The wealth of stagnation:
Falling growth, rising valuations[*]

James D. Paron[†]

January 4, 2025

Abstract

Over the last half-century, economic growth stagnated but stock-market wealth boomed. I present evidence that declining innovation productivity reconciles these trends. At the macro level, I document that R&D spending has fallen relative to value, while M&A spending has doubled relative to R&D. At the micro level, most of the increase in aggregate valuation ratios is explained by a reallocation of sales shares toward high-valuation firms. Using a Schumpeterian model of growth and asset prices, I find that declining innovation productivity explains these facts. When innovation productivity falls, R&D falls and M&A rises. This concentrates production into the hands of the most efficient (high-valuation) incumbents, causing aggregate value to boom. Quantitatively, this explains most of the decline in growth and the rise in valuations. It also helps explain other salient trends, including declining firm entry, rising concentration, and falling interest rates. While stock-market wealth boomed, the present value of consumption (consumer welfare) stagnated with output.

Firms have a greater incentive to innovate when the present value of new profits is high. Over the last fifty years, this relationship seems to have broken down. The value of the stock market has far outpaced output, suggesting that incentives to innovate have reached historic highs. And yet, over the same period, aggregate economic growth has fallen. How can we reconcile stagnating economic growth with a booming stock market?

I present evidence that one structural shift—declining innovation productivity—explains both of these trends. Intuitively, when innovation gets harder, firms invest less in R&D and economic growth slows. This accords with evidence from the growth literature that new ideas have gotten harder to find (Gordon, 2016; Bloom, Jones, Van Reenen, & Webb, 2020).

Declining research productivity also causes the valuation of the stock market to rise relative to output. This rise comes from a reallocation of market share toward the most productive firms in the economy. The intuition is as follows. More productive firms earn higher profits, grow faster, and therefore have higher valuations relative to sales and cashflows. They also grow through M&A, which creates value by reallocating ideas to more efficient producers. As research productivity falls, high-productivity incumbents spend relatively more on M&A than R&D, concentrating production in their hands. Concurrently, the risk of creative destruction by entrants falls, so these incumbents become entrenched. The reallocation-driven growth of incumbents comes at the expense of labor income and new entrants, so stock-market wealth booms as the rest of the economy—including welfare—stagnates.

I present evidence of this mechanism from two sources. First, using firm-level micro data, I lay out empirical evidence of declining innovation productivity and a reallocation of sales toward high-valuation firms. Second, using a Schumpeterian growth model, I estimate the decline in innovation productivity and show that the estimated decline explains a substantial fraction of the decline in growth and the rise in valuations over the last fifty years.

I first document empirical evidence that firms have shifted away from innovation and toward reallocation. Two stylized facts point to this shift. First, over the last fifty years, aggregate R&D expenditures have fallen significantly relative to market value. R&D spending has not kept pace with the rising present value of profits, consistent with declining innovation productivity driving a wedge between these series. Second, total M&A spending has risen substantially relative to R&D spending. Over time, firms have been growing less and less

through innovation and increasingly through reallocation.

I next show that most of the secular rise in the valuation of the aggregate market is explained by a reallocation of market share toward high-valuation firms. I arrive at this conclusion by decomposing changes in aggregate valuation ratios (value relative to sales or cashflows) into two components: within-firm changes, holding fixed the composition of the market (e.g., sales shares); and compositional changes, holding fixed firm-level valuations. Over the last fifty years, firm-level valuations have been relatively stable.[1] The long-run increase in aggregate value is instead a consequence of firms with high valuation ratios gaining market share. Importantly, high-valuation firms have gained market share not only in terms of value, but also in terms of *sales*, so this is a reallocation of real economic activity.[2] This new finding poses a challenge to representative-firm explanations of rising valuations, which necessarily imply that all of the valuation increase came from within firms.

I investigate and rule out alternative explanations for these stylized facts. The first alternative is that R&D has been outsourced to private firms, which are increasingly being acquired by public firms. I find no evidence of an increase in private-firm R&D intensity or of an increase in M&A spending by public acquirers on private targets. A second alternative is that the reallocation between firms was caused by shifting sectoral composition—for example, tech displacing construction. I find that almost all of the reallocation was not between sectors, but between firms within sectors. In fact, in support of the main mechanism, I find that those sectors with the largest decline in R&D-to-value also saw the largest increase in M&A and valuation ratios and the most reallocation to high-valuation firms.

I build a Schumpeterian growth model that incorporates the tradeoff between R&D and M&A and use it to study valuations. Firms compete in product markets and grow by inventing new ideas (R&D), in the spirit of Klette and Kortum (2004). I introduce three ingredients into this framework. The first new ingredient is the buying and selling of existing ideas (M&A). M&A makes possible the exchange of intellectual property between firms

---

[1]At higher frequencies, within-firm changes explain most aggregate movements. This includes the collapse and recovery of valuations around 1980 and the dot-com boom around 2000.

