregate demand, $B = \int_j b_j dj$, is stationary. Therefore, changes in firm-specific demand occur within a "zero-sum game", whereby products which enjoy increases in demand do so at the expense of other products for which demand declines.

The representative household maximizes (1) subject to the following budget constraint:

$$A_{t+1} + \int_{j\in\Omega_t} p_{j,t} c_{j,t} = W_t N_t + (1+r_t) A_t, \tag{4}$$

where A_t are total assets, $p_{j,t}$ are variety-specific prices relative to the aggregate price

[11] In what follows, we will use the terms utility weight and demand interchangeably.

index P_t which we take as the numeraire and normalize to 1, W_t is the wage and r_t is the interest rate. Given that the representative household owns all the firms, the asset market clearing condition implies

$$A_t = \int_{j \in \Omega_t} V_{j,t} dj, \tag{5}$$

where $V_{j,t}$ is the value of a firm producing goods variety j in period t. The optimality conditions of the household can be written as:

$$1 = \beta \mathbb{E} \frac{C_t}{C_{t+1}} (1 + r_{t+1}), \tag{6}$$

$$W_t = \upsilon C_t, \tag{7}$$

$$c_{j,t} = b_{j,t} p_{j,t}^{-\eta} C_t. \tag{8}$$

The conditions above constitute, respectively, the Euler equation (6), optimal labor supply (7) and the demand for individual goods varieties (8).

Incumbent firms. Goods variety are produced by heterogeneous firms using labor supplied by the household. In addition to using labor in production, businesses can also hire workers to conduct research and development (R&D) allowing them to increase their production efficiency, $q_{j,t}$.

We assume that consumption varieties are produced using the following linear technology:

$$c_{j,t} = q_{j,t} n_{j,t}^c, \tag{9}$$

where $n_{j,t}^c$ is labor used in production at firm j in period t. In order to improve their production efficiency, firms can hire $n_{j,t}^r$ workers to conduct R&D, yielding an innovation probability of $x_{j,t}$. If successful, innovations lead to an increase in production efficiency by a factor of $(1 + \lambda)$, where $\lambda > 0$. We assume that R&D costs are given by:

$$n_{j,t}^r = \frac{p_{j,t} c_{j,t}}{P_t C_t} \frac{x_{j,t}^{\psi}}{\gamma}, \tag{10}$$

where $\psi > 1$ and γ and constants. Note that R&D costs are proportional to the firm-specific sales share. First, this implies that a relatively more sought after product is more expensive to innovate on further. Similar specifications, albeit purely productivity-driven, are common in the literature (see e.g. Akcigit and Kerr, 2018; Acemoglu et al., 2018). Second, as will become clear below, this specification allows us to analytically characterize the outcomes of our model.

Before describing the firm's optimization problem, let us note that a firm's idiosyncratic state is given by its production efficiency, $q_{j,t}$, its demand level, $b_{j,t}$ and the growth of idiosyncratic demand, $\theta_{j,t}$. To lighten the exposition, we group these into a vector of firm-specific state variables $s_{j,t} = (q_{j,t}, b_{j,t}, \theta_{j,t})$. In addition, let us define next period's state – depending on the outcome of the innovation process – as $s^+_{j,t} = (q_{j,t}(1+\lambda), \theta_{j,t}b_{j,t}, \theta_{j,t+1})$ and $s^-_{j,t} = (q_{j,t}, \theta_{j,t}b_{j,t}, \theta_{j,t+1})$.

We are now ready to describe the value of a firm producing good j as

$$V(s_{j,t}) = \max_{p_{j,t}, n^c_{j,t}, n^r_{j,t}} \left\{ \begin{array}{c} p_{j,t}c_{j,t} - W_t(n^c_{j,t} + n^r_{j,t}) + \\ \mathbb{E}\beta_t(1-\delta)\left[x_{j,t}V(s^+_{j,t+1}) + (1-x_{j,t})V(s^-_{j,t+1})\right] \end{array} \right\}, \tag{11}$$

where δ is an exogenous firm exit rate and where $\beta_t = \beta\mathbb{E}C_t/C_{t+1} = 1/(1+r_{t+1})$ is the household's stochastic discount factor. The optimality conditions of an incumbent firm can then be written as:[12]

$$p_{j,t} = \frac{\eta}{\eta-1}\frac{W_t}{q_{j,t}}, \tag{12}$$

$$\psi W_t \frac{p_{j,t}c_{j,t}}{P_tC_t}\frac{x^{\psi-1}_{j,t}}{\gamma} = \mathbb{E}\beta_t(1-\delta)\left[V(s^+_{j,t+1}) - V(s^-_{j,t+1})\right]. \tag{13}$$

The conditions above describe optimal pricing as a constant markup over marginal costs (12) and the optimal innovation investment as a balance between marginal costs and benefits (13). As is common in models of endogenous growth, the latter depends on the change in firm value brought about by a successful innovation.

Finally, notice that our setting incorporates *two* sources of firm-level growth. First, a higher production efficiency enables firms to produce their goods at a lower price. This, in turn, attracts higher demand from the side of the household (8) enabling the firm to expand. Second, the demand for a firm's good j is also governed by the household's demand weight, $b_{j,t}$, which evolves over time independently of a firm's production efficiency. Importantly, however, we will show that demand variation interacts with productivity growth because it alters the incentives to conduct R&D.

Entrants. Every period, a unit mass of potential startups has the option to enter the economy. In order to do so, they must first pay a fixed entry cost κ (denoted in labor units). It is assumed that, upon paying the entry cost, startups obtain a random draw of the idiosyncratic state, s_e, comprising of production efficiency, $q_{e,t}$, assumed to be proportional to last period's aggregate productivity (Q_{t-1}) defined below.[13] In addition, entrants also draw an initial demand level, $b_{j,0}$ and its life-cycle growth profile, $\{\theta_{j,s}\}^{\infty}_{s=0}$.

[12] Note that with $P = 1$, $p_jc_j/(PC) = bq_j^{\eta-1}/Q^{\eta-1}$.

[13] This setup is characterized by entrants "standing on the shoulders of giants" since aggregate production efficiency determines their initial productivity draws as is common in the literature (see e.g. Akcigit and Kerr, 2018).

These initial draws are identically and independently distributed across firms and time, described by a cumulative distribution function $H(s_e)$. After the realization of the initial draws, entrants decide on investment into R&D, x_e.

The free entry condition is given by

$$\kappa W_t = \int_{s_e} \left\{ \begin{array}{c} -W_t \frac{p_{e,t}c_{e,t}}{P_t C_t} \frac{x_{e,t}^{\psi}}{\gamma} + \\ \mathbb{E}\beta_t (1-\delta) \left[x_{e,t} V(s_{e,t+1}^{+}) + (1-x_{e,t}) V(s_{e,t+1}^{-}) \right] \end{array} \right\} dH(s_e), \tag{14}$$

where $p_{e,t}c_{e,t} = b_{e,t}\frac{\eta}{\eta-1}\left(W/q_{e,t}\right)^{1-\eta} C_t$.[14]

The associated optimal entrant innovation probability, x_e, is then defined by the following optimality condition which mirrors that of incumbent businesses:

$$W_t \psi \frac{p_{e,t}c_{e,t}}{P_t C_t} \frac{x_e^{\psi-1}}{\gamma} = \mathbb{E}\beta_t (1-\delta) \left[V(s_{e,t+1}^{+}) - V(s_{e,t+1}^{-}) \right]. \tag{15}$$

Balanced growth equilibrium. To close the model, we define labor market clearing, the law of motion for the mass of firms and aggregate economic growth. We can then define the balanced growth path equilibrium of our economy.

Labor market clearing simply requires that all labor demanded by firms (for production, R&D and entry costs) is supplied by the household

$$N_t = \int_{j \in \Omega_t} (n_{j,t}^c + n_{j,t}^r) dj + \kappa M_t. \tag{16}$$

The mass of firms evolves according to the following law of motion

$$\Omega_{t+1} = \underbrace{(1-\delta)\Omega_t}_{\text{surviving incumbents}} + \underbrace{(1-\delta)M_t}_{\text{new entrants}}, \tag{17}$$

where M is the mass of entrants determined by free entry (14).

Finally, we turn to defining aggregate economic growth. We focus on the balanced growth path (BGP) of the economy, along which all growing variables grow at the same rate $1+g$ given by:

$$1 + g = \frac{Q_{t+1}}{Q_t}, \tag{18}$$

[14] Note that while $p_{e,t}c_{e,t}$ enters the R&D cost of potential entrants, they only produce consumption goods in the next period following payment of the entry cost. Therefore, $p_{e,t}c_{e,t}$ serves merely as a scaling parameter in their R&D cost function, symmetric to incumbent businesses.

where Q_t is an aggregate productivity index, defined as

$$Q_t \equiv \frac{1}{B}\left(\int_{j\in\Omega_t} b_{j,t} q_{j,t}^{\eta-1} dj\right)^{\frac{1}{\eta-1}}. \tag{19}$$

Intuitively, our economy grows at the pace of (demand) weighted average firm-level productivity growth, adjusted for the elasticity of substitution in consumption.[15]

Therefore, along the BGP we can stationarize our economy by dividing all growing variables by Q. In what follows, we denote stationarized variables with "hats", e.g. $\hat{C} = C/Q$.

Definition 1 (Balanced Growth Path Equilibrium). *A balanced growth path equilibrium of our model consists of the following tuple in every period t with $j \in \Omega$: b_j, c_j, p_j, x_j, x_e, n_j^c, n_j^r, $V(s_j)$, r, W, C, A, M, Ω, Q, g, such that (i) demand, output and prices, b_j, c_j and p_j, satisfy (3), (8) and (12), (ii) optimal innovation probabilities of incumbents and entrants, x_j and x_e, satisfy (13) and (15), (iii) labor demand, n_j^c and n_j^r, satisfy (9) and (10), (iv) firm values, $V(s_j)$, satisfy (11), (v) the interest and wage rates, r and W, satisfy (6) and (16), (vi) aggregate consumption and assets, C and A, are defined by (2) and (5), (vii) the mass of entrants and firms, M and Ω, satisfy (14) and (17), (viii) the aggregate productivity index and its growth, Q and g, are defined by (19) and (18).*

2.2 Analytical characterization of the decentralized economy

Having laid out our baseline economy, we now turn to analytically characterizing some of its properties. We focus on the decentralized economy, but we also compare some of the model's predictions to those implied by a social planner. This comparison highlights the key interaction between demand and productivity growth at the firm level. We defer all the proofs to the Appendix.

Parametric assumption, linear firm value and the innovation rate. We start by a particular parametric assumption enabling us to derive closed form solutions to the firms' optimality problem. We relax this assumption and extend our model along several dimensions in the next section where we take our framework to the data.

Assumption 1. *Assume (i) parameter values $\eta = \psi = 2$ and (ii) constant and common demand growth across firms and time, $\theta_{j,t} = \theta$ and $b_{j,0} = b_0$ for all j and t.*

[15] Note that while we assume aggregate demand to be fixed, i.e. $B_t = B$ for all t, our results would remain to hold if we allowed for aggregate growth in demand $1 + g_b = B_{t+1}/B_t$, with $g_b > 0$. Moreover, note that firm-level productivity, q, grows at the rate of average productivity, $\overline{q} = 1/\Omega \int_j q_j dj$. However, since firm-level productivity is always paired with a demand weight, this simple productivity average is inconsequential for the model.

Note that Assumption 1 is not particularly restrictive. First, empirical evidence on the curvature of the R&D cost function typically points to a value of $\psi = 2$ (see e.g. Hall et al., 2001; Bloom et al., 2002). Second, while at the lower end of estimates, an elasticity of substitution of $\eta = 2$ still falls within the range documented in Broda and Weinstein (2006). Finally, a common and constant demand growth across firms and time, $\theta_{j,t} = \theta$ and $b_{j,0} = b_0$, of course naturally nests the case considered in existing models of endogenous growth, where firm-level demand is assumed fixed, i.e. $\theta = 1$.[16]

Therefore, the following analysis extends existing models by allowing for the possibility of a non-zero growth rate in firm-level demand, i.e. $\theta \neq 1$. We begin by showing that under Assumption 1, firm value is in fact proportional to the product of demand and productivity. This feature then enables us to derive a number of model properties.

Proposition 1 (Firm value and the innovation rate). *Firm value has the following form*

$$\hat{V} = \mathcal{A}(\theta) b \hat{q}$$

where $\mathcal{A}$ is implicitly defined as the positive real solution to

$$\mathcal{A}(\theta) = \frac{\frac{1}{4v} - \frac{x(\mathcal{A})^2}{2\gamma}}{1 - \frac{\beta(1-\delta)}{(1+g)} \left(1 + \lambda x\left(\mathcal{A}\right)\right) \theta} > 0,$$

and where the innovation rate $x(\mathcal{A})$ is given by

$$x(\mathcal{A}) = \gamma \frac{\beta(1-\delta)}{(1+g)} \theta \lambda \mathcal{A}(\theta).$$

All else equal, firm-level innovation increases with firm-level demand growth

$$\frac{\partial x(\mathcal{A})}{\partial \theta} > 0.$$

Proposition 1 shows that the optimal, firm-specific, innovation rate is constant and independent of particular values of production efficiency q and demand b. Therefore, it is common across all businesses and thus independent of firm size.

However, optimal innovation rates do depend on the common demand growth rate θ. Therefore, profitability growth unrelated to productivity directly impacts incentives to conduct R&D. In particular, firms which face faster profitability growth (higher θ), will optimally choose higher rates of productivity investments. The reason for this is that firms with higher expected demand growth foresee that they will be able to reap more

[16] Recall that we assume, for the sake of simplicity, that B is constant. This naturally occurs if $\theta = 1$, which does not affect any of our analytical results.

benefits from successful innovations, increasing their incentives to conduct R&D. This is akin to the "market size effect" identified at the aggregate level in earlier vintages of endogenous growth models (see e.g. Jones, 1995).

Note, however, that our framework does not feature market size effects at the aggregate level.[17] This is because the costs of conducting R&D are expressed relative to aggregate expenditure and, because of the free entry condition, firm mass is a function of labor supply. An increase in the scale of the economy due to, say, an increase in the level of labor supply N, would translate into a proportional rise in the mass of firms, and hence the mass of offered varieties, Ω. As in e.g. Howitt (1999) and Young (1998), this means that consumption is spread more thinly over a larger number of products nullifying the impact of increased scale on aggregate growth.[18]

Firm size distribution. We now turn to characterizing the model-implied firm size distribution.

Proposition 2 (Firm size growth). *Expected growth of surviving incumbents is*

$$\frac{n_{t+1}}{n_t} = \frac{\theta(1 + \lambda x(\mathcal{A}))}{1 + g}.$$

Proposition 2 makes explicit that firm size growth depends both on demand growth, θ, and production efficiency increases, $1 + \lambda x$. As will become clear below, the optimal innovation rate x is constant and common to all firms. Therefore, our theoretical framework satisfies Gibrat's law.[19]

Proposition 3 (Firm size distribution). *Assume that the initial joint distribution of firm-level productivity and demand, $H(s_e)$, is such that the size distribution of entrants follows a Pareto distribution with shape parameter $\kappa_n > 1$. Then, the equilibrium firm size distribution is also Pareto with shape parameter κ_n.*

Proposition 3 shows that the firm size distribution can follow a Pareto distribution, consistent with the highly skewed firm size distribution observed in the data.

Aggregate economic growth. As a final step, let us now turn to aggregate economic growth in our model.

[17] See the Appendix for a formal proof.

[18] The intuition that in a general-equilibrium firm-dynamics model an increase in supply of labor leads to a higher firm entry, rather than price adjustment, goes back to Hopenhayn and Rogerson (1993) and is more recently utilized by e.g. Hopenhayn et al. (2020) or Karahan et al. (2021).

[19] While Gibrat's law has been widely documented for large firms, young and small businesses seem to deviate from it (see e.g. Evans, 1987).

Proposition 4 (Aggregate growth). *Aggregate growth is given by*

$$\frac{Q_{t+1}}{Q_t} = 1 + g = 1 + \lambda x(\mathcal{A}).$$

All else equal, aggregate growth increases with firm-level demand growth

$$\frac{\partial g}{\partial \theta} = \lambda \frac{\partial x}{\partial \theta} > 0.$$

Proposition 4 first makes clear that aggregate economic growth simply reflects the endogenous choices of firms to invest into their firm-specific productivity. However, aggregate growth does not *directly* depend on demand growth. This is intuitive since aggregate demand, B, is assumed to be fixed and there are no aggregate market size effects.[20] It is only the allocation of firm-specific demand within B, which changes over the life-cycle of firms. As we have seen from Proposition 1 these firm-level demand changes, however, have a fundamental impact on R&D incentives. Therefore, firm-level demand growth impacts aggregate economic growth *indirectly* through its effect on firm-level innovation decisions.

2.3 Analytical characterization of the planner's solution

The paragraphs above have shown that firm-level changes in profitability which are unrelated to productivity growth are important in driving innovation and, in turn, aggregate economic growth.

Before taking our model to the data we turn to analyzing the socially optimal level of innovation. In particular, we turn to the problem of a benevolent social planner maximizing economy-wide welfare. In doing so, we identify a novel channel, the "profitability distortion", which may lead firms to *over-invest* in R&D.

Socially optimal innovation at the firm-level. Given that we are not interested in monopolistic distortions as such, we assume that the planner is restricted to the decentralized production and pricing choices and only focus on R&D decisions.

Proposition 5 (Socially optimal innovation). *The socially optimal innovation rates satisfy*

$$\underbrace{2\upsilon \frac{x(\theta)^*}{\gamma}}_{\textit{marginal costs}} = \underbrace{\frac{\beta(1-\delta)}{(1+g^*)(1-\beta)}}_{\textit{discounting}} \underbrace{\lambda\theta}_{\textit{flow benefits}}. \tag{20}$$

[20] Note that while $\partial B/\partial \theta > 0$, we are concerned with the new long-run steady state in which aggregate demand is fixed (albeit at a higher level if θ increases). Therefore, aggregate demand cancels out in the definition of economic growth $Q_{t+1}/Q_t = 1 + g$.

For convenience, we repeat the innovation rates in the market allocation, where we explicitly write out the expression for $\mathcal{A}(\theta)$:

$$\underbrace{2\hat{W}\frac{x(\theta)}{\gamma}}_{\text{marginal costs}} = \underbrace{\frac{\beta(1-\delta)}{(1+g)\,(1-\beta(1-\delta)\theta)}}_{\text{discounting}} \underbrace{\lambda\theta\widetilde{\pi}(\theta)}_{\text{flow benefits}}\,, \tag{21}$$

where $\widetilde{\pi}(\theta) = 1/(4\upsilon) - x(\theta)^2/(2\gamma)$.

The above incorporates three well known differences between the socially optimal and market innovation rates in endogenous growth models. First, there is a difference in the (marginal) costs of innovation because the individual firm faces wage costs, while the planner takes into account the disutility of labor. This is the so called *monopolistic distortion*, which arises because wages are partly determined by production workers in monopolistically competitive firms. The wage may, therefore, be inefficiently low and thus lead to over-investment into R&D.

Second, the planner and the individual firm differ in their *discounting*. While the planner takes into account that any innovation lasts forever, the individual firm only takes into account its own lifetime. Therefore, because the individual firm discounts the future more heavily, this leads to under-investment in R&D.

Third, the individual firm only takes into account the private flow benefits (implicitly defined through $\widetilde{\pi}(\theta)\lambda\theta$), while for the planner the benefit is the contribution to growth ($\lambda\theta$). This effect is referred to as *limited appropriability.* The fact that the private firm cannot appropriate the entire benefits from its own innovation typically leads to under-investment in R&D.

Note that demand growth θ shows up directly in the flow benefits of both the planner and the individual firm in the decentralized economy. This simply reflects the fact that firms are characterized by different weights in the aggregate productivity index Q, which the planner also takes into account. However, this direct effect does not drive a wedge between the decisions made by the planner and those carried out in the decentralized economy.

Profitability distortion. Finally, we now analyze how the presence of firm-level demand growth creates a difference between the privately and socially optimal level of innovation rates. Typically, endogenous growth models exhibit *under-investment* in R&D as the monopolistic distortion discussed above is not strong enough to overturn the other two effects (see e.g. Aghion and Howitt, 1994; Denicolo and Zanchettin, 2014).[21] Our model,

[21] Our model lacks another distortion, the *business stealing effect* present in models based on the tradition of Klette and Kortum (2004). However, Denicolo and Zanchettin (2014) highlight that business stealing itself cannot lead to over-investment in standard endogenous growth models.

however, puts forward a novel channel through which individual firms may over-invest in R&D.

To understand this, notice that the discount rate of individual firms in (21) is affected by demand growth. In contrast to this, the discounting of the social planner in (20) does not feature demand growth. The intuition for this rests on the fact that innovation of individual firms is governed by growth in future profitability, which is partly driven by life-cycle growth unrelated to productivity enhancements. Firm-level demand growth, therefore, provides an extra boost to the incentives of individual businesses to conduct R&D. In contrast, the social planner internalizes the fact that aggregate demand is fixed and focuses only on productivity growth.[22]

We dub this effect the *profitability distortion* because it comes about only if firm profitability is affected by factors other than productivity growth. In principle, the profitability distortion could lead to *over-investment* in R&D if demand growth is strong enough.[23] But even if this channel is not strong enough, and the economy is characterized by under-investment in R&D at the aggregate level, the profitability distortion creates heterogeneity in the extent to which firms under-invest. In particular, firms with high demand growth are likely to under-invest less, potentially opening the door to more targeted R&D policies.

2.4 Possible extensions and discussion

Before moving on to the quantitative analysis, we turn to a brief discussion of our theoretical findings and of possible extensions to our model framework.

Endogenous exit and firm selection. A crucial aspect of firm-level growth is the process of business churn, or so called up-or-out dynamics (see e.g. Haltiwanger et al., 2013). In addition, Foster et al. (2008) have documented that demand side factors seem to play a dominant role in determining the firm selection process.

Up until now our theoretical model featured only exogenous business exit. While this allowed for closed form solutions, it precludes us from understanding how firm selection shapes the results discussed above. Therefore, one of the key advantages of the quantitative analysis is the ability to incorporate endogenous entry and exit and in turn rich up-or-out dynamics in our model.

Heterogeneity in R&D ability and innovation speed. At least since Lentz and Mortensen (2005), researchers have emphasized the importance of accounting for hetero-

[22]Note firms (products) with higher demand growth become more important for welfare as their weight in aggregate demand increases. This is reflected in the flow benefits term, but this effect is present in both the the planner's allocation (20) as well as the decentralized economy (21).

[23]Indeed, when $\theta > \frac{1}{1-\delta}$, individual firms will discount the future *less* than the social planner.

geneity in the ability to conduct R&D, γ. For example, Lentz and Mortensen (2008b) estimate that about half of aggregate growth can be attributed to a reallocation of resources from less to more efficient firms in terms of generating innovations (also called the "selection effect"). Similarly, several models assume that certain types of firms differ in innovation "speed" or the "step size" of innovations, λ (see e.g. Mukoyama and Osotimehin, 2019, for a recent analysis).

While our framework does not feature heterogeneity in R&D ability or innovation speed, the qualitative conclusions from our framework remain unchanged in their presence. This can be easily seen from Lemma 1 which is qualitatively unaffected in the presence of an R&D "firm fixed effect" (γ_j or λ_j). Moreover, as we discuss in detail in the next Section, our parametrization strategy leads to realistic productivity and demand shocks. In particular, we show that our model-implied productivity and demand patterns are consistent with empirical estimates provided by Foster et al. (2008).

That said, Acemoglu et al. (2018) highlight the interaction between firm selection and ex-ante heterogeneity in R&D, especially when it comes to policy evaluation. As will become clear, one of the key conclusions from our model is that the nature of firm selection, and in particular whether it occurs on productivity or profitability, is key for understanding the impact of pro-growth policies. Therefore, incorporating ex-ante heterogeneity into our framework and analyzing optimal policies may be a promising avenue for future research.

Imperfect information and endogenous demand accumulation. Information frictions play a key role in firm dynamics models (see Jovanovic, 1982, for a seminal contribution). Similarly, endogenous accumulation of demand has featured in several recent theoretical and empirical studies of firm growth (see e.g. Gourio, 2014; Arkolakis, 2016; Foster et al., 2016).

In our framework, individual firms have perfect information about their ability to conduct R&D as well as about the demand that they are facing. Moreover, demand evolution is exogenous. Sterk r al. (2021) show that information frictions have relatively little impact in a similar type of model, and Eslava and Haltiwanger (2021) document that markups play only a modest role for the cross-sectional variability of firm-level sales.

Nevertheless, in the Appendix we provide an extension to our baseline model in which we allow firms to endogenously accumulate demand. We recalibrate the extended model to match the same targets as our baseline and show that firm-level demand is somewhat *more* important for aggregate economic growth compared to our baseline results.

3 Estimation and Quantitative Analysis

In this section, we relax the parametric assumption which allowed us to develop closed form characterizations of our model and we extend it to include endogenous firm exit. This leads to two important advantages. First, this generalized version of our theoretical model allows us to estimate a realistic combination of demand and productivity dynamics at the firm level and, in turn, to quantitatively evaluate our structural model. Second, we can analyze the role played by firm selection - along both productivity and demand - in driving aggregate growth.

In what follows, we first describe how our model can be extended to include endogenous exit. Next, we discuss the parametrization of our framework, explicitly discussing how we are able to separately discipline productivity and profitability dynamics which are key for our quantitative analysis. Finally, this section offers descriptive statistics on the fit of our estimated model. The quantitative results are presented in the next section.

3.1 Endogenous firm exit

In order to introduce endogenous firm exit, we assume that firms must pay a per-period fixed cost ϕ (denoted in units of labor) in order to stay in operation. If businesses choose not to pay the entry cost, they shut down and obtain a return of zero. We assume that firms pay the cost at the beginning of the period before demand realizes.[24] Specifically, the beginning-of-period firm value can be written as

$$V^c(s_{j,t}) = \max[0, \mathbb{E}_{t-1}V(s_{j,t}) - W_t\phi],$$

where $V(s_{j,t})$ now represents beginning-of-period firm value, conditional on remaining in operation, defined as

$$V(s_{j,t}) = \max_{p_{j,t},n^c_{j,t},n^r_{j,t}} \left\{ \begin{array}{c} p_{j,t}c_{j,t} - W_t(n^c_{j,t} + n^r_{j,t}) + \\ \mathbb{E}\beta_t(1-\delta)\left[x_{j,t}V^c(s^+_{j,t+1}) + (1-x_{j,t})V^c(s^-_{j,t+1})\right] \end{array} \right\}, \tag{22}$$

All optimality conditions remain the same with future firm values redefined accordingly. The above setting results in a cutoff rule for firm exit. In particular, there exists a threshold $\widetilde{s}_{j,t}$ (a combination of idiosyncratic productivity and demand) at which businesses are exactly indifferent between shutting down and remaining in operation, s.t. $\mathbb{E}_{t-1}V(\widetilde{s}_{j,t}) = W_t\phi$. Finally, labor market clearing now also takes into account the labor

[24]Note that innovations realize at the end of a given period and, therefore, productivity is given at the beginning of the period.

used in firm operation:

$$N_t = \int_{j \in \Omega_t} (n^c_{j,t} + n^r_{j,t} + \phi) dj + \kappa M_t. \tag{23}$$

3.2 Model parametrization

We now match our model to firm data and a set of aggregate moments. The quantitative results crucially depend on realistic productivity and demand dynamics and the next section shows that our parametrized model delivers such patterns. In particular, within our model we replicate the analysis of Foster et al. (2008), who estimate productivity and demand shocks using detailed firm-level quantity and price data. We show that the productivity and demand patterns implied by our model are very close to those estimated by Foster et al. (2008).

The majority of our model is disciplined by aggregated firm-level information taken from the Business Dynamics Statistics (BDS) of the U.S. Census Bureau. The moments that we will utilize are the firm size and exit life-cycle profiles and the autocovariance structure of log-employment at the firm-level.[25] Sterk ⓡ al. (2021) highlight the importance of the latter for correctly pinning down the nature of driving forces at the firm level and we discuss in detail how we use these moments to discipline our framework. Finally, aggregate economic growth is measured by average real GDP growth in the U.S. National Accounts. The sample period is $1979 - 2012$, dictated by the available BDS data.

Productivity, profitability and the autocovariance of employment. Let us begin by considering the autocovariance structure of log-employment predicted by a large class of endogenous growth models and how it relates to the data. The empirical autocovariance structure of log-employment (see Panel (c) in Figure 1) features strongly decreasing autocovariances in horizon, i.e. $\text{cov}(\ln n_a, \ln n_{a+h_1}) > \text{cov}(\ln n_a, \ln n_{a+h_2})$ with $h_1 < h_2$. In contrast, the following proposition shows that a large class of endogenous growth models predicts constant autocovariances with horizon, i.e. $\text{cov}(\ln n_a, \ln n_{a+h_1}) = \text{cov}(\ln n_a, \ln n_{a+h_2})$ with $h_1 < h_2$.

Proposition 6 (Autocovariance of log-employment). *Let j indicate individual firms and a indicate their age. Consider a class of endogenous growth models in which (i) employment is proportional to productivity at the firm-level, $n_{j,a} = \chi \hat{q}_{j,a}$ with $\chi > 0$, (ii) realized firm-level productivity growth is independent of past productivity levels*

$$\text{cov}(\ln \hat{q}_{j,a} - \ln \hat{q}_{j,a-h}, \ln \hat{q}_{j,a-h}) = 0, \quad for \;\; h = 1, ..., a \;\; and \;\; a > 0,$$

[25]While the BDS does not offer these moments, they form the central focus in Sterk ⓡ al. (2021) who also provide them on their websites.

(iii) demand is fixed at the firm-level, $b_{j,a} = b_j$ for all ages a.

In this class of models, the firm-level autocovariance of log-employment is constant with horizon $h > 0$

$$\mathrm{cov}(\ln n_{j,a}, \ln n_{j,a+h}) = \mathrm{var}(\ln n_{j,a}).$$

The Appendix A.3 provides a proof of the proposition and a discussion of model features which deliver such autocovariance patterns. Examples include models in which firm-level innovation rates are constant, though possibly heterogeneous across firms (see e.g. Klette and Kortum, 2004; Lentz and Mortensen, 2008a), but also features such as idiosyncratic and transitory exogenous variation in innovation step size, λ (such as that considered in e.g. Sedláček, 2019). Our current structural framework falls under the conditions in Proposition 6 when $\eta = \psi = 2$ and $\theta_{j,a} = 1$ for all firms j.

Therefore, a large class of endogenous growth models features employment patterns which are inconsistent with the data. Extensions that have the potential to reconcile such models with the data include those adopted in Akcigit and Kerr (2018) or Mukoyama and Osotimehin (2019). The latter consider labor adjustment frictions (firing costs) and, in addition to endogenous innovation, also exogenous, stationary and persistent shocks affecting sales. Akcigit and Kerr (2018), instead, allow for heterogeneous innovation types which do not perfectly scale with firm size, resulting in innovation rates that vary over the firm size distribution. These extensions lead to departures from Gibrat's law and, therefore, are better placed to match the autocovariance structure of log-employment.[26]

Our framework features similar mechanisms. In particular, as in Akcigit and Kerr (2018), our model predicts innovation rates that differ across the firm size distribution even if firm-level demand is fixed.[27] As in Mukoyama and Osotimehin (2019), sales in our model are also affected by persistent stochastic shocks. The key difference is that our framework allows for rich firm-level demand variation, which is quantitatively disciplined by the autocovariance structure of log-employment.

We believe that our approach is justified for at least four reasons. First, as we show in the next subsection, the productivity and demand shocks resulting from our parametrization are very close to those estimated in the data using detailed information on firm-level quantities and prices (see Foster et al., 2008). Second, our approach is firmly grounded in existing research as it combines standard features of endogenous growth models with those found in demand-driven models of firm dynamics (see e.g. Klette and Kortum, 2004; Sterk ⓡ al., 2021). Third, our framework conforms with existing research into the

[26]Mukoyama and Osotimehin (2019) use the empirical first-order autocorrelation coefficient of employment growth to discipline their extension with exogenous, stationary and persistent shocks.

[27]Because R&D costs scale with $\hat{q}^{\eta-1}$ in our model, more productive (and thus larger) businesses face higher innovation costs.

determinants of firm-level R&D, which suggest an important role for market power and frictions in expanding market size (see e.g. Acemoglu and Linn, 2004; De Ridder, 2020). Fourth, our model is consistent with existing empirical evidence on the importance of demand-side factors for firm selection and growth (see e.g. Foster et al., 2008; Hottman et al., 2016).

Demand and productivity processes. Let us now provide details of how we incorporate firm-level demand variation into our framework. Towards this end, we follow Sterk ⓡ al. (2021) and specify the following process for firm-level demand:

$$\begin{array}{rcll}
\ln b_{j,a} & = & \ln u_{j,a} + \ln v_{j,a} + \ln z_{j,a} & \\
\ln u_{j,a} & = & \overline{u}_j + \rho_u \ln u_{j,a-1}, & \ln u_{j,-1} \sim N(0, \sigma_u^2), \quad \overline{u}_j \sim N(\mu_{\overline{u}}, \sigma_{\overline{u}}^2), \quad |\rho_u| \leq 1 \\
\ln v_{j,a} & = & \rho_v v_{j,a-1} + \epsilon_{j,a}, & \ln v_{j,-1} \sim N(0, \sigma_v^2), \quad \epsilon_{j,a} \sim N(0, \sigma_\epsilon^2) \quad |\rho_v| \leq 1, \\
\ln z_{j,a} & \sim & N(0, \sigma_z^2) &
\end{array}$$

where $\ln u$ represents a *demand profile* (with a stochastic initial value, $u_{j,-1}$) which gradually evolves, eventually settling at $\ln u_{j,\infty} = \overline{u}_j/(1-\rho_u)$. The term $\ln v$ represents *demand shocks* (with a stochastic initial value, $v_{j,-1}$) and $\ln z$ represents transitory demand disturbances.

The theoretical model highlighted the importance of expected demand growth, $\theta_{j,a} = \mathbb{E}_a[\Delta b_{j,a+1}]$, for firm-level decisions. Note that in this case, expected demand growth is given by

$$\theta_{j,a} = \mathbb{E}_a[\Delta b_{j,a+1}] = \overline{u}_j + (\rho_u - 1) \ln u_{j,a} + (\rho_v - 1) \ln v_{j,a} - \ln z_{j,a}. \tag{24}$$

In addition to the initial values of demand specified above, entrants also draw stochastic productivity levels. In particular, we assume that

$$q_{j,0,t} \quad = \quad Q_{t-1} q_{j,e}, \quad q_{j,e} \sim N(0, \sigma_q^2),$$

where entrants obtain a productivity draw proportional to last period's aggregate productivity index.

Parameters set a priori and normalizations. We set three parameters a priori. First, we assume the model period to be one year and therefore we set the discount factor to $\beta = 0.97$. Second, we set the elasticity of substitution between goods to $\eta = 6$, within the range of values reported in Broda and Weinstein (2006) and implying a markup of 20%. Third, following the microeconometric evidence on innovation, we set the elasticity of the R&D cost with respect to the success probability of innovation to $\psi = 2$ (see e.g. Hall and Ziedonis, 2001; Blundell et al., 2002; Bloom et al., 2002).

In addition, we choose two normalizations. First, we set the fixed cost of entry, κ, which controls the mass of firms in the economy such that aggregate consumption is

Table 1: Parameter values

parameter		value	parameter		value
β	discount factor	0.970	$\sigma_{\overline{u}}$	$\overline{u}$, standard deviation	1.216
η	elasticity of substitution	6.000	$\mu_{\overline{u}}$	$\overline{u}$, mean	-1.733
υ	disutility of labor	1.000	σ_u	$\ln u$, standard deviation	1.256
κ	entry cost	0.242	ρ_u	$\ln u$, persistence	0.397
ϕ	fixed cost of operation	0.388	σ_v	$\ln v$, standard deviation	0.594
δ	exogenous exit rate	0.021	ρ_v	$\ln v$, persistence	0.984
γ	R&D efficiency	0.158	σ_ϵ	ϵ, standard deviation	0.285
λ	innovation step size	0.137	σ_z	$\ln z$, standard deviation	0.285
			σ_q	q_e, standard deviation	0.090

Note: β, η, υ and κ are calibrated as discussed in the main text. The remaining parameters are set such that the model matches the empirical age profiles of average size, exit rates, and the autocovariance of log-employment from startup (age 0) to age 19 in the BDS.

normalized to $\hat{C} = 1$. Second, we set the disutility of labor such that the aggregate wage is normalized to $\hat{W} = 1$.

Remaining parameters. The remaining 14 parameters – those of the demand and productivity processes, the exogenous exit rate and the operational cost – are set by matching our model to key moments in the data.

In particular, we target 250 moments which can be grouped into four sets: (i) average growth of real GDP (1 moment), (ii) the firm size life-cycle profile (20 moments), (iii) the firm exit life-cycle profile (19 moments) and (iv) the upper triangle of the autocovariance matrix of log-employment, by age and for a balanced panel of firms surviving up to at least the age of 19 years (210 moments). All model parameters are shown in Table 1 and we defer the details of the solution and simulation procedures to the Appendix.

3.3 Model-implied productivity and demand dynamics

The relative importance of productivity and demand variation at the firm level is crucial for our quantitative analysis. Therefore, the following paragraphs document that our parametrization strategy delivers firm-level variation in productivity and demand which very closely matches existing empirical evidence. This provides additional validation of our structural model and the quantitative results presented in the next Section.

Estimating productivity and demand shocks. To gauge the realism of the driving forces of our model, we draw upon the methodology and empirical evidence presented in Foster et al. (2008). We begin by replicating the estimation procedure in Foster et al. (2008) in order to obtain physical total factor productivity (TFPQ) and demand shocks.

Specifically, TFPQ is estimated as the residual from

$$\ln c_{j,t} = \alpha_0 + \alpha_1 \ln n_{j,t} + \chi_{j,t},$$

where $c_{j,t}$ and $n_{j,t}$ are physical output and employment in firm j.[28] Given the productivity estimates $\chi_{j,t}$, demand shocks are estimated from

$$\ln c_{j,t} = \beta_0 + \beta_1 \ln p_{j,t} + \zeta_{j,t},$$

where prices are instrumented using the TFPQ estimates, $\chi_{j,t}$.[29]

Comparison to the data. Panel A of Table 2 shows the estimated persistence and standard deviations of TFPQ and demand shocks estimated on simulated data from our model. The model-implied dynamics of both productivity and demand shocks are very close to those in the data (see Tables 1 and 3 in Foster et al., 2008).[30]

In addition to the above, in Section 4.5 we estimate productivity (TFPR) and demand shocks using firm-level data from Compustat. Also with this dataset we find that the productivity and demand dynamics implied by our model are empirically realistic. These estimates further validate our parameterization strategy and model. The realism of the model-implied productivity and demand shocks is reassuring, since the firm-level driving forces are key for our quantitative results, to which we turn to next.

3.4 Model fit

Before turning to our quantitative results, we document that the model does well not only in matching the targeted moments, but also in accounting for a range of untargeted ones.

Targeted moments. The empirical targeted moments, together with their model-based counterparts, are depicted in Figure 1. The model does well in matching all three sets of the targeted moments: average exit rate and size by age and the autocovariance

[28]In addition to employment, Foster et al. (2008) also consider capital, material and energy inputs, none of which, however, are present in our model. They use the respective cost shares at the 7-digit industry level to measure input elasticities. We, instead, estimate α_1 directly in our model. However, imposing it to be equal to 1, the elasticity in production, changes very little.

[29]Using prices directly, which are known in the model, changes little of our results.

[30]In order to separate demand from productivity, Foster et al. (2008) focus on specific industries producing physically homogeneous products for which they also observe price information. In contrast, our model is parameterized to the economy as a whole. Therefore, while the model-implied productivity and demand dynamics need not be exactly the same as those estimated in the industries with homogeneous products, the similarity shown below is encouraging and suggests that our model is driven by realistic shocks.

Table 2: Model fit: untargeted moments

	model	data
Panel A: Demand and productivity moments		
demand shock persistence	0.92	0.91
demand shock standard deviation	1.08	1.16
TFPQ persistence	0.83	0.79
TFPQ standard deviation	0.21	0.26
Panel B: Firm dynamics moments		
job creation rate	19%	17%
job destruction rate	19%	15%
job creation share from entry	12%	9%
job destruction share from exit	14%	17%
Panel C: R&D moments		
R&D/GDP	1.5%	2.6%
growth contribution from entry	22%	19%

Note: Panel A shows persistence (from an unweighted regression) and volatility estimates for TFPQ and demand shocks from Foster et al. (2008). Panel B shows firm dynamics moments, taken from the BDS. Finally, in Panel C the R&D/GDP ratio is taken from the Bureau of Economic Analysis and the contribution of entrants to aggregate growth is from Acemoglu et al. (2018).

structure of log-employment. The model-implied aggregate growth rate is 1.52%, close to the empirical value of 1.5%.

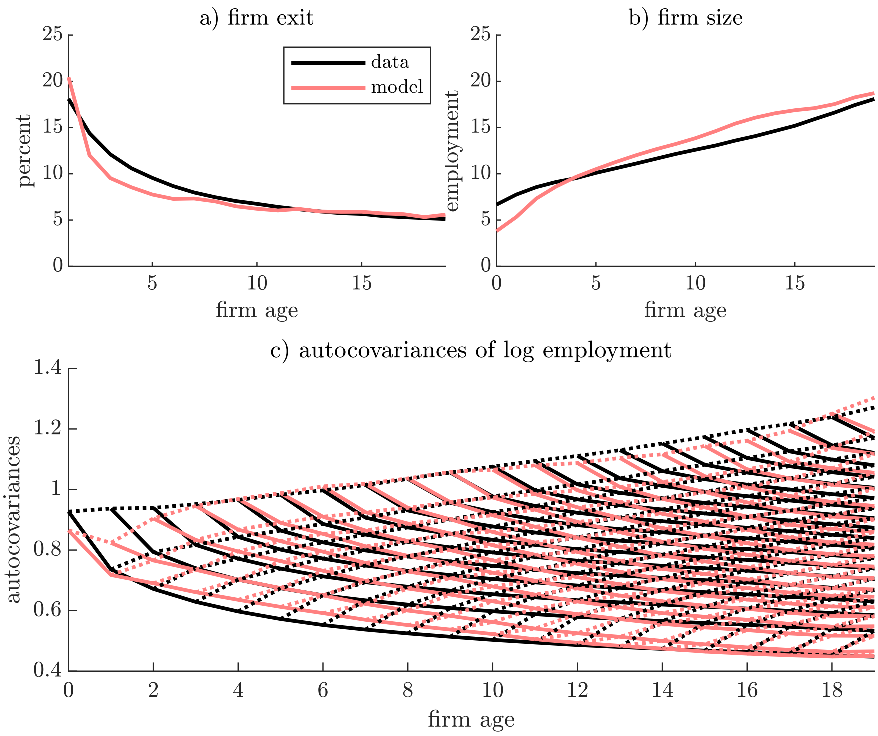
Employment distribution by firm size and age. Figure 2 shows that the model not only fits the autocovariance structure of log-employment, but is also consistent with the employment distribution by firm size and age. This includes the employment share in large firms (500 and more employees), even among young businesses which is typically difficult to match by models.

Job creation and destruction patterns. The model replicates the average job creation and destruction rates observed in the aggregate economy, see the Panel B of Table 2. In addition, it also features realistic contributions to job creation and destruction by entrants and exiting firms, respectively.

R&D and growth patterns. Panel C of Table 2 shows moments related to R&D, innovation and growth. In particular, the ratio of R&D expenditures to GDP and the contribution of entrants to aggregate growth.[31] The model also attributes about 22 percent of aggregate growth to entrants, similar to findings in e.g. Bartelsman and Doms (2000); Foster et al. (2008); Acemoglu et al. (2018).

[31] The empirical value of R&D expenditures reflects "Real Gross Domestic Product: Research and Development" of the Bureau of Economic Analysis.

Figure 1: Model fit



Note: Top panels show average firm size (employment) and exit rates by age in the model and the BDS data. The bottom panel shows the observed and model-implied autocovariance matrix of log employment for a balanced panel of firms surviving at least up to age 19.

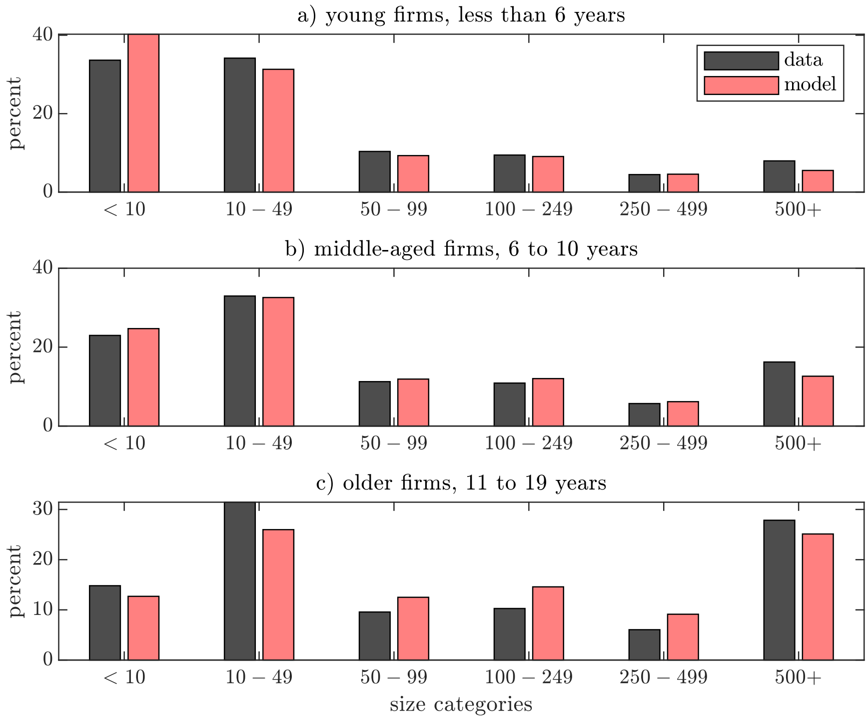
4 Quantitative results

This section quantitatively evaluates our model, presenting our results in five steps. First, we begin with describing a counterfactual economy used to isolate the contribution of changes in profitability which are unrelated to productivity dynamics on firm-level choices and aggregate outcomes. Second, we use our counterfactual economy to document how demand variation impacts firm-level innovation, productivity growth and up-or-out dynamics. Third, we turn our attention to analyze the implications of our framework for aggregate economic growth. Fourth, we show that explicitly making a distinction between productivity and profitability is crucial for quantitative conclusions about the aggregate impact of pro-growth policies. Finally, we end this section by providing empirical support to our key channel, which links business-level expected demand growth to firms' innovation decisions.

4.1 Decomposing the baseline model: a counterfactual economy

The key choices of heterogeneous firms in our model pertain to investment into R&D and to remaining in operation or shutting down. These decisions form the backbone of

Figure 2: Firm size-age distribution: data and model



Note: Size-age distribution in the BDS data and in the model. Both distributions are expressed as shares of employment in a given age category (young: < 6 years, middle-aged: $6 - 10$ years and older: $11 - 19$ years).

aggregate economic growth. Unlike labor demand or pricing decisions, which are static in our framework, R&D investment and firm exit are forward-looking – based on expected firm values. In contrast to existing models of endogenous growth, our framework allows for profitability, and in turn firm values, to be affected by idiosyncratic changes in demand. This possibility creates a wedge between profitability and productivity which is absent in existing models of endogenous growth.

To isolate the contribution of this new channel, we now present a counter-factual economy. The key goal is to *decompose* the predictions of the baseline model into those made because of variation in demand and a remainder. Therefore, the counterfactual economy retains firm-level demand variation exactly as in the baseline model, however, we assume that firms ignore it. Instead, they believe that idiosyncratic demand levels are fixed.[32] This effectively mimics our theoretical analysis in which expected demand growth, θ_j, was key in determining firm-level and aggregate outcomes. Therefore, in our counterfactual economy, firms will expect $\theta_{j,t} = \mathbb{E}_t[\Delta b_{j,t+1}] = 1$.

[32]We choose this counterfactual, because assuming fixed demand at the firm level is conceptually identical to existing models of endogenous growth where demand is effectively fixed at 1. An alternative would be to allow firms to recognize the stochastic nature of the demand process, but require $\mathbb{E}_t[\Delta b_{j,t+1}] = 1$. However, this restriction effectively pins down expected realizations of the demand shock $\epsilon_{j,t+1}$, which does not mimic existing models of endogenous growth.

The counterfactual economy: no expected demand growth Our counterfactual economy is identical to the baseline model, except that firms believe, in each period, that their observed demand levels will remain fixed in the future. Therefore, instead of firm value being described by (22), in the counterfactual economy it is given by

$$\underline{V}(q_{j,t}, b_{j,t}) = \max_{\underline{p}_{j,t}, \underline{n}^c_{j,t}, \underline{n}^r_{j,t}} \left\{ \begin{array}{c} \underline{p}_{j,t} \underline{c}_{j,t} - W_t(\underline{n}^c_{j,t} + \underline{n}^r_{j,t}) + \\ \\ \mathbb{E}\beta_t(1-\delta) \left[\begin{array}{c} \underline{x}_{j,t} \underline{V}^c(q_{j,t}(1+\lambda), b_{j,t}) + \\ \\ (1 - \underline{x}_{j,t}) \underline{V}^c(q_{j,t}, b_{j,t}) \end{array} \right] \end{array} \right\}, \tag{25}$$

where "underbars" indicate firm-level variables in the counterfactual economy and where we've made the firm's state variables explicit. Notice that the only difference between (25) above and (22) in the baseline model are expected demand levels in the future.

We solve the firm-level optimality problem using the above firm values, but we retain all parameter values and equilibrium variables as in the baseline model. Therefore, the outcomes from this counterfactual economy can be interpreted as a *decomposition* of the results from our baseline model attributed to the presence of changes in profitability unrelated to productivity dynamics.

The counterfactual economy: three distinct scenarios. In our analysis we highlight three distinct versions of our counterfactual economy. These scenarios are based on the fact that ignoring expected demand growth will affect not only individual choices of incumbent firms, but also their composition as firms' exit decisions change.

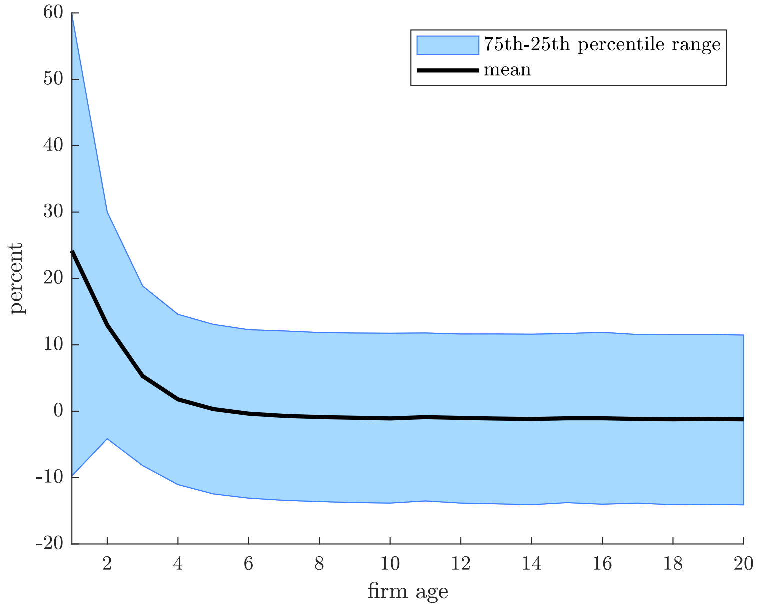
The first, "intensive" margin scenario, allows only firm choices to change, but keeps the composition of firms the same as in the baseline. The second, "extensive" margin scenario, allows the composition of firms to change, but keeps individual firm choices the same as in the baseline economy. Finally, a "overall" scenario, allows for both individual choices of firms and their composition to change compared to the baseline as a consequence of firms ignoring future demand growth.

4.2 Productivity, profitability and firm dynamics

We begin our quantitative analysis be revisiting the theoretical results which showed that expected demand growth impacts firm-level innovation decisions and, in turn, productivity and firm size growth.

Expected demand growth in the baseline model. Before describing how our baseline economy compares to the counterfactual with fixed demand levels, we give a flavor of how expected demand growth in our baseline economy.

Figure 3: Expected demand growth in the baseline model



Note: The figure shows expected demand growth on average, together with the associated 75th-25th percentile ranges in the baseline model.

Figure 3 shows the evolution of expected demand growth over firms' lifecycles. It depicts the average over surviving firms, together with the 25th and 75th percentiles at each firm age. The figure shows that average expected demand growth quickly declines with age, falling slightly below zero after about age 5. This may suggest that especially young firms will be the ones whose choices will be heavily impacted by demand variation. That said, there exists a lot of heterogeneity in expected demand growth, even among old businesses.

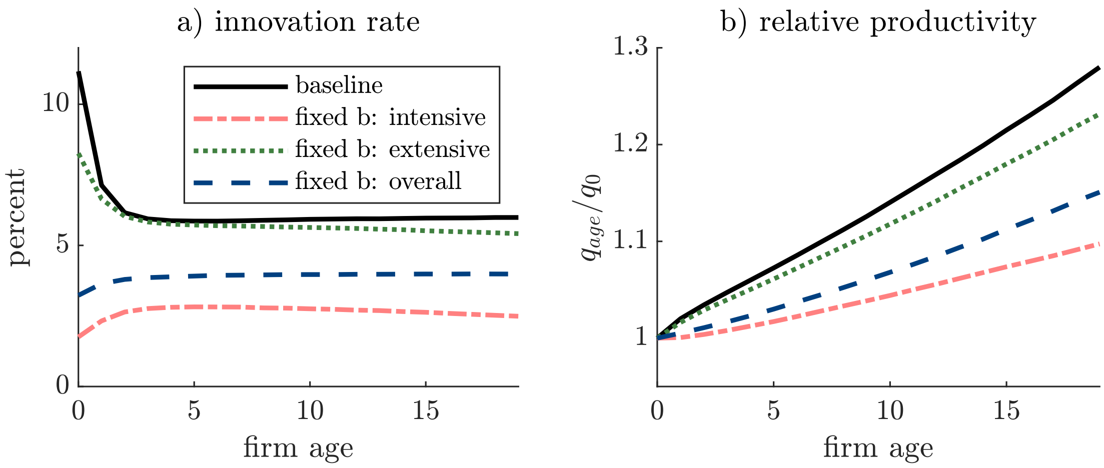
Firm-level innovation: average patterns. Panel a) of Figure 4 shows firm-level innovation rates, while Panel b) depicts productivity levels relative to those of startups in the same cohort. The figure plots values in the baseline economy, together with those from our three counterfactual scenarios.

Three points stand out when it comes to the innovation rate. First, in the baseline economy, the innovation rate declines strongly with firm age. Since young firms are also on average smaller than incumbents (see Panel b in Figure 1), the model also predicts that smaller businesses innovate more, consistent with the empirical evidence (see e.g. Akcigit and Kerr, 2018).[33]

Second, turning to the impact of demand variation on innovation choices, we find that innovation is only slightly affected by the change in composition of firms stemming

[33] Akcigit and Kerr (2018) find a negative relationship between patents per worker and log employment. A similar relationship holds in our model. Specifically, estimating $x_{j,t}/c_{j,t} = \alpha + \beta \ln c_{j,t} + \epsilon_{j,t}$ yields $\beta = -0.02$. Similar results hold when replacing sales with employment.

Figure 4: Profitability vs productivity and innovation



Note: The figure shows average innovation rates (panel a) and productivity relative to that of startups (panel b) by age in the baseline and in the counterfactual economy in which firms ignore expected demand growth. The figures shows results from the "baseline" and three counterfactual scenarios: (i) "fixed b: intensive" in which only firm choices change, but the composition is fixed to that of the baseline, (ii) "fixed b: extensive" in which the composition of firms changes, but firm decisions are fixed to those in the baseline and (iii) "fixed b: overall" which is the combination of the previous two.

from firm-level demand ("fixed b: extensive"). In contrast, firms decide to scale down innovation substantially in the counterfactual economy ("fixed b: intensive"), suggesting that firm-level demand variation is quantitatively important for R&D incentives. In particular, innovation rates decline by about one half (3 percentage points).

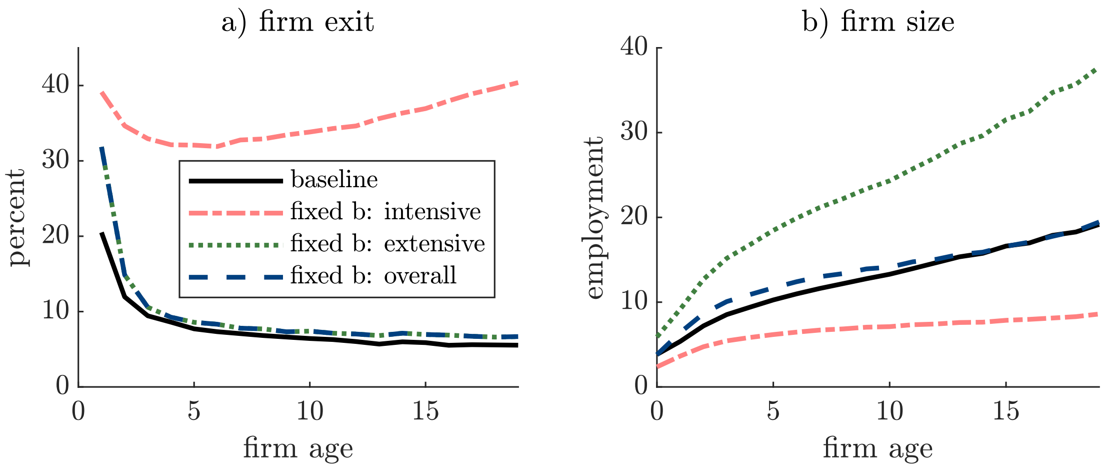
Third, the negative relationship of the innovation rate with age in the baseline disappears when firms do not expect demand to change. These results, therefore, highlight that expected demand growth is not only quantitatively important for firm-level innovation decisions, but that it also is the source of the negative innovation-age relationship.[34]

The right panel then reflects what can be seen in the innovation patterns. Because innovation rates are lower in the counterfactual (and especially so for young businesses), productivity grows much more slowly on average. Therefore, expected demand growth is indirectly responsible for a substantial portion of firm-level *productivity growth*. We will return to this point when analyzing the implications of expected profitability on aggregate growth.

Firm-level innovation: heterogeneity across firms. The average patterns depicted in Figure 4 hide a substantial amount of heterogeneity across firms. On average, the 75th percentile of the innovation rate distribution is about 2/3 higher than the 25th percentile. Moreover, this value does not decay as firms age with innovation heterogeneity being present even among very old businesses.

[34]This holds also for the innovation-size relationship. Estimating $\underline{x}_{j,t}/\underline{c}_{j,t} = \alpha + \beta \ln \underline{c}_{j,t} + \epsilon_{j,t}$ in the counterfactual yields $\beta = 0.0001$.

Figure 5: Profitability vs productivity and up-or-out dynamics



Note: The figure shows average exit rates (panel a) and average size (panel b) by age in the baseline and in the counterfactual economy in which firms ignore expected demand growth. The figures shows results from the "baseline" and three counterfactual scenarios: (i) "fixed b: intensive" in which only firm choices change, but the composition is fixed to that of the baseline, (ii) "fixed b: extensive" in which the composition of firms changes, but firm decisions are fixed to those in the baseline and (iii) "fixed b: overall" which is the combination of the previous two.

In addition, much of the innovation differences are driven by heterogeneity in expected demand growth, as is suggested by Proposition 1. In particular, the correlation between innovation rates and expected demand growth is 0.48 in the baseline economy. The same correlation in the counterfactual where firms ignore future demand growth is 0.04.

Therefore, as in the theoretical analysis, also in the full model expected demand growth is a strong determinant of firm-level innovation. This link between demand and productivity dynamics, which is absent from existing models of endogenous growth, has important aggregate consequences as well as implications for the efficacy of growth policies. We turn to these questions in the next subsections.

Up-or-out dynamics: average patterns. Figure 5 depicts average exit rates (left panel) and average firm size (right panel). Keeping the composition of firms the same as in the baseline, businesses would choose to shut down much more often if demand were fixed ("fixed b: intensive"). This is because firms which are at the brink of exit in the baseline economy, but choose to remain, do so on the basis of higher expected firm values. If these firms ignore such developments, they are more likely to shut down.[35]

This effect is particularly strong for older businesses suggesting that demand plays an important role for preventing older, potentially less productivity firms, from shutting down. Overall, the quantitative impact of demand on exit rates is very large. About 40

[35] The opposite does not hold, i.e. if firms in the baseline choose to remain in operation despite an expected drop in future demand, they will certainly do so under the assumption of fixed future demand.

percent of businesses in the baseline economy would shut down on average if they were to ignore future demand evolution.

Note, however, that because of this increase in exit rates, the composition of firm changes. The exit rates in the counterfactual in which the composition of firms does change are only slightly higher than those in the baseline ("fixed b: extensive" and "fixed b: overall"). The reason for this is that while fixed demand levels induce firms to shut down more often, the composition of businesses shifts towards relatively more productive firms with lower exit rates. The extensive margin adjustment, therefore, undoes most of the intensive margin effect.

The right panel of Figure 5 shows that expected demand growth also plays an important role for the "up" part of up-or-out dynamics. Ignoring expected demand growth flattens the average size-age profile considerably. This is for two reasons. First, individual firms hire fewer workers to conduct R&D, consistent with the lower innovation rates in the left panel of Figure 4. Second, because of their diminished productivity (see the right panel of Figure 4), firms also produce less and therefore hire fewer production workers.

Interestingly, average firm size *increases* as a result of a change in firm selection. This is consistent with the previous paragraphs and is driven by increased exit rates tilt the composition of firms towards businesses with higher demand and therefore larger firms. Hence, the combined effect of ignoring expected demand growth ("fixed b: overall") results in a similar average size profile as in the baseline model.

Up-or-out dynamics: selection on productivity or profitability? The previous paragraphs showed that expected demand growth is very important for firm's exit decisions. However, this does not imply that productivity is inconsequential. We now quantify the extent to which selection in the baseline economy occurs on productivity or profitability.

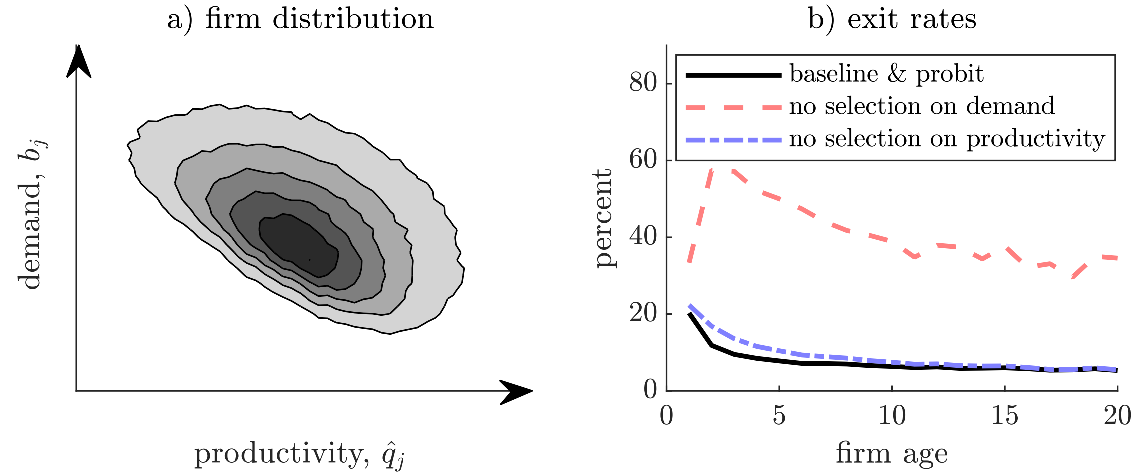
The left panel of Figure 6 shows the distribution of demand, b_j, and (relative) productivity, $\hat{q}_j$, from the simulated model with darker shades indicate more densely populated parts of the state-space. Since firms select on firm values and firm values are a combination of demand and productivity, there is a clear negative relationship between the two among surviving incumbents. Intuitively, if firms enjoy high demand for their product, they can survive while having lower relative productivity and vice versa.

The inverse relationship between productivity and demand has profound implications for selection over firms' lifecycles. To understand this, we estimate the following regression

$$y_{j,t} = \alpha + \beta_{en}\mathbb{1}_{\text{entry}} + \alpha_{ex}\mathbb{1}_{\text{exit}} + \omega_{j,t}, \tag{26}$$

where $y_{j,t}$ is our variable of interest – either log demand, $\ln b_{j,t}$ or relative productivity $\ln \hat{q}_{j,t}$ of incumbent firm j in period t. The variables $\mathbb{1}_{\text{entry}}$ and $\mathbb{1}_{\text{exit}}$ are indicator functions

Figure 6: Productivity and demand distribution in model



Note: Panel a) of the figure shows the distribution of demand and (relative) productivity in the simulated model. Darker shades indicate densely populated areas. White areas indicate areas of the state space which are not populated. Panel b) shows exit rates in the baseline (and the age-dependent probit which coincides with it), and two counterfactuals. One in which demand happens only on productivity ("no selection on demand") and one in which selection happens only on demand ("no selection on productivity").

equal to 1 if firm j in period t is an entrant or a firm that shuts down, respectively. Therefore, the estimated coefficients can be interpreted as percentage differences in demand or productivity of entrants or exiting firms relative to incumbent businesses.

Table 3 shows estimates of β_{en} and β_{ex}, respectively. The results suggest that entering firms are characterized by relative productivity which is about 4 percent higher than that of incumbents. On the other hand, firms which decide to shut down have relative productivity about 4 percent below that of incumbents. In contrast, both entering and exiting firms have demand levels which are below the average incumbent. That said, entrants are burdened by a substantially lower demand level compared to incumbents. Both the model-predicted productivity and demand patterns of entrants and exiters are consistent with empirical evidence (see e.g. Foster et al., 2008, 2016).

Table 3: Productivity and demand: entering and exiting firms

	entry	exit
productivity	0.041	−0.041
demand	−1.168	−0.615

Note: The table shows estimates of coefficients on indicator variables for "entry" and "exit" in regression (26) in the main text. Demand stands fo $\ln b_{j,t}$ and productivity stands for $\ln \hat{q}_{j,t}$.

To quantify the extent to which firm selection is driven by productivity or demand we estimate the following Probit model for firm exit

$$Pr(exit_{j,a} = 1 | b_{j,a}, \hat{q}_{j,a}) = \Phi(\beta_{0,a} + \beta_{b,a} \ln b_{j,a} + \beta_{q,a} \ln \hat{q}_{j,a}), \tag{27}$$

where $exit_{j,t}$ is an indicator function equal to 1 if business j of age a decides to shut down (i.e. is out of operation in the next period).[36] The right panel of Figure 6 shows that the age-dependent Probit model (which mimics the exit pattern in the baseline) together with two counterfactual exit rates. "No selection on demand" is average firm exit predicted by the estimated Probit model when we ignore the effect of demand on exit decisions, i.e. setting $\beta_{b,a} = 0$ for all ages. In contrast, "no selection on productivity" represents firm exit predicted by our Probit model when we ignore the effect of productivity on exit decisions, i.e. setting $\beta_{q,a} = 0$ for all a.

The right panel of Figure 6 shows that demand is a dominant force when it comes to firm selection, consistent with our counterfactual exercise, but also with existing empirical evidence (see e.g. Foster et al., 2008). Ignoring the impact of demand on exit decisions would result in considerably higher exit rates, especially for young businesses. On the other hand, while selection on productivity also plays a role, it's quantitative effect is much weaker. As will become clear below, these results will be key for understanding the aggregate efficacy of growth policies.

Up-or-out dynamics: gazelles. The literature on firm dynamics has developed a consensus that high-growth firms, so called "gazelles", are crucial for aggregate job creation despite the fact that they account for only a small fraction of all businesses (see e.g. Haltiwanger et al., 2016). We now investigate the role of gazelles in our framework.

Following the existing literature, we define gazelles as firms which increase their employment at an annual rate of at least 15% for at least five consecutive years. In our baseline model, we identify about 7% of such businesses, similar to the fraction found in the data (see Haltiwanger et al., 2016).

When inspecting the sources of their employment growth, demand comes out as the dominant factor, consistent with existing empirical evidence (see e.g. Foster et al., 2016). To see this, we regress employment growth on productivity and demand growth and inspect the adjusted R-square contribution of productivity and demand growth. The contribution of demand growth is about eight times as large as that of productivity growth. We will return to the role of gazelles for the rest of the economy in the next subsection when analyzing the drivers of aggregate economic growth.

4.3 Productivity, profitability and aggregate economic growth

The previous paragraphs have highlighted the importance of distinguishing between productivity and profitability for firm-level outcomes. We now turn to quantifying the impact of firm-level demand variation on aggregate economic growth.

[36] While we estimate separate Probit models for each age, similar results are obtained for a single Probit on the pooled dataset.

Demand-driven aggregate growth. We begin by decomposing aggregate economic growth into the contribution of demand variation, once again with the help of our counterfactual economy. The top row of Table 4 shows aggregate growth in the baseline and the three counterfactual scenarios.

The table shows that demand accounts for about 20 percent of aggregate economic growth ("fixed b: overall"). This is entirely driven by firm-level innovation decisions ("fixed b: intensive"). In fact, in isolation the changes in the composition of firms brought about by firms expecting fixed demand levels ("fixed b: extensive") leads to a slight *increase* in aggregate economic growth. This is consistent with our firm-level analysis discussed earlier, which showed that changes in the composition of firms alone have only moderate effects on firm-level productivity and demand, but lead to a composition shift towards more productive businesses.

These results paint a very different picture of the drivers of aggregate economic growth. While aggregate growth has typically been considered a purely supply-side phenomenon, we provide a framework which posits firm-level demand as a quantitatively important source of aggregate growth. This not only alters our understanding of the driving forces of growth, but also its sensitivity to policy interventions, to which we turn to in the next subsection. Moreover, our framework potentially opens the door to a new set of pro-growth policy instruments operating through the stabilization or support of firm-level demand.

Creative destruction and aggregate growth. Section 3.2 already discussed the fact that in the baseline economy startups account for about 22% of aggregate growth, consistent with empirical estimates. This is shown again in the first column of the second row of Table 4.

Note that the contribution of entrants to growth is somewhat larger in the counterfactual economy. This is intuitive since expected demand growth helps incumbents survive and incentivizes them to conduct R&D, see Figures 4 and 5.

Therefore, not taking account of demand-driven profitability variation, which is independent of productivity dynamics, may skew the view on entrants by attributing them too large of a contribution to aggregate economic growth. Such conclusions may have important policy implications which often stress the importance of targeting young businesses which, on average, grow faster than incumbents.

Gazelles and aggregate growth. We have already discussed the role of gazelles in up-or-out dynamics. We now turn our attention to their contribution to aggregate economic growth.

Recall that in the baseline economy gazelles account for about 7% of all firms. However, the last row of Table 4 documents that without this small fraction of firms aggregate

Table 4: Aggregate economic growth: baseline and counterfactual

		Counterfactual: fixed b		
	Baseline	Intensive	Extensive	Overall
Aggregate growth	0.0153	0.0121	0.0154	0.0126
Entrant contribution	22%	27%	27%	32%
Gazelle contribution	30%	38%	26%	37%

Note: The table shows results from the "baseline" and three counterfactual scenarios: (i) "fixed b: intensive" in which only firm choices change, but the composition is fixed to that of the baseline, (ii) "fixed b: extensive" in which the composition of firms changes, but firm decisions are fixed to those in the baseline and (iii) "fixed b: overall" which is the combination of the previous two.

economic growth would be substantially lower. In particular, without gazelles aggregate growth in the baseline would be only 1.05 percent. This means that gazelles alone account for about 30% of aggregate economic growth further enforcing their importance identified in existing studies.[37]

4.4 Implications for Growth Policies

A natural question to consider is: how much do we miss in understanding aggregate economic growth, when considering models in which productivity is the only source of profitability variation? This subsection explores the extent to which failing to account for the distinction between productivity and profitability changes our understanding of the macroeconomic impact of growth policies.

In doing so, we consider a restricted ("productivity-only") economy, versions of which are often used in existing studies. This restricted economy is assumed to completely abstract from demand variation at the firm level.[38] We then analyze two distinct policy experiments often considered in existing studies and in practice: subsidizing firm operation and R&D costs. Our key questions is whether the presence of firm-level demand variation changes the model-implied policy conclusions.[39]

As will become clear, the impact of these policies is substantially different once we consider the distinction between profitability and productivity, as we do in our baseline model. Therefore, our results suggest that failing to account for demand variation at the firm-level can fundamentally change the policy prescriptions stemming from models of

[37]The contributions of gazelles to growth in the counterfactual are similar to that in the baseline, lying roughly between 1/4 and 1/3. Note that this does not mean that firm-level demand variation is unimportant for gazelles. To the contrary, recall that the counterfactual retains the demand variation of the baseline economy, but firms are assumed to ignore it in their decisions.

[38]Note that this differs from the counterfactual "fixed b" economy in the previous sections which served as a tool to *decompose* the results of our baseline model. While in the restricted "productivity-only" economy demand is truly fixed and common to all businesses, the "fixed b" counterfactual features demand variation at the firm level, but assumes that businesses do not take it into account.

[39]Given our interest in the comparison between our baseline and restricted economy, in particular when it comes to aggregate growth, we leave the quantitative analysis of optimal policies and welfare to future research.

endogenous growth.

A restricted "productivity-only" economy. The restricted economy is assumed to be exactly the same as our baseline except that firm-level demand is fixed and common to all firms. This is achieved by assuming that $\sigma_z = \sigma_\epsilon = \sigma_{\overline{u}} = \sigma_u = \sigma_v = 0$.

In order to achieve a meaningful comparison between our baseline model and the productivity-only counterfactual, we recalibrate the latter to match the same targets as discussed in Section 3.2, with the exception of the autocovariance structure of firm-level employment. As we discuss in Section 3.2, a model without firm-level demand variation is not able to match the autocovariance structure well. Instead, we target the baseline R&D/output ratio, a common empirical target (see e.g. Akcigit and Kerr, 2018).[40]

Introducing subsidies to R&D and operational costs. We introduce subsidies to R&D, τ_r, and to the costs of operation, τ_ϕ, in both our baseline and the restricted economy. Specifically, subsidies – assumed to be financed by lump-sum taxes on households – are provided to each business at a common rate. Firm values, therefore, become

$$V(s_{j,t}) = \max_{p_{j,t}, n^c_{j,t}, n^r_{j,t}} \left\{ \begin{array}{c} p_{j,t}c_{j,t} - W_t(n^c_{j,t} + (1-\tau_r)n^r_{j,t}) + \\ \\ \mathbb{E}\beta_t(1-\delta)\left[x_{j,t}V^c(s^+_{j,t+1}) + (1-x_{j,t})V^c(s^-_{j,t+1})\right] \end{array} \right\}, \tag{28}$$

where the continuation values are given by

$$V^c(s_{j,t}) = \max[0, \mathbb{E}_{t-1}V(s_{j,t}) - W_t\phi(1-\tau_\phi)].$$

Clearly, with $\tau_r = \tau_\phi = 0$ we recover our baseline economy. In what follows, we will consider two cases in each economy, one with positive R&D subsidies (but zero subsidies to operation costs) and one with positive subsidies to costs of operation (but zero R&D subsidies). In doing so, we ensure comparability between the baseline and the restricted economy by calibrating the subsidies in the latter such that the overall fraction of output spent on the respective subsidies is equal to the fraction spent in the baseline economy.

Finally, note that in both the baseline and the restricted model we analyze the results in general equilibrium. Therefore, we recompute equilibrium variables: wages, consumption, mass of firms and the aggregate growth rate. Finally, note that we consider the long-run effects of the two subsidies, abstracting from the transition path.

Impact of subsidies to firm operation. Our first policy experiment considers subsidizing firms' costs of operation, assuming $\tau^b_\phi = 0.25$ in the baseline model. The operation cost subsidy in the restricted model is $\tau^r_\phi = 0.20$, ensuring that the resources spent on

[40] See the Appendix for further details on the productivity-only model and its fit to the data.

Table 5: Impact of growth policies

	$\overline{n}$	$\delta(1-\overline{F})$	$\overline{x}$	N	W	g
		Panel A: Operational cost subsidies				
restricted model	-9.6	-17.9	$+9.5$	$+15.7$	$+2.1$	-7.3
baseline model	-24.1	$+1.6$	-1.2	$+39.0$	$+0.0$	-0.8
		Panel B: R&D subsidies				
restricted model	$+3.7$	$+3.3$	$+3.3$	-14.6	-0.5	$+5.6$
baseline model	$+3.4$	$+5.2$	$+11.4$	-29.4	-1.3	$+12.6$

Note: The table shows long-run effects of subsidies to the operation of firms (Panel A) and to R&D costs (Panel B). The table shows results for the baseline model and a restricted, productivity-only, economy in which firms face constant and common levels of demand. The table depicts effects on average firm size ($\overline{n}$), average exit rates ($\delta(1-\overline{F})$), average innovation rates ($\overline{x}$), aggregate employment (N), aggregate wage (W) and aggregate growth (g). Values are reported in percent deviations from the respective economies without policies.

subsidies in both the baseline and restricted economies account for the same share of output.

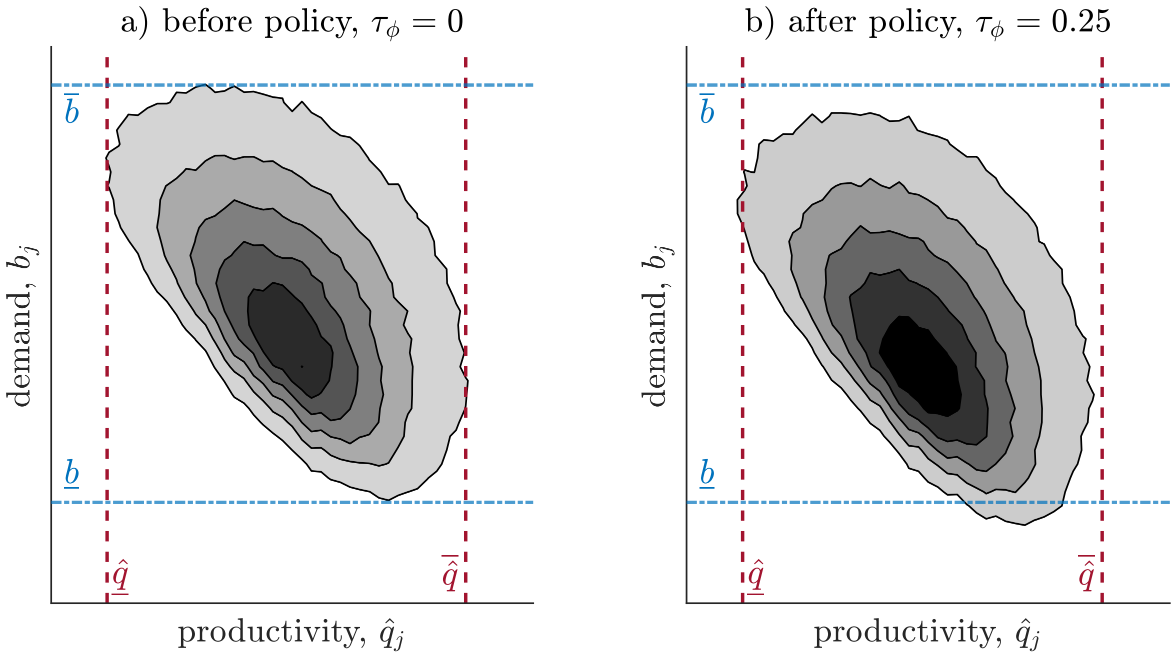
Panel A of Table 5 show the long-run impact of subsidies to operational costs of firms. Let us consider the impact of this policy in the restricted productivity-only economy first. The lower cost of operation makes it easier for businesses to survive, resulting in lower exit rates. The decreased risk of firm exit raises the returns from innovation and therefore firms invest more into R&D, raising the average probability of innovating. Lower exit rates also increase the number of businesses in the economy, raising labor demand and thus wages (consumption) and equilibrium employment. Finally, despite the higher probability of innovation, aggregate growth slows. This is because the distribution of firms shifts towards less productive firms which can afford to survive under the new policy.

In contrast, the baseline economy is largely insensitive to the subsidy on operational costs. There is a decrease in average firm size, accompanied by an increase in the number of firms (and hence employment), but the impact on other margins, including aggregate growth, is negligible. The reason for this stark contrast to the restricted model lies in the firm selection process. While in the restricted model firms select purely on productivity, the baseline model features selection on profitability.

From Section 4.2 we know that selection in fact happens predominantly on demand, rather than on productivity. Figure 7 depicts this by showing how the joint distribution of productivity and demand before (Panel a) and after the policy change (Panel b). In addition to the respective distributions, the figure also plots their respective lower ($\underline{b}$ and $\underline{\hat{q}}$) and upper ($\overline{b}$ and $\overline{\hat{q}}$).

It is apparent from Figure 7 that the distribution of firms mainly shifts towards businesses with lower demand levels, explaining the lower average firm size after the policy implementation. In contrast, the distribution along the productivity dimension

Figure 7: Demand and productivity distribution in baseline model



Note: The figure shows the distribution of demand and (relative) productivity in the simulated model. Darker shades indicate densely populated areas. White areas indicate areas of the state space which are not populated. Panel a) depicts the baseline model without any subsidies. Panel b) depicts the baseline model with a subsidy to operation costs equal to $\tau_\phi = 0.25$. The values $\underline{\hat{q}}$, $\bar{\hat{q}}$, $\underline{b}$ and $\bar{b}$ in both the left and right panels indicate, respectively, the lower and upper bounds of firm-level productivity and the lower and upper bounds of firm-level demand in the baseline with $\tau_\phi = 0$.

changes very little (vertical lines indicate the lower, $\underline{\hat{q}}$, and upper, $\bar{\hat{q}}$, bounds of the distribution without the policy).

Impact of R&D subsidies. The second policy we discuss is one of subsidizing firms' R&D costs. We assume that $\tau_r^b = 0.1$ in the baseline economy. The subsidy is assumed to be $\tau_r^r = 0.055$ in the restricted model, ensuring that the share of output spent on subsidizing businesses is the same as in the baseline model.

Panel B of Table 5 show the long-run impact of R&D subsidies in the baseline economy and the restricted productivity-only counterfactual. The implications of R&D subsidies are qualitatively the same in both models. Supporting firms in their investment into R&D increases innovation rates and, in turn, raises aggregate growth. Faster economic growth, however, comes with higher exit rates as firms which are unsuccessful in their innovation attempts become unprofitable relatively faster. Higher firm exit (especially among younger businesses) tilts the distribution of firms towards older businesses, raising average firm size. Despite this, aggregate employment falls as the economy is inhabited by fewer firms.[41]

[41] We note that our framework is particularly flexible on two important margins of adjustment: firm mass and labor supply. This is because we have assumed free entry of new businesses and a linear

The key difference between the impact of the R&D subsidy in the two economies is quantitative. In particular, the subsidy is much more powerful in the baseline economy where the innovation rate rises by more than 10%. In contrast, the restricted productivity-only economy sees an increase in average innovation rates of only about 3%. These differences are then reflected in the response of aggregate growth to the R&D subsidy which is more than twice as effective in the baseline economy, compared to the productivity-only model. While in the latter aggregate growth increases by about 5%, it rises by 12.6% in the baseline economy.

Intuitively, this result hinges on the distribution of firm values which drive innovation incentives (see Equation 13).[42] While in the restricted model only firm-level productivity determines firm values, in the baseline model it is also demand-side factors which play an important role. Similarly as with expected demand growth being crucial for the level of innovation rates, it is important in determining their sensitivity to changes in marginal costs. These conclusions are similar in spirit to those in Sterk ⓡ al. (2021), who argue that the autocovariance structure of firm-level employment – which we use to estimate our model – is important for disciplining the sources of firm-level heterogeneity which ultimately determine the distribution of firm values.

The above results therefore show that accounting for the distinction between productivity and profitability is important for at least two reasons. First, it changes our understanding of the sources of aggregate growth with *firm-level* demand variation partly driving about *aggregate* economic growth. Second, it has the potential to fundamentally change the policy conclusions derived from models which ignore the distinction between productivity and profitability.

4.5 Empirical support for model mechanism

This subsection provides empirical support for our key channel which links expected demand growth at the firm-level to firms' innovation decisions and productivity growth. Towards this end, we first review existing studies which empirically document this relationship in various settings. Second, we draw on firm-level data from Compustat to estimate this key relationship and show that our model is both qualitatively and quantitatively consistent with the data in this regard.

Review of existing evidence on firm-level demand effects. While the market size effect has been questioned (see e.g. Jones, 1995), a number of existing empirical studies find support of such demand effects influencing innovation at the firm-level.

disutility of labor. However, these two margins are modelled in the same way both in the baseline and in the productivity-only counterfactual.

[42]Indeed, using the restrictions in Section 2.2, it is possible to show that the responsiveness of innovation rates to R&D policies is an increasing function of the level of the innovation rate. The latter, in turn, is a function of firm values.

For instance, there is a range of papers focusing on the pharmaceutical industry which identify significant effects between market size expansions and firms' innovation. Acemoglu and Linn (2004) measure exogenous changes in market size (demand for particular drugs) using demographic trends and show that such changes lead to the entry of new drugs. Finkelstein (2004) studies the effects of health policies aimed at increasing the usage of existing vaccines. The results suggest that such policies alter the expected returns from developing new inoculations, leading to a 2.5-fold increase in the number of vaccine clinical trials for diseases affected by the policies.

Kyle and McGahan (2012) use cross-country variation in patent laws to find that a increased patent protection is associated with higher R&D efforts. Similarly, Dubois et al. (2015) use the worldwide number of deaths from diseases in relevant therapeutic classes to instrument expected market size and find large positive elasticities of changes in expected market size and innovation. Finally, using scanner data from the U.S. retail sector, Jaravel (2019) finds that increasing relative demand leads to increasing product variety.

An empirical study, which is more closely related to the structure of our framework, is by Aghion et al. (2020) who document that firm level innovation rises in response to exogenous demand shocks. They measure the latter using detailed firm-level data on French exporters, utilizing variation in firm-level export destinations and the respective aggregate conditions of the destination economies. Measuring innovation outcomes using priority patents, the paper shows that following shocks to the growth in export destinations, firms respond by patenting more.

The above examples of existing research show that our key channel, namely changes in expected market size spurring innovation at the firm-level, is well documented in the data. However, to the best of our knowledge, it remains an open question whether firm-level demand variation has an impact on *aggregate* innovation and growth. The next section takes our framework to the data in order to addresses precisely this question.

Estimating the demand-productivity link using firm-level data. As a final step in our analysis, we now draw on firm-level data to directly estimate the key channel predicted by our model. In particular, we ask to what extent idiosyncratic expected demand growth affects firms' decisions to innovate and, in turn, their productivity growth.

Towards this end, we draw on firm-level data from Compustat between 1962 and 2019. We exclude observations of financial firms and utilities and consider an unbalanced panel resulting in about 90 thousand firm-year observations with roughly 10.5 thousand unique firms. More details on the data cleaning process, variable definitions, as well as the estimation procedure and additional results can be found in the Appendix.

Using the Compustat data, we proceed in four steps. First, we follow Foster et al.

(2016) and define total factor productivity as

$$\ln(\text{tfp}_{j,t}) = \ln(y_{j,t}) - \alpha_l \ln(l_{j,t}) - \alpha_k \ln(k_{j,t}) - \alpha_m \ln(m_{j,t}), \tag{29}$$

where, in the absence of firm-level price information, $y_{j,t}$ is the revenue of firm j in year t and where l, k and m corresponds to labor, capital, and intermediate inputs (materials), respectively. The elasticities, α, are estimated using the methodology described in Levinsohn and Petrin (2003).[43]

Second, as in Section 3.2 we follow Foster et al. (2008) and use estimated idiosyncratic TFP to proxy for firm-level prices and estimate demand shocks, $\ln(b_{j,t})$, as residuals from the following regression

$$\ln(\text{sales}_{j,t}) = \alpha_q^s \ln(\text{tfp}_{j,t}) + \alpha_x^s \mathbf{X}_{j,t} + \ln(b_{j,t}), \tag{30}$$

where we control for a range of firm-specific and aggregate variables, $\mathbf{X}_{j,t}$, including 3-digit-industry-year fixed effects, dividend value, total assets, and liquidity ratio.[44]

Third, we assume that expected firm-level demand growth, $\mathbb{E}_{t-1}\Delta\ln(b_{j,t})$, corresponds to the predicted values from the following regression

$$\Delta\ln(b_{j,t}) = \alpha_b^b \Delta\ln(b_{j,t-1}) + \alpha_x^b \mathbf{X}_{j,t-1} + \epsilon_{j,t}^b, \tag{31}$$

effectively assuming that firms use mean squared error linear forecasts to form expectations.

The final step in our procedure is to directly estimates the model-predicted link between expected demand growth, $\mathbb{E}_t\Delta\ln(b_{j,t})$, estimated in (31) and productivity growth, $\Delta\ln(\text{tfp}_{j,t})$, estimated in (48). Specifically, we estimate

$$\Delta\ln(\text{tfp}_{j,t}) = \alpha_j^b + \beta_b \mathbb{E}_{t-1}\Delta\ln(b_{j,t}) + \alpha_x^q \mathbf{X}_{j,t} + \epsilon_{j,t}^q, \tag{32}$$

where α_j are firm fixed effects and where the key coefficient of interest is β_b.

Results from firm-level estimation. Table 6 shows the results, focusing on the key coefficient β_b reported in the first column. The estimate is positive and statistically significant. Moreover, conducting the same estimation procedure in the baseline model shows an almost identical coefficient.[45]

[43] The Appendix shows that our results are robust to several alternative specifications, including the use of labor productivity, instead of TFP, and the use of the approach of Olley and Pakes (1996) instead of that in Levinsohn and Petrin (2003).

[44] As documented for example in Cloyne et al. (2018) these variables capture adequately firm's financial constraints. Note that we do not include firm fixed effects, since our model precisely predicts firm fixed effects in demand which we want to capture in our estimation.

[45] While in the data we proxy physical TFP with revenue-based TFP, we do not need to do so in the model. Foster et al. (2008) document that the two are highly correlated. In addition, the model does

Section 3.2 documented that the model-based productivity and demand shocks, driving our heterogeneous firm framework, have realistic properties, closely mimicking those estimated in the data (see Foster et al., 2008). The results presented in this subsection further validate our model and its quantitative analysis. While TFP and demand estimates should always be taken with care, especially when lacking firm-level price data, our results are encouraging. In particular, they suggest that our key mechanism – a link between expected demand growth and productivity growth at the firm level – is not only present in the data, but also quantitatively reasonable.

Table 6: Estimated impact of expected demand growth on productivity growth, β_b

	(I)	(II)
Compustat data	0.060^{**}	0.057^{**}
	(2.48)	(2.19)
Baseline model	0.041	0.041
Industry & year fixed effects	✓	✓
Additional controls, $\mathbf{X}_{j,t}$		✓
Number of observations	93876	93876

Note: Table shows estimates of β_b, the relationship between expected demand growth and productivity growth at the firm level, from (32). The first column does not include additional controls (dividend value, total assets, and liquidity ratio), while the second does. None of these are included in the model regressions. Brackets indicate bootstrapped z-statistics and ** indicates statistical significance at the 5 percent level. Standard errors are calculated by bootstrap, see Appendix E.1 for details.

5 Conclusion

We build on the recent and expanding evidence that demand-side factors are crucial for driving firm-level outcomes. Incorporating this feature into a new model of endogenous growth in which heterogeneous firms innovate and operate based on profitability, instead of productivity alone, we show that firm-level demand variation partly drives aggregate growth. Estimating our model using firm data suggests that this link is quantitatively important and that ignoring the distinction between profitability and productivity can fundamentally alter model predictions about the impact of growth policies.

We believe that our framework opens the door to several intriguing questions which we have left unanswered in our paper. First, what new tools could policy-makers use to spur aggregate growth? Our framework suggests that recognized demand-oriented tools, such as monetary policy, could be used to impact long-run growth. By the same token,

not allow for industry fixed effects or additional controls.

our model could be extended for aggregate uncertainty to analyze the extent to which transitory demand changes affect long-run outcomes.

Second, our model would also be suitable for the study of more unconventional and understudied growth policies, such as public procurement or fiscal transfers. What is the resulting optimal policy mix? While demand-side growth policies have been debated in policy circles (see e.g. European Commission, 2003a,b, 2006), they have been - by and large - missing from systematic academic analysis within state-of-the-art models of endogenous growth.[46]

Finally, can we learn something about the secular trends in productivity and growth, observed in many developed economies, by studying the *demand* side of the economy? Our model suggests that there may be new channels and driving forces at play when it comes to understanding secular stagnation.

[46] A notable exception is Slavtchev and Wiederhold (2016) who study the impact of public procurement within a multi-sector endogenous growth model with heterogeneous innovation step sizes.

References

Acemoglu, Daron, Ufuk Akcigit, Nicholas Bloom, and William Kerr (2018) "Innovation, Reallocation and Growth," *American Economic Review*, 108 (11), 3450–3491.

Acemoglu, Daron and Joshua Linn (2004) "Market Size in Innovation: Theory And Evidence from the Pharmaceutical Industry," *Quarterly Journal of Economics*, 119 (3), 1049–1090.

Aghion, Philipe, Antonin Bergeaud, Matthieu Lequien, and Marc Melitz (2020) "The Heterogenous Impact of Market Size on Innovation: Evidence from French Firm-Level Exports," mimeo.

Aghion, Philippe and Peter Howitt (1992) "A Model of Growth through Creative Destruction," *Econometrica*, 60, 323–351.

——— (1994) "Growth and Unemployment," *Review of Economic Studies*, 61, 477–494.

Akcigit, Ufuk and William Kerr (2018) "Growth through Heterogeneous Innovations," *Journal of Political Economy*, 126 (4), 1374–1443.

Anzoategui, Diego, Diego Comin, Mark Gertler, and Joseba Martinez (2019) "Endogenous Technology Adoption and R&D as Sources of Business Cycle Persistence," *American Economic Journal: Macroeconomics*, 11 (3), 67–110.

Arkolakis, Costas (2016) "A Unified Theory of Firm Selection and Growth," *Quarterly Journal of Economics*, 131 (1), 89–155.

Bachmann, Ruediger and Peter Zorn (2020) "What Drives Aggregate Investment? Evidence from German Survey Data," *Journal of Economic Dynamics and Control*, 115, 103873.

Bartelsman, Eric and Mark Doms (2000) "Understanding Productivity: Lessons from Longitudinal Microdata," *Journal of Economic Literature*, 38, 569–594.

Benigno, Gianluca and Luca Fornaro (2018) "Stagnation Traps," *Review of Economic Studies*, 85 (3), 1425–1470.

Beraja, Martin and Christian Wolf (2021) "Demand Composition and the Strength of Recoveries," mimeo.

Bernard, Andrew, Emmanuel Dhyne, Glenn Magerman, Kalina Manova, and Andreas Moxnes (forthcoming) "The Origins of Firm Heterogeneity: A Production Network Approach," *Journal of Political Economy.*

Blanchard, Olivier, Guido Lorenzoni, and Jean Paul L'Huillier (2017) "Short-Run Effects of Lower Productivity Growth: A Twist on the Secular Stagnation Hypothesis," NBER Working Paper 23160.

Bloom, Nicholas, Rachel Griffith, and John Van Reenen (2002) "Do R&D Tax Credits Work?" *Journal of Public Economics*, 85 (1), 1–31.

Blundell, Richard, Rachel Griffith, and Frank Windmeijer (2002) "Individual Effects and Dynamics in Count Data Models," *Journal of Econometrics*, 108 (1), 113–131.

Broda, Christian and David Weinstein (2006) "Globalization and the Gains form Variety," *Quarterly Journal of Economics*, 121 (2), 541–585.

Brynjolfsson, Erik and Lorin M Hitt (2003) "Computing productivity: Firm-level evidence," *Review of economics and statistics*, 85 (4), 793–808.

Cavenaile, Laurent, Murat Alp Celik, Pau Roldan-Blanco, and Xu Tian (2021) "Style over Substance? Advertising, Innovation, and Endogenous Market Structure," mimeo.

Cavenaile, Laurent and Pau Roldan-Blanco (2021) "Advertising, Innovation and Economic Growth," *American Economic Journal: Macroeconomics*, 13 (3).

Clementi, Gian Luca and Berardino Palazzo (2016) "Entry, exit, firm dynamics, and aggregate fluctuations," *American Economic Journal: Macroeconomics*, 8 (3), 1–41.

Cloyne, James, Clodomiro Ferreira, Maren Froemel, and Paolo Surico (2018) "Monetary policy, corporate finance and investment," NBER Working Paper 25366.

Comin, Diego, Danial Lashkari, and Marti Mestieri (2021) "Structural Change with Long-Run Income and Price Effects," *Econometrica*, 89 (1), 311–374.

De Ridder, Maarten (2020) "Market Power and Innovation in the Intangible Economy," mimeo.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda (2017) "Changing Business Dynamism and Productivity: Shocks vs. Responsiveness," mimeo.

Denicolo, Vincenzo and Piercarlo Zanchettin (2014) "What Causes Over-Investment in R&D In Endogenous Groth Models," *Economic Journal*, 124, 1192–1212.

Dubois, Pierre, Olivier de Mouzon, Fiona Scott-Morton, and Paul Seabright (2015) "Market Size and Pharmaceutical Innovation," *The RAND Journal of Economics*, 46 (4), 844–871.

Efron, B. (1979) "Bootstrap Methods: Another Look at the Jackknife," *The Annals of Statistics*, 7 (1), 1–26, `http://www.jstor.org/stable/2958830`.

Eslava, Marcela and John Haltiwanger (2021) "The Size and Life-Cycle Growth of Plants: The Role of Productivity, Demand and Wedges," NBER Working Paper, 27184.

European Commission (2003a) "Investing in research: an action plan for Europe,"Technical Report 226 2003/2148, Brussels.

——— (2003b) "Raising EU R&D intensity Improving the effectiveness of the mix of public support mechanisms for private sector research and development.,"Technical Report EUR 20713, Luxembourg.

——— (2006) "Creating an Innovative Europe - Report of the Independent Expert Group on R&D and innovation appointed following the Hampton Court Summit and chaired by Mr. Esko Aho.,"Technical Report EUR 22005, Luxembourg.

Evans, David (1987) "Tests of Alternative Theories of Firm Growth," *Journal of Political Economy*, 95, 657–674.

Finkelstein, Amy (2004) "Static and Dynamic Effects of Health Policy: Evidence from the Vaccine Industry," *Quarterly Journal of Economics*, 119 (2), 527–564.

Fornaro, Luca and Martin Wolf (2021) "The Scars of Supply Shocks," mimeo.

Foster, Lucia, John Haltiwanger, and Chad Syverson (2008) "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review*, 98 (1), 394–425.

——— (2016) "The Slow Growth of New Plants: Learning about Demand?" *Economica*, 83 (329), 91–129.

Fujii, Daisuke, Yukiko Saito, and Tatsuro Senga (2017) "The Dynamics of Inter-firm Networks and Firm Growth," RIETI Working Paper, no. 17-E-110.

Gourio, Francois (2014) "Financial Distress and Endogenous Uncertainty," mimeo, Federal Reserve Bank of Chicago.

Gourio, François and Leena Rudanko (2014) "Customer Capital," *The Review of Economic Studies*, 81 (3), 1102–1136.

Grossman, Gene and Elhanan Helpman (1991) "Quality Ladders in the Theory of Growth," *Review of Economic Studies*, 58, 43–61.

Hall, Bronwyn H (1990) "The manufacturing sector master file: 1959-1987,"Technical report, National Bureau of Economic Research.

Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg (2001) "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools," NBER Working Paper 8498.

Hall, Bronwyn and Rosemarie Ziedonis (2001) "The Patent Paradox Revisited: An Empirical Study of Patenting in the US Semiconductor Industry, 1979-1995," *RAND Journal of Economics*, 101–128.

Haltiwanger, John (2012) "Job Creation and Firm Dynamics in the U.S.," *Innovation Policy and the Economy*, 17–38.

Haltiwanger, John, Ron Jarmin, Robert Kulick, and Javier Miranda (2016) "High Growth Young Firms: Contribution to Job, Output and Productivity Growth," mimeo.

Haltiwanger, John, Ron Jarmin, and Javier Miranda (2013) "Who Creates Jobs? Small Versus Large Versus Young," *Review of Economics and Statistics*, 95 (2), 347–361.

Hansen, Gary (1985) "Indivisible Labor and the Business Cycle," *Journal of Monetary Economics*, 16 (3), 309–327.

Hopenhayn, Hugo A. and Richard Rogerson (1993) "Job Turnover and Policy Evaluation: A General Equilibrium Analysis," *Journal of Political Economy*, 101 (5), 915–938, 10.1086/261909.

Hopenhayn, Hugo, Julian Neira, and Rish Singhania (2020) "From Population Growth to Firm Demographics: Implications for Concentration, Entrepreneurship and the Labor Share," mimeo.

Hottman, Collin, Stephen Redding, and David Weinstein (2016) "Quantifying the Sources of Establishment Heterogeneity," *Quarterly Journal of Economics*, 131 (3), 1291–1364.

Howitt, Peter (1999) "Steady Endogenous Growth with Population and R & D Inputs Growing," *Journal of Political Economy*, 107 (4), 715–730.

Jaravel, Xavier (2019) "The Unequal Gains from Product Innovations: Evidence from the U.S. Retail Sector," *Quarterly Journal of Economics*, 134 (2), 715–783.

Jones, Charles (1995) "R&D-Based Models of Economic Growth," *Journal of Political Economy*, 103 (4), 759–784.

Jorda, Oscar, Moritz Schularick, and Alan Taylor (2021) "Disasters Everywhere: The Costs of Business Cycles Reconsidered," NBER Working Paper 26962.

Jovanovic, Boyan (1982) "Selection and the Evolution of Industry," *Econometrica*, 50 (3), 649–670.

Kaas, Leo and Bihemo Kimasa (forthcoming) "Firm Dynamics with Frictional Product and Labor Markets," *International Economic Review*.

Karahan, Fatih, Benjamin Pugsley, and Aysegul Sahin (2021) “Demographic Origins of the Startup Deficit,” mimeo.

Kehrig, Matthias and Nicolas Vincent (2021) “The Micro-Level Anatomy of the Labor Share Decline,” *Quarterly Journal of Economics*, 136 (2), 1031–1087.

Keller, Wolfgang and Stephen Yeaple (2009) “Multinational Enterprises, International Trade, and Productivity Growth: Firm-Level Evidence from the United States,” *The Review of Economics and Statistics*, 91 (4), 821–831, `https://EconPapers.repec.org/RePEc:tpr:restat:v:91:y:2009:i:4:p:821-831`.

Klette, Tor Jakob and Samuel Kortum (2004) “Innovating Firms and Aggregate Innovation,” *Journal of Political Economy*, 112 (5), 986–1018.

Kyle, Margaret and Anita McGahan (2012) “Investments in Pharmaceuticals Before and After TRIPS,” *The Review of Economics and Statistics*, 94 (4), 1157–1172.

Lee, Yoonsoo and Toshihiko Mukoyama (2015) “Productivity and employment dynamics of US manufacturing plants,” *Economics Letters*, 136, 190–193.

Lentz, Rasmus and Dale T Mortensen (2005) “Productivity growth and worker reallocation,” *International Economic Review*, 46 (3), 731–749.

——— (2008a) “An empirical model of growth through product innovation,” *Econometrica*, 76 (6), 1317–1373, Publisher: Wiley Online Library.

——— (2008b) “An empirical model of growth through product innovation,” *Econometrica*, 76 (6), 1317–1373.

Levinsohn, James and Amil Petrin (2003) “Estimating Production Functions Using Inputs to Control for Unobservables,” *The Review of Economic Studies*, 70 (2), 317–341. Ωİmrohoroğlu and Tüzel

İmrohoroğlu, Ayşe and Şelale Tüzel (2014) “Firm-Level Productivity, Risk, and Return,” *Management Science*, 60 (8), 2073–2090, 10.1287/mnsc.2013.1852.

Mukoyama, Toshihiko and Sophie Osotimehin (2019) “Barriers to Reallocation and Economic Growth: The Effects of Firing Costs,” *American Economic Journal: Macroeconomics*, 11 (4), 235–270.

Olley, G. Steven and Ariel Pakes (1996) “The Dynamics of Productivity in the Telecommunications Equipment Industry,” *Econometrica*, 64 (6), 1263–1297.

Perla, Jesse (2019) “A Model of Product Awareness and Industry Life Cycles,” mimeo.

Rachel, Lukasz (2021) "Leisure-Enhancing Technological Change," mimeo.

Rogerson, Richard (1988) "Indivisible Labor, Lotteries, and Equilibrium," *Journal of Monetary Economics*, 21 (1), 3–16.

Rouwenhorst, K Geert (1995) "Asset Pricing Implications of Equilibrium Business Cycle Models," in *Cooley, T.F. (Ed.), Frontiers of Business Cycle Research*: Princeton University Press.

Rozsypal, Filip et al. (2016) "Schumpeterian business cycles," *Mimeo.*

Sedláček, Petr (2019) "Creative Destruction and Uncertainty," *Journal of the European Economic Association*, 18 (4), 1814–1843.

Sedláček, Petr and Vincent Sterk (2017) "The Growth Potential of Startups over the Business Cycle," *American Economic Review*, 107 (10), 3182–3210.

Sedláček, Petr (2019) "Creative Destruction and Uncertainty," *Journal of the European Economic Association*, 18 (4), 1814–1843, 10.1093/jeea/jvz047.

Slavtchev, Viktor and Simon Wiederhold (2016) "Does the Technological Content of Government Demand Matter for Private R&D? Evidence from US States," *American Economic Journal: Macroeconomics*, 8 (2), 45–84, `https://pubs.aeaweb.org/doi/10.1257/mac.20130069`.

Stein, Jeremy (1997) "Waves of Creative Destruction: Firm-Specific Learning-by-Doing and the Dynamics of Innovation," *Review of Economic Studies*, 64 (2), 265–288.

Sterk, Vincent ⓡ Petr Sedláček ⓡ Benjamin Pugsley (2021) "The Nature of Firm Growth," *American Economic Review*, 111 (2).

Wooldridge, Jeffrey M. (2009) "On estimating firm-level production functions using proxy variables to control for unobservables," *Economics Letters*, 104 (3), 112–114, https://doi.org/10.1016/j.econlet.2009.04.026.

Young, Alwyn (1998) "Growth without Scale Effects," *Journal of Political Economy*, 106 (1), 41–63.

Appendices

A Proofs

For what follows, we note that the following expressions will be useful in the proofs presented below

$$
\begin{aligned}
C &= \left(\int_j b_j^{\frac{1}{\eta}} c_j^{\frac{\eta-1}{\eta}}\,\mathrm{d}j\right)^{\frac{\eta}{\eta-1}} = \left(\int_j b_j^{\frac{1}{\eta}} \left[b_j \left[\frac{\eta}{\eta-1}\right]^{-\eta} W^{-\eta} q_j^{\eta} C\right]^{\frac{\eta-1}{\eta}} \mathrm{d}j\right)^{\frac{\eta}{\eta-1}} \\
&= \left[\frac{\eta}{\eta-1}\right]^{-\eta} W^{-\eta} C \left(\int_j b_j q_j^{\eta-1}\,\mathrm{d}j\right)^{\frac{\eta}{\eta-1}}
\end{aligned}
$$

and hence

$$
\begin{aligned}
W^{\eta} &= \left[\frac{\eta}{\eta-1}\right]^{-\eta} \left(\int_j b_j q_j^{\eta-1}\,\mathrm{d}j\right)^{\frac{\eta}{\eta-1}} \\
C &= \upsilon^{-1} \left[\frac{\eta}{\eta-1}\right]^{-1} \left(\int_j b_j q_j^{\eta-1}\,\mathrm{d}j\right)^{\frac{1}{\eta-1}} = \frac{Q(\eta-1)}{\upsilon\eta},
\end{aligned}
\tag{33}
$$

where we used $Q = \left(\int_j b_j q_j^{\eta-1}\,\mathrm{d}j\right)^{1/(\eta-1)}$ and $W = \upsilon C$.

A.1 The decentralized economy

Proof of Proposition 1. Consider the stationarized economy and decisions made by an incumbent with production efficiency q, market size b. Let us define the stationarized idiosyncratic state by $\hat{s} = (b\hat{q}, b, \theta)$, where $\hat{q} = q/Q$, $\hat{s}^+ = (b\hat{q}\theta(1+\lambda)/(1+g), b\theta, \theta)$ and $\hat{s}^- = (b\hat{q}\theta/(1+g), b\theta, \theta)$. Where possible, in what follows we suppress the dependence of endogenous choice variables on the firm-specific state, $\hat{s}$.

Under Assumption 1, and using the labor supply condition (7), optimal pricing condition (12) and the fact that $W = (\eta-1)/\eta Q$, per-period profits are given by

$$
\pi = pc - W(c/q + n_r) = bq\frac{1}{2}\frac{C}{W} - bq\frac{1}{4}\frac{C}{W} - bq\frac{W}{Q}\frac{x^2}{\gamma} = bq\left[\frac{1}{4\upsilon} - \frac{x^2}{2\gamma}\right].
$$

Therefore, stationarized profits are given by

$$
\hat{\pi} = b\hat{q}\left[\frac{1}{4\upsilon} - \frac{x^2}{2\gamma}\right].
$$

We can then write stationarized firm value as

$$
\hat{V}(\hat{s}) = \hat{\pi} + \beta(1-\delta)\left[x\hat{V}(\hat{s}^+) + (1-x)\hat{V}(\hat{s}^-)\right]. \tag{34}
$$

The first-order condition for innovation rate is

$$x = \frac{\gamma}{b\hat{q}}\beta(1-\delta)\left(\hat{V}(\hat{s}^{+}) - \hat{V}(\hat{s}^{-})\right) \tag{35}$$

Let us now guess that the stationarized value function is an affine transformation $\hat{V} = \mathcal{A}b\hat{q}$ and verify our guess using the above optimality conditions. Using our guess yields the following expression for the innovation probability

$$x(\mathcal{A}) = \gamma\frac{\beta(1-\delta)}{(1+g)}\theta\lambda\mathcal{A}(\theta), \tag{36}$$

Note that the innovation rate common across all firms, since it is independent of q and b.

Using the guess in the firm value (34) delivers

$$\mathcal{A}b\hat{q} = b\hat{q}\left[\frac{1}{4\upsilon} - \frac{x^2}{2\gamma} + \mathcal{A}\frac{\beta(1-\delta)}{1+g}\theta\left(1+\lambda x\right)\right],$$

which is linear in $b\hat{q}$ and hence our initial guess is verified. $\mathcal{A}$ is the positive real solution to the following quadratic equation

$$\mathcal{A}(\theta) = \left[\left(\frac{1}{4\upsilon} - \frac{x^2}{2\gamma}\right)\right]\left[1 - \frac{\beta(1-\delta)}{(1+g)}\left(1+\lambda x\left(\mathcal{A}\right)\right)\theta\right]^{-1} > 0.$$

It then also directly follows that

$$x(\mathcal{A}) = \gamma\frac{\beta(1-\delta)}{(1+g)}\theta\lambda\mathcal{A}(\theta) \tag{37}$$

□

Proof of Proposition 2. Firm size is given by $n = n_c + n_r$. Proposition 1 implies that x is constant. Using the demand constraint, we can then write firm size as

$$n(\theta) = b\hat{q}\left[\frac{1}{2\upsilon} + \frac{x(\theta)^2}{\gamma}\right], \tag{38}$$

where the term in square brackets is constant. Firm size is, therefore, proportional to $b\hat{q}$. Therefore, we can write the expected firm size growth rate as

$$\frac{n'(\theta)}{n(\theta)} = \frac{\theta(1+\lambda x(\theta))}{1+g} \tag{39}$$

□

Proof of Proposition 3. In a stationary equilibrium, the in- and outflows have to balance out or in every state $\hat{s} = (b, \hat{q}, \theta)$. The stationary firm density satisfies

$$\begin{aligned}\mu(b, \hat{q}, \theta) = (1-\delta)&\left[\mu\left(\frac{b}{\theta}, \frac{\hat{q}(1+g_q)}{1+\lambda}, \theta\right) + \delta h_e\left(\frac{b}{\theta}, \frac{\hat{q}(1+g_q)}{1+\lambda}, \theta\right)\right]x(\theta)\\ + (1-\delta)&\left[\mu\left(\frac{b}{\theta}, \hat{q}(1+g_q), \theta\right) + \delta h_e\left(\frac{b}{\theta}, \hat{q}(1+g_q), \theta\right)\right](1-x(\theta))\end{aligned}$$

The left-hand side captures outflows, which are simply $\mu(\hat{s})$, while normalized productivity adjusts accordingly. The right-hand sides marks inflows into state state $\hat{s}$ from all surviving firms that either successfully innovated from $\frac{\hat{q}}{1+\lambda}$ or their innovation efforts failed and their productivity remained at $\hat{q}$. We also used the fact that the innovation rate depends only on θ. Finally, the last term captures the new entrants. The entrant distribution satisfies

$$h_e(b, \hat{q}, \theta) \propto \begin{cases} x b^{-(\kappa_n+1)} & \text{if } \hat{q} = \frac{1}{1+g_q} \\ (1-x) b^{-(\kappa_n+1)} & \text{if } \hat{q} = \frac{1+\lambda}{1+g_q} \\ 0 & \text{otherwise} \end{cases}$$

which follows from the premise of the proposition as well as the fact that entrants start with $q = \overline{q}$.

In the simple model, the demand growth rates are common to all firms, and hence we can focus on the marginal distribution $\mu(b, \hat{q}) = \int \mu \, \mathrm{d}\theta$. The innovation rate x is independent of b and $\hat{q}$. Then, the firm distribution is Pareto with shape κ_n. To see this, guess that $\mu(b, \hat{q}) \propto (b\hat{q})^{-(\kappa_n+1)}$. We need to consider three cases, since the mass of entrants is positive only in states $\hat{s} = (b, 1)$. Firstly, for any $\hat{q} \notin \left\{\frac{1}{1+g_q}, \frac{1+\lambda}{1+g_q}\right\}$, there are no entrants and we have

$$\begin{aligned}\mu(b\hat{q}) \propto (b\hat{q})^{-(\kappa_n+1)} &= (1-\delta)\begin{bmatrix} \left(\frac{b}{\theta}\frac{\hat{q}(1+g_q)}{1+\lambda}\right)^{-(\kappa_n+1)} x \\ + \left(\frac{b}{\theta}\hat{q}(1+g_q)\right)^{-(\kappa_n+1)}(1-x)\end{bmatrix}\\ &= (1-\delta)\,(b\hat{q})^{-(\kappa_n+1)}\left[\left(\frac{1+g_q}{(1+\lambda)\theta}\right)^{-(\kappa_n+1)} x + \left(\frac{1+g_q}{\theta}\right)^{-(\kappa_n+1)}(1-x)\right]\\ &\propto (b\hat{q})^{-(\kappa_n+1)}\end{aligned}$$

verifying the guess. Next, consider the case of $\hat{q} = \frac{1+\lambda}{1+g_q}$, hence $\mu(b\hat{q}) \propto \left(b\frac{1+\lambda}{1+g_q}\right)^{-(\kappa_n+1)}$.

It follows that

$$\mu(b\hat{q}) \propto \left(b\frac{1+\lambda}{1+g_q}\right)^{-(\kappa_n+1)} = (1-\delta)\begin{bmatrix} \left(\frac{b}{\theta}\right)^{-(\kappa_n+1)}(1+\delta)x \\ +\left(\frac{b}{\theta}(1+\lambda)\right)^{-(\kappa_n+1)}(1+\delta)(1-x) \end{bmatrix}$$
$$\propto \left(b\frac{1+\lambda}{1+g_q}\right)^{-(\kappa_n+1)}$$

Which verifies the guess. The case of $\hat{q}=\frac{1+\lambda}{1+g_q}$ is analogous.

Finally, note that since the size satisfies $n = b\hat{q}\left[\frac{\hat{W}}{\eta^\eta\upsilon}+\frac{x(\theta)^2}{\gamma}\right]$, it follows that the cdf of the firm size distribution is Pareto with shape parameter κ_n which yields the desired result. □

Proof of Proposition 4. Under Assumption 1, we can express the productivity index as

$$\begin{aligned} Q_t =& \frac{1}{B}\sum_{a=0}\int_{j\in\Omega_a} b_a q_{j,t}dj \\ =& \frac{1}{B}\left[\sum_{a=1}\int_{j\in\Omega_a} b_a q_{j,t}dj + \int_{\hat{q}^e} b_0\hat{q}^e Q_{t-1}(1+\lambda x)\mu_0(\hat{q}^e)\right], \end{aligned}$$

where $\int_{\hat{q}}\mu_a(\hat{q}) = \Omega_a$ is the mass of firms of a particular age a, with $\Omega_{a+1} = (1-\delta)\Omega_a$ and $\Omega_0 = (1-\delta)M = (1-\delta)\int_{\hat{q}}\mu_e(\hat{q})$, and where $\hat{q}^e$ is a time-invariant initial draw of relative productivity. Next period's productivity index can then be expressed as

$$\begin{aligned} Q_{t+1} =& \frac{1}{B}\sum_{a=0}\int_{j\in\Omega} b_a q_{j,t+1}dj \\ =& \frac{1}{B}\left[\sum_{a=0}\int_{\hat{q}} b_a\theta\hat{q}Q_t(1+\lambda x)(1-\delta)\mu_a(\hat{q}) + \int_{\hat{q}^e} b_e\theta\hat{q}^e Q_t(1+\lambda x)(1-\delta)\mu_e(\hat{q}^e)\right] \\ =& Q_t\frac{1+\lambda x}{B}\left[\sum_{a=1}\int_{\hat{q}} b_a\hat{q}\mu_a(\hat{q}) + \int_{\hat{q}^e} b_0\hat{q}^e\mu_0(\hat{q}^e)\right] = Q_t(1+\lambda x) \end{aligned}$$

where the last equality follows from the fact that $B = \sum_a\int_{j\in\Omega_a} b_a\hat{q}_{j,t}dj$. Therefore,

$$1+g = \frac{Q_{t+1}}{Q_t} = 1+\lambda x. \tag{40}$$

□

Aggregate market size effect. Note that the growth rate in the economy does not depend on the scale of the economy N nor the aggregate market size C. The underlying reason is that, as the scale of the economy increases, so does the equilibrium mass of firms and offered variates Ω. As a result, although the firm-level market size rises incentivizing innovation, the consumption is spread more thinly over larger number of products nullifying impact of scale on aggregate growth (Howitt, 1999; Young, 1998).

A.2 The planner's allocation

In what follows, we consider a utilitarian planner, who gives equal weight to all workers, in an economy adhering to Assumption 1.

$$\sum_{t=0}^{\infty} \beta^t \left[\ln C_t - \upsilon N_t\right].$$

We start by noting that, on the BGP, the above objective function can be simplified to[47]

$$\left[\ln C + \ln(1+g)\frac{\beta}{1-\beta} - \upsilon N\right],$$

where $C = \left(\int_j b_j^{\frac{1}{2}} c_j^{\frac{1}{2}} dj\right)^2$ and $N = \int_j n_{j,c} + n_{j,r} dj = \int_j \frac{c_j}{q_j} + \frac{b_j q_j}{\bar{q}} \frac{x_j^2}{\gamma} dj$.

Let us first discuss the static decision of allocating consumption varieties c_j in every period. The planner's allocation, denoted by a star, takes the following form

$$c_j^* = b_j \left(\frac{\upsilon C}{q_j}\right)^{-2} C^*, \tag{41}$$

which gives rise to the following aggregate consumption value[48]

$$C^* = \frac{Q^*}{\upsilon}, \tag{42}$$

where $Q^* = \left(\int_j b_j q_j dj\right)$.

Proof of Proposition 5. Let us now use the results from the previous section and substitute out the optimal allocation of consumption varieties (41) in the planner's problem[49] The planner is left with choosing optimal innovation rates. Given the common growth profiles of demand, this amounts to choosing age- and productivity-specific innovation rates for incumbents, $x_a(\hat{q})$, and for entrants, $x_e(\hat{q}^e)$, to solve:

$$\max_{x_a(\hat{q}), x_e(\hat{q}^e)} \left[\ln Q^* - \ln \upsilon + \ln(1+g)\frac{\beta}{1-\beta} - \upsilon N\right],$$

[47] Note that $\sum_t \beta^t \ln C_t = \sum_t \beta^t \left(\ln C_0 t \ln(1+g)\right) = \frac{\ln C_0}{1-\beta} + \ln(1+g)\frac{\beta}{(1-\beta)^2}$. In addition, N is stationary and therefore its discounted value is equal to $N/(1-\beta)$. Finally, we denote C_0 as C since the composition of the consumption good is fixed over time.

[48] We use (41) in the definition of the consumption bundle to express C^*.

[49] We use the fact that $C^* = Q^*/\upsilon$ to write aggregate labor supply as $N = \int_j c_j^*/q_j + n_{j,r} dj = 1/\upsilon + \int_j n_{r,j} dj$.

where

$$N = \frac{1}{\upsilon} + \sum_{a=0} \int_{\hat{q}} b_a \hat{q} \frac{x_a(\hat{q})^2}{\gamma} \mu_a(\hat{q}) + \int_{\hat{q}^e} b_e \hat{q}^e \frac{x_e(\hat{q}^e)^2}{\gamma} \mu_e(\hat{q}^e), \tag{43}$$

$$1 + g = \frac{Q'}{Q} = \left[\begin{array}{c} \sum_{a=0} \int_{\hat{q}} b_a \theta \hat{q}(1 + \lambda x_a(\hat{q}))(1-\delta)\mu_a(\hat{q}) \\ \\ + \int_{\hat{q}^e} b_e \theta \hat{q}^e (1 + \lambda x_e(\hat{q}^e))(1-\delta)\mu_e(\hat{q}^e) \end{array} \right]. \tag{44}$$

The resulting socially optimal innovation rates are given by

$$x_a(\hat{q}) = x_e(\hat{q}^e) = \frac{\gamma}{2\upsilon} \frac{\beta(1-\delta)}{(1+g)(1-\beta)} \theta \lambda. \tag{45}$$

A.3 The autocovariance structure of employment

Proof of Proposition 6. By premise *i)* of the Proposition, the log-employment of a firm j of age a reads $\ln n_{j,a} = \ln \chi + \ln \hat{q}_{j,a}$. Therefore, we can write the autocovariance of log-employment as

$$\begin{aligned} \text{cov}(\ln n_{j,a}, \ln n_{j,a-h}) &= \text{cov}(\ln \chi + \ln \hat{q}_{j,a}, \ln \chi + \ln \hat{q}_{j,a-h}) \\ &= \text{cov}(\ln \hat{q}_{j,a}, \ln \hat{q}_{j,a-h}) \end{aligned} \tag{46}$$

By premise *ii)* of the Proposition, we know that

$$0 = \text{cov}(\ln \hat{q}_{j,a} - \ln \hat{q}_{j,a-h}, \ln \hat{q}_{j,a-h}) = \text{cov}(\ln \hat{q}_{j,a}, \ln \hat{q}_{j,a-h}) - \text{var}\left(\ln \hat{q}_{j,a-h}\right)$$

This implies $\text{cov}(\ln \hat{q}_{j,a}, \ln \hat{q}_{j,a-h}) = \text{var}\left(\ln \hat{q}_{j,a-h}\right)$. Plugging the intermediate results into (46) delivers

$$\text{cov}(\ln n_{j,a}, \ln n_{j,a-h}) = \text{var}\left(\ln \hat{q}_{j,a-h}\right) = \text{var}(\ln n_{j,a-h})$$

□

Discussion on the class of models satisfying the premise of Proposition 6 Consider a general specification of the productivity process of a firm j of age a. Upon a successful innovation, firm's log-productivity rises by $\ln(1 + \lambda_{j,a})$. Formally,

$$\ln \hat{q}_{j,a+1} = \mu_{j,a} + \ln \hat{q}_{j,a} + \epsilon_{j,a}, \tag{47}$$

where $\epsilon_{j,a}$ are $\ln(1+\lambda_{j,a})$ with probability $x_{j,a}$ and equal to zero with probability $1-x_{j,a}$. This implies that $\mathbb{E}\,\epsilon_{j,a} = x_{j,a}\ln(1+\lambda_{j,a})$.[50]

The drift component $\mu_{j,a}$ captures aggregate forces, for example aggregate growth rate, or any passive technology diffusion process unrelated to firm's current innovation or productivity.

Consider a balanced panel of firms, such that there is no firm exit. Iterating (47) implies

$$\ln\hat{q}_{j,a+h} = \underbrace{\sum_{k=0}^{h-1}\mu_{j,a+k}}_{:=M_{j,a,h}} + \underbrace{\sum_{k=0}^{h-1}\epsilon_{j,a+k}}_{:=E_{j,a,h}} + \ln\hat{q}_{j,a}.$$

Therefore,

$$\operatorname{cov}_{a,h}\left(\ln\hat{q}_{j,a+h} - \ln\hat{q}_{j,a}, \ln\hat{q}_{j,a}\right) = \operatorname{cov}_{a,h}\left(M_{j,a,h}, \ln\hat{q}_{j,a}\right) + \operatorname{cov}_{a,h}\left(E_{j,a,h}, \ln\hat{q}_{j,a}\right)$$

The robust feature of the firm-level data is that $\operatorname{cov}_{a,h}\left(\ln\hat{q}_{j,a+h} - \ln\hat{q}_{j,a}, \ln\hat{q}_{j,a}\right) < 0$. This places empirical restrictions on models of endogenous growth.

Consider a class of models characterized by

- $\mu_{j,a} \equiv \mu$, that is the deterministic drift component is constant; for example aggregate growth rate (as in our model).
- $\lambda_{j,a}$ is an i.i.d. random variable for all j, a with mean $\lambda > 0$, along the lines of randomly determined size of realized quality improvements in Akcigit and Kerr (2018); Rozsypal et al. (2016); Sedláček (2019), for example.
- $x_{j,a+h} \equiv x_j$ is potentially firm-specific but otherwise independent of $\ln q_{j,a}$ for all $h > 0$. This is the case for example in our theoretical model when $\eta = \psi = 2$ and $\theta_{j,t} = \theta$ and in Klette and Kortum (2004); Lentz and Mortensen (2008a).

Then, the components of the cross-sectional age-specific autocovariance are, in turn,

$$\begin{aligned}\operatorname{cov}_{a,h}\left(M_{j,a,h}, \ln\hat{q}_{j,a}\right) &= \mathbb{E}\left[h\mu\ln q_{j,a}\right] - \mathbb{E}\left[h\mu\right]\cdot\mathbb{E}\left[\ln q_{j,a}\right] \\ &= 0\end{aligned}$$

[50] The associated cross-sectional, age-specific variance is therefore $\operatorname{var}_a(\epsilon_{j,a}) = \int x_{j,a}\left(1-x_{j,a}\right)\left[\ln(1+\lambda_{j,a})\right]^2 \mathrm{d}F_j$.

where $\mathbb{E}$ marks cross-sectional mean. By the same token

$$\begin{aligned}\operatorname{cov}_{a,h}\left(E_{j,a,h}, \ln \hat{q}_{j,a}\right) &= \sum_{k=0}^{h-1} \mathbb{E}\left[\epsilon_{j,a+k} \ln q_{j,a}\right] - \sum_{k=0}^{h-1} \mathbb{E}\,\epsilon_{j,a+k}\, \mathbb{E}\left[\ln q_{j,a}\right] \\ &= 0,\end{aligned}$$

where the second line follows because $\lambda_{j,a+h}$ and $x_{j,a+h}$ are independent of $\ln q_{j,a}$ for all j and all $h > 0$ and, hence, $\operatorname{cov}_{a,h}\left(\ln \hat{q}_{j,a+h} - \ln \hat{q}_{j,a}, \ln \hat{q}_{j,a}\right) = 0$.

Proposition 6, therefore, shows that a large class of endogenous growth models – without additional extensions – cannot match the empirical autocovariance structure of log-employment.

B Solution and estimation of baseline model

B.1 Solution method

The aggregate state of the economy consists of aggregate growth rate, g, consumption, C, and wage, W, which are all time-invariant, since we consider a balanced growth path equilibrium. The firm-specific state consists of logs of the current accumulated demand, $\ln b$, relative productivity, $\ln \hat{q}$, permanent component of the demand profile, $\ln \bar{u}$, and the current value of autoregressive demand shock, $\ln v$. We discretize the four dimensions of the state space using 41, 101, 21, and 31 equidistant grid points, respectively. These relatively large grids allow us to capture the full extent of the firm lifecycle dynamics induced by the evolution of demand. The grid for $\ln b$ spans the interval $[-5, 7]$. The choice of the upper and lower bound is dictated by the desire to replicate the empirical size distribution of firms (see Section 3.4 for details). We opt for a relatively dense productivity grid as we want to capture adequately the discrete productivity jumps by factor $1 + \lambda$ while at the same time allowing for cross-sectional productivity dispersion consistent with the empirical evidence. The resulting equidistant grid for productivity ranges from -0.5 to 0.99 and generates TFPQ and TFPR dispersion in line with existing literature and our estimates based on Compustat data (see Section 3.4 in the main text and Section E.3 below). The grid for $\ln \bar{u}$ spans the interval $[-3, 4]$ chosen to capture the rich pattern of log size autocovariance in the data, as illustrated in Figure 1. We discretize the AR(1) process for $\ln v$ using the Rouwenhorst (1995) method.

We use value function iteration to solve numerically the quantitative model. We use a global solution method to capture adequately non-linearities in firms' policy function and complex interactions between expected demand growth, innovation, and productivity improvements. Given the current guess for the value function on each grid point $V(s)$, we can solve the firm's problem in closed form. The CES demand structure lends itself to simple characterization of the pricing decision. This exit choice is described by the following cutoff rule (see Section 3.1)

$$\max\left\{0, \mathbb{E}_{t-1} V(s_t) - W_t\phi\right\}.$$

Given out guess for $V(s_t)$ and the distributional assumptions on the shocks processes, we can compute $\mathbb{E}_{t-1} V(s_t)$. The first-order condition for the optimal R&D expenses (13) implies the following innovation probability

$$x = \left[\gamma \frac{PC}{\psi W pc} \mathbb{E}\beta(1-\delta)\left[V^c(s^+) - V^c(s^-)\right]\right]^{\frac{1}{\psi-1}}$$

where V^c marks the firm's value conditional on survival in the next period. The pricing decision together with R&D expenses yields the total labor demand. Given those choices we can compute the implied new value function at each grid point using formula (22). We use the implied value as a new guess for the value function. We repeat the above procedure until the Euclidean distance between the current guess and the implied new value function is less than 10^{-6}.

In the baseline model we normalize wage W and consumption C to unity, using the disutility of labor and mass of potential entrants as normalizing constants. The remaining aggregate state - the aggregate growth rate g - is a part of estimation targets as described in Section B.2 below.

When searching for new equilibria corresponding to the policy experiments in Section 4.4, we iterate on aggregate growth, consumption, and wages until convergence. That is, given a guess for these aggregate states, we solve the firm problem using value function iteration, we simulate the economy and calculate the implied values for g, C, and W. We iterate on the values for the aggregates state until convergence.

B.2 Simulation and estimation

To estimate moments pertaining to firm lifecycle dynamics or aggregate economic growth, we need to approximate the stationary firm distribution. Towards this end, we use on-grid stochastic simulation. We draw potential entrants from the initial distribution $H(s_e)$ over the idiosyncratic states. We determine which entrants continue operating based on their exit decision as described Section 3.1. We draw new potential entrants up to a point in which we have 100 thousand surviving startups. From then on, we simulate the lifecycle of each firm for 51 years. This simulation length is roughly in line with the empirical sample length in the BDS and Compustat datasets. Moreover, as the demand process settles eventually at $\frac{\bar{u}}{1-\rho}$, there is little to be gained in terms of precision when allowing for a longer lifespan.

In each period, firms face demand and innovation shocks, and make decision regarding pricing, employment, investment into R&D, and exit. Based on these choices, firms move between grid points along three dimensions of the statespace (the firm-specific demand parameter $\ln \bar{u}$ is determined at entry and fixed over the firm's life).

While the firm is alive, in each period we simulate the following steps. At the beginning of each period, before any shock is realized, firms decide whether to pay the fixed cost and continue operating or avoid paying the cost and exit. Firms that decided to continue operating receive demand shocks z and ϵ, and observe the realization of the innovation shock determining their current period productivity. The probability of a successful innovation was determined by R&D expenses in the previous period. Next, firms decide on investment into R&D based on the first-order condition. Then, the firm transitions to the next period, decides on exit, and so on.

The resulting unbalanced panel of firms serves as the data source behind all our quantitative experiments in Section 3 and is used to compute the relevant moments for estimation as described in Section 3.2. We estimate 14 model parameters that pertain to the demand and productivity processes, the exogenous exit rate and the operational cost, as described in the main text. Towards this end, we minimize the equally-weighted Euclidean distance between 250 moments in the model and in the data. Our targets can be grouped into four sets: (i) average growth of real GDP (1 moment), (ii) the firm size life-cycle profile (20 moments), (iii) the firm exit life-cycle profile (19 moments) and (iv) the upper triangle of the autocovariance matrix of log-employment, by age and for a balanced panel of firms surviving up to at least the age of 19 years (210 moments).

C Restricted "productivity-only" model

As mentioned in the main text, the restricted "productivity-only" version of our baseline model is obtained by assuming $\sigma_z = \sigma_\epsilon = \sigma_{\overline{u}} = \sigma_u = \sigma_v = 0$. The latter results in firm-level demand being common and constant across all businesses.

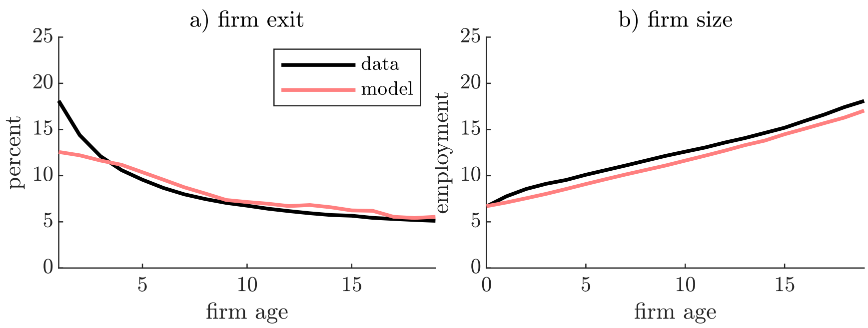
We parametrize the restricted model in exactly the same way as the baseline, with the exception that we do not target the autocovariance structure of employment. This because, as we show in 3.2, the productivity-only model would not be able to match the empirical patterns. Finally, in order for the model to match the steep decline in exit rates by age, we allow for the mean of the initial draws of productivity μ_q to differ from 0. Instead, we target the aggregate R&D to output ratio obtained in the baseline model, as is common in the literature. Table 7 shows the model parameters and Figure 8 depicts the model fit.

Table 7: Parameter values of restricted "productivity-only" model

parameter		value	parameter		value
β	discount factor	0.970	$\sigma_{\overline{u}}$	$\overline{u}$, standard deviation	0
η	elasticity of substitution	6.000	$\mu_{\overline{u}}$	$\overline{u}$, mean	3.600
υ	disutility of labor	1.000	σ_u	$\ln u$, standard deviation	0
κ	entry cost	0.475	ρ_u	$\ln u$, persistence	0
ϕ	fixed cost of operation	0.730	σ_v	$\ln v$, standard deviation	0
δ	exogenous exit rate	0.020	ρ_v	$\ln v$, persistence	0
γ	R&D efficiency	0.260	σ_ϵ	ϵ, standard deviation	0
λ	innovation step size	0.147	σ_z	$\ln z$, standard deviation	0
			σ_q	q_e, standard deviation	0.200
			μ_q	q_e, mean	−0.505

Note: β, η, υ and κ are calibrated as discussed in the main text. The remaining parameters are set such that the model matches the empirical age profiles of average size, exit rates from startup (age 0) to age 19 in the BDS and the aggregate R&D to output ratio in the baseline model.

Figure 8: Model fit of restricted "productivity-only" model



Note: The Figure shows average firm size (employment) and exit rates by age in the restricted model and the BDS data.

D Details on the endogenous demand extension

In this section we provide details of the model extension in which we endogenize the evolution of firm-level demand. After observing the realization of the demand shock, $\epsilon_{j,a}$, firms can decide whether to accept this new draw or reject it. However, in order to accept the new shock realization, firms must pay a fixed adjustment cost. Formally, the decision problem of a firm j of age a can be described as follows. The transitory component of firm-level demand evolves as

$$\ln v_{j,a} = \begin{cases} \rho_v v_{j,a-1} + \epsilon_{j,a} & \text{if realization accepted} \\ \rho_v v_{j,a-1} & \text{otherwise} \end{cases}$$

Firms accept or reject the shock so as to maximize their value. Let ϕ_ϵ be the fixed cost associated with the acceptance of the realization of the demand shock. Firm accepts the new shock $\epsilon_{j,a}$ if and only if the new shock realization increases firm value net of the fixed

Table 8: Parameter values of "endogenous-demand" model

parameter		value	parameter		value
β	discount factor	0.970	$\sigma_{\overline{u}}$	$\overline{u}$, standard deviation	1.249
η	elasticity of substitution	6.000	$\mu_{\overline{u}}$	$\overline{u}$, mean	-1.523
υ	disutility of labor	1.000	σ_u	$\ln u$, standard deviation	1.290
κ	entry cost	0.319	ρ_u	$\ln u$, persistence	0.395
ϕ	fixed cost of operation	0.347	σ_v	$\ln v$, standard deviation	0.576
δ	exogenous exit rate	0.022	ρ_v	$\ln v$, persistence	0.941
γ	R&D efficiency	0.144	σ_ϵ	ϵ, standard deviation	0.295
λ	innovation step size	0.149	σ_z	$\ln z$, standard deviation	0.163
ϕ_ϵ	demand adj. cost	1.292	σ_q	q_e, standard deviation	0.055

Note: β, η, υ and κ are calibrated as discussed in the main text. ϕ_ϵ is calibrated such that the aggregate cost paid by all adjusting firms amounts to 1% of their sales. The remaining parameters are set such that the model matches the empirical age profiles of average size, exit rates from startup (age 0) to age 19 in the BDS and the aggregate R&D to output ratio in the baseline model.

cost, i.e

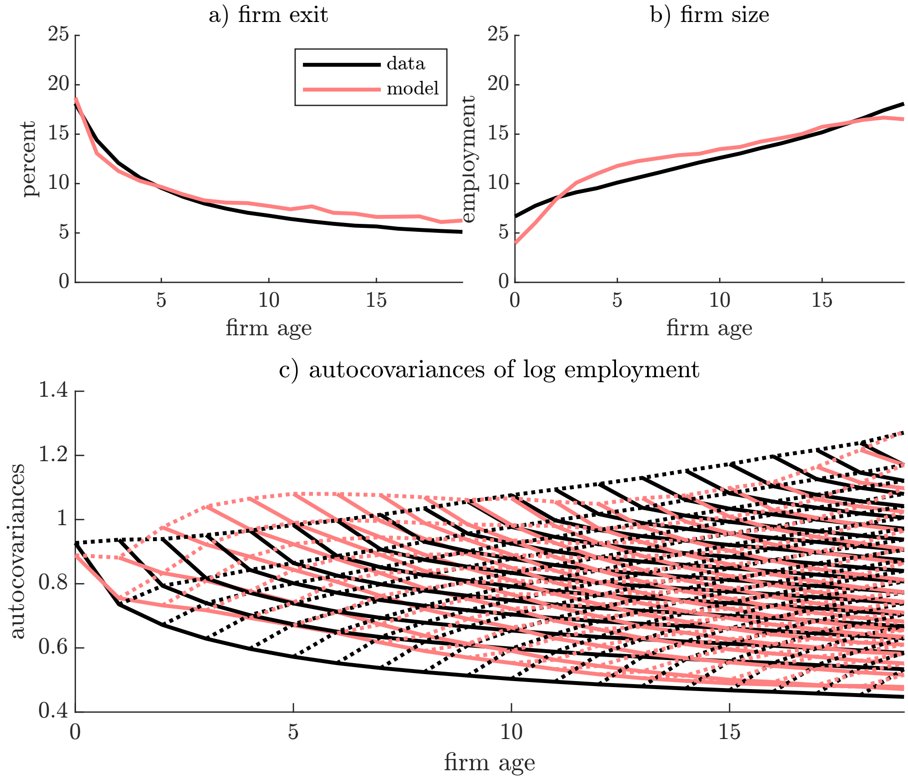
$$V(\rho_v v_{j,a-1} + \epsilon_{j,a}, \ldots) - \phi_\epsilon \geq V(\rho_v v_{j,a-1}, \ldots).$$

We recalibrate the extend model to match the same targets as in the baseline. We calibrate the value of adjustment costs ϕ_ϵ such that the aggregate cost paid by all adjusting firms amounts to 1% of their sales. Table 8 collects the values of calibrated parameters while Figure 9 illustrates that the model fits data very well. Lastly, Figure 10 shows that the model with endogenous demand accumulation matches (non-targeted) size distribution of firms across age categories equally as well as the baseline version.

We then perform the same growth decomposition as in the baseline economy (see Table 4 and the text in Section 4.3 for the details of the exercise). The firm-level demand growth is responsible for 28% of aggregate economic growth. In comparison with the baseline economy, the impact of firm-level demand variation on aggregate growth is slightly stronger (28% vs 22% in the baseline model).

The reason for this is that the adjustment cost on demand accumulation prevents low-growth firms from accumulating demand. Therefore, in order for the model to match the observed lifecycle patterns, more weight is given to businesses with high expected demand growth (and therefore higher innovation rates). Notice that compared to the baseline model, the version with endogenous demand accumulation requires both a higher mean and wider dispersion of the permanent demand component $\overline{u}$ (compare Tables 8 and 1). "Switching" off this channel then comes at the expense of a larger drop in aggregate growth.

Figure 9: Model fit of "endogenous-demand" model



Note: The top two panels in the Figure show average firm size (employment) and exit rates by age in the endogenous-demand model and the BDS data. The bottom panel shows the observed and model-implied autocovariance matrix of log employment for a balanced panel of firms surviving at least up to age 19.

E Details on estimation of the demand-productivity link using Compustat data

In this section we provide more details on the procedure used to estimate the link between firm-specific demand growth and subsequent productivity improvements.

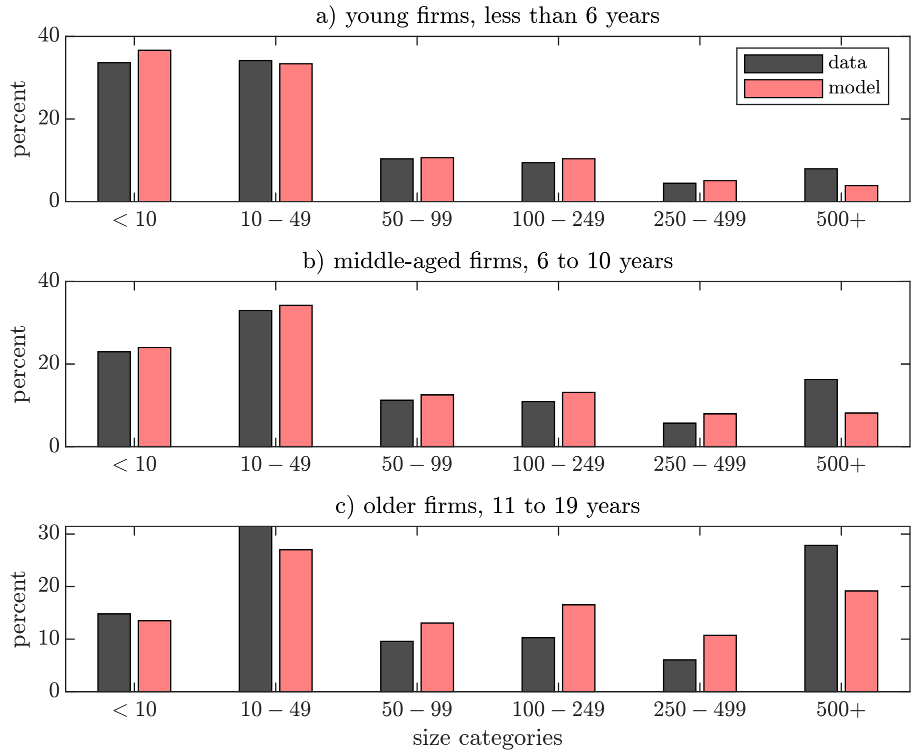
E.1 Production function estimation

W Using the Compustat data, we proceed in four steps. First, we follow Foster et al. (2016) and define total factor productivity as

$$\ln(\mathrm{tfp}_{j,t}) = \ln(y_{j,t}) - \alpha_l \ln(l_{j,t}) - \alpha_k \ln(k_{j,t}) - \alpha_m \ln(m_{j,t}), \tag{48}$$

where, in the absence of firm-level price information, $y_{j,t}$ is the revenue of firm j in year t and where l, k and m corresponds to labor, capital, and intermediate inputs (materials), respectively. The choice of production inputs is endogenous and more productive firms are likely to use more production factors. As a result, OLS estimates of the production function may not recover the true underlying factor elasticities. The Levinsohn and

Figure 10: Non-targeted moments in the data and the “endogenous-demand” model



Note: Size-age distribution in the BDS data and in the endogenous demand model. Both distributions are expressed as shares of employment in a given age category (young: < 6 years, middle-aged: $6 - 10$ years and older: $11 - 19$ years)..

Petrin (2003) (LP) approach is to use data on intermediate inputs to control for this simultaneity problem. The procedure entails an instrumental variable estimation of the elasticities in which the explanatory variables are instrumented with regressors that are correlated with the inputs but uncorrelated with current technology shocks. The LP approach uses lagged values of capital and materials as these instruments. While using LP approach as a baseline, we make sure that our results are consistent across various alternative ways to estimate the production function.

Specifically, we consider Olley and Pakes (1996) approach, who opt for using investment, rather than materials, as an instrument in the first stage of the production function estimation. However, as observed by Levinsohn and Petrin (2003), non-convexities in the capital adjustment costs lead to inaction and kinks in the investment function and can affect responsiveness of investment to the transmitted shocks. For this reason, we use the approach of Levinsohn and Petrin (2003) as our baseline, but we show that our results are virtually the same when using Olley and Pakes (1996) approach, estimated in rolling annual windows as proposed and applied to Compustat data by İmrohoroğlu and Tüzel (2014). Our results are materially unchanged when using Wooldridge (2009) approach who proposes a procedure to estimate the system of equations in Levinsohn and Petrin (2003) in a more efficient, single step procedure. For reference, we also provide the results when production function is estimated via plain OLS. Finally, we consider a simple measure of labor productivity defined as the revenues per employee. In the latter case, the results are of the same sign as in the baseline estimation (and the same sign as in the model), albeit quantitatively much larger.

Bootstrap The capture adequately the uncertainty in our estimates of β_b in (32), we use a bootstrap approach (Efron, 1979). We resample firms with replacement from our original data set to create bootstrap samples containing the same number of firms as the original data. Given a bootstrap sample, we perform all four steps of our empirical procedure: we estimate TFP via equation (48); we estimate demand as a residual from equation (30); we instrument expected demand growth with its past value and controls as in (31); we estimate the elasticity of productivity growth with respect to expected demand growth, eq. (32). We repeat the procedure 500 times and use the resulting distribution of estimated coefficients $\hat{\beta}_b$ to approximate the z-statistic.

E.2 Data

The main data source is Compustat. We use the Compustat fundamental annual data from 1962 to 2019. We exclude observations of financial firms (SIC classification between 6000 and 6999) and utilities (SIC classification 4900-4999).[51] We consider only companies incorporated in the U.S. and traded in the U.S. stock exchange. We exclude firms that have negative entries for sales, total assets, number of employees, depreciation, accumulated depreciation, as well as gross property, plant, and equipment. We remove observations with negative values for our estimates of value added or materials (to be described in detail below). If for a given firm-year pair there are missing values for sales, employment, total assets, or capital, we impute the missing values by linearly interpolating between values recorded by the firm in the two adjacent years. The observations that cannot not be interpolated in this way due to missing data are dropped from the sample. The resulting sample is an unbalanced panel and in our baseline regressions we utilize around 90 thousand firm-year pairs with approximately 10.5 thousand unique firms.

Vale added is defined as net sales from Compustat (Compustat variable name SALE) net of materials. As in for example İmrohoroğlu and Tüzel (2014); Keller and Yeaple (2009), we measure materials measured as total expenses minus labor expenses. Total expenses is approximated as sales net of operating income before depreciation and amortization (OIBDP). Labor expenses is calculated by multiplying the number of employees (EMP), our measure of the stock of labor, by average wages from the Social Security Administration. The nominal values such as sales, assets, materials, or other expenses

[51]As is standard in the literature, we remove financial firms because the balance reporting conventions are very different for those companies. We delete utilities because it is a non-representative and regulated sector the latter.

are deflated with GDP price deflator from the Bureau of Economic Analysis. To further account for industry-specific time trends in prices, in all our regressions we control for fixed effects corresponding to each pair of 3-digit industry and year.[52]

Capital stock is given by gross property, plant, and equipment (PPEGT) from Compustat, deflated by the price deflator for non-residential fixed investment deflator from the Bureau of Economic Analysis. Since investment is made at various times in the past, we follow İmrohoroğlu and Tüzel (2014) - who build on the methods of Hall (1990) and Brynjolfsson and Hitt (2003) - and calculate the average age of capital at every year for each company. Then, we apply the appropriate deflator corresponding to the year the investment was made as implied by the calculated average capital stock age. The latter is calculated by dividing accumulated depreciation (DPACT) by current depreciation (DP). As in İmrohoroğlu and Tüzel (2014), age is smoothed by means of 3-year moving average.[53] Due to the balance sheet reporting convention, we lag the resulting capital stock by one period to reflect the capital stock available for production at the beginning of the period.

When comparing the results in data with model-implied coefficients, we follow the same steps as employed in the data whenever possible. To construct the model-implied counterpart of TFPQ, we use difference between log output and log employment.[54] Given the values for TFPQ we run the same regressions as in the data (see Section 4.5 in the main text). The only difference being that we do not include year-industry fixed effects, as we consider a representative industry in a stationary equilibrium nor the additional controls measuring firm's lifecycle position or financial constraints as there are no confounding factors in our model-simulated data.

E.3 Descriptive statistics

Next, we discuss the descriptive statistics concerning the estimated values of revenue productivity and firm-specific demand in the data and in the model. The estimated TFPR is defined in eq. (48) and estimated as described above, while the demand is the residual from estimated equation (30). Table 9 presents the statistics. The properties of productivity and demand in the model and data are very similar. The estimated demand and productivity process is slightly less dispersed in the data than in the model. The

[52]In our baseline estimation, we do not utilize deflators and wages for individual industries. Detailed deflators and wages at the 4-digit SIC code level are available for manufacturing firms in the NBER Productivity Database. However, restricting our attention to manufacturing firms severely limits the sample size. The use industry-time fixed effects alleviates the need for industry-specific deflators to some extent. We made sure that our results are robust to using industry-specific deflators and wages in the restricted sample.

[53]If the firm's history at a given point in time is shorter than three years, we use all available years.

[54]In the model there are two types of employees: production workers and researchers. Because of that, to recover the underlying TFPR$\hat{q}_{j,t}$ we would need to subtract only the production workers from the log output. However, to be consistent with measurements in the data, we use the total employment.

Table 9: Descriptive statistics of the estimated TFPR and demand.

variable	mean	standard dev.	90-10 pct range	autocorrelation
			Panel A: data	
$\ln \hat{b}_{j,t}$	0	0.41	0.99	0.90
$\ln \mathrm{tfpr}_{j,t}$	0.004	0.18	0.35	0.81
			Panel B: model	
$\ln \hat{b}_{j,t}$	0	1.08	2.66	0.92
$\ln \mathrm{tfpr}_{j,t}$	-0.07	0.21	0.42	0.83

Note: Table shows descriptive statistics for the estimated TFPR and demand processes. Column titled *90-10 pct range* contains the difference between value of 90th and 10th percentile.

reason may be that the firms in Compustat tend be larger and older that the population of firms in the U.S. on which the model is estimated. The mean log demand is zero, since it is estimated as a residual from regression (30).

E.4 Robustness to alternative approaches to TFPR estimation

In this section we compare the value of estimated parameter of interest – the elasticities of the productivity growth with respect to expected demand growth – across various alternative measures of the total factor productivity. We consider Olley and Pakes (1996) method following İmrohoroğlu and Tüzel (2014) who utilize rolling-window annual estimates of production function (effectively allowing the elasticities to vary over time). In addition, we consider Wooldridge (2009) method, simple OLS, and a measure of labor productivity defined as revenues per employee. All these methods paint very similar picture: the estimated firm-specific demand is associated with the subsequent productivity improvements. The effects are both economically and statistically significant in all cases. The magnitude of the effect is very similar to the one implied by the model. The one exception is the case in which we use revenues per employee as a measure of labor productivity. In this case the magnitude is much larger than in the model and the other approaches to approximate TFPQ. However, this measure of labor productivity may not capture adequately the underlying TFPQ.

We also consider a specification in which production function elasticities are estimated separately for each 3-digit industry. Table 11 illustrates that the results are virtually unchanged. In this robustness check, we do not use İmrohoroğlu and Tüzel (2014) approach, since slicing the sample by 3-digit industry in addition to performing annual rolling window estimates of TFP results in small sample sizes and noisy estimates.

Table 10: Estimated impact of expected demand growth on productivity growth, alternative approaches to TFPR estimation.

	İmrohoroğlu and Tüzel (2014)	Wooldridge (2009)	OLS	labor prod.
β_b	0.133***	0.0878***	0.0565***	0.336***
	(9.44)	(14.66)	(11.22)	(31.92)
Industry & year fixed effects	✓	✓	✓	✓
Additional controls, $\mathbf{X}_{j,t}$	✓	✓	✓	✓
Number of observations	93876	93876	93876	93876

Note: Table shows estimates of β_b, the relationship between expected demand growth and productivity growth at the firm level, from (32). Each column corresponds to an alternative method of estimating TFPR; see the text for their description. Brackets indicate non-bootstrapped t-statistics and *** indicates statistical significance at the 1 percent level.

Table 11: Estimated impact of expected demand growth on productivity growth, industry-specific elasticities.

	Levinsohn and Petrin (2003)	Wooldridge (2009)	OLS
β_b	0.0983***	0.115***	0.134***
	(19.81)	(19.57)	(26.18)
Industry & year fixed effects	✓	✓	✓
Additional controls, $\mathbf{X}_{j,t}$	✓	✓	✓
Number of observations	43620	43597	43660

Note: Table shows estimates of β_b, the relationship between expected demand growth and productivity growth at the firm level, from (32). Each column corresponds to an alternative method of estimating TFPR; see the text for their description. Brackets indicate t-statistics and *** indicates statistical significance at the 1 percent level.