[2]In support of the hypothesis that this was also a reallocation toward high-growth firms, I document that high-valuation firms tend to be high-R&D firms in the cross-section. An analogous decomposition of aggregate R&D-to-sales reveals that within-firm R&D-to-sales has been falling, but high-R&D firms have been gaining market share, consistent with both declining innovation productivity and reallocation.

and leads to a tradeoff between innovation and acquisition. The second new ingredient is heterogeneous firm-level productivity: high-productivity firms can produce goods at a lower marginal cost, giving them a competitive advantage in product markets. In equilibrium, high-productivity firms buy low-productivity firms via M&A. The third ingredient is population growth and new-variety creation (Peters & Walsh, 2022). This feature endogenizes the degree of creative destruction in the economy, which turns out to be critical for quantifying asset pricing implications, as firms' valuations rise when the risk of creative destruction falls.[3]

The assumption of heterogeneous firm-level productivity yields two equilibrium results that are key for understanding reallocation. First, high-productivity firms have higher valuation ratios. This is because high-productivity firms are more profitable—and, specifically, have more profitable growth opportunities—so they choose to invest more in R&D and arise as high-growth firms. Second, high-productivity firms are the only firms that invest in M&A. Consequently, M&A transfers ideas from less productive sellers to more productive buyers, and the surplus from this exchange creates value for the aggregate market.

In the model, a decline in research productivity causes M&A spending to rise relative to R&D spending, reallocating market share toward high-valuation firms. Aggregate M&A-to-R&D rises both because all firms reduce R&D and because high-productivity firms shift skilled labor from R&D to M&A. Consequently, high-productivity firms start to buy up low-productivity firms faster than low-productivity firms grow by inventing new ideas. This reallocation concentrates production into the hands of high-valuation firms.

I use the model to quantify how much of the long-run historical trends are explained by declining innovation productivity. I proceed in three steps. First, I estimate the model using micro data from before the 1980s, assuming historically balanced growth. Second, I estimate the path of innovation productivity from the time series of R&D-to-value. And third, I compare the model's predicted transition path for economic trends—including growth and valuations—with the actual trends in the data.

I find that the decline in R&D-to-value since 1975 implies an estimated 47% decline in innovation productivity—in other words, innovation productivity fell almost in half. The

---

[3]Beyond valuations, Luttmer (2011) shows that this is essential for matching the firm-size distribution. It also implies a long-run link between population and economic growth (C. I. Jones, 1995, 2022a).

intuition for identification is as follows. Firms spend more on R&D if either innovation productivity is high (ideas are easy to find) or the present value of successful innovation is high (ideas are valuable). Dividing aggregate R&D by value identifies long-run changes in research productivity because it nets out the effect of changing valuations. In the model, any structural change in the economy that changes R&D through the present value of ideas—for example, falling discount rates or rising market power—will also change the value of the stock market roughly one-for-one, leaving R&D-to-value relatively unchanged. In contrast, a decline in innovation productivity causes R&D-to-value to fall. R&D-to-value therefore provides good identification for the level of innovation productivity.

The estimated decline in innovation productivity explains most of the fall in growth and the rise in valuations. To see this, I feed the estimates into the model and study the economy's transition path. The only exogenous change along the transition is the fall in innovation productivity; all other parameters are held fixed. Per-capita output growth falls by 1.1 percentage points in the model, virtually the same as in the data. The mechanism also explains the rise in aggregate M&A-to-R&D.[4] Finally, the transition explains most of the rise in aggregate valuation ratios. The level of the value-to-sales ratio rises by 83% over fifty years, compared with the roughly 85% rise in the data. As in the data, virtually all of this increase comes from a similar increase in value-to-cashflows. Notably, the decline in innovation productivity does not explain the large decline and recovery of valuations around 1980, which coincided with a large rise and fall in discount rates (namely, real interest rates).[5]

As in the data, the model implies that most of the long-run rise in aggregate valuation ratios came from a reallocation of market share toward high-valuation firms. Firm-level valuations are, on average, stable. Why does the model imply stable firm-level valuations? While declining innovation does reduce firm-level growth rates and, in turn, valuations, this is offset by three forces. First, equilibrium discount rates fall with aggregate growth. Second, the shift to M&A enables high-productivity firms to keep growing without innovating. And third, because competitors and entrants are also doing less R&D, the risk of creative

---

[4]The transition does not explain the late-1990s M&A wave; however, the cumulative rise in M&A over this period is similar in the model and data, which is what matters for quantifying total reallocation.

[5]This episode was mainly driven by a large "within-firm" change in the micro decompositions of aggregate valuations, further supporting the idea that this is explained by a common discount-rate shock.

destruction falls, decreasing the rate of entry and raising the expected growth of incumbents.

Declining innovation productivity can also help explain trends in other macro and micro series. I highlight three trends of interest. First, the model explains all of the decline in the rate of new-firm entry, which in the data fell from about 12% to about 8% since 1980. The reason is that entrepreneurs also have a harder time creating new firms. This results in not only less entry, but also less creative destruction. Second, and closely related to this fact, the model implies an increase in firm concentration. With less entry and more acquisitions of small, low-productivity firms, production becomes concentrated into high-productivity incumbents, which become large and entrenched. And third, model-implied interest rates fall with the rate of economic growth. As already stated, this cannot account for the dramatic rise and fall of rates around 1980, but it does account for much of the longer-run decline in real rates since the 1960s.[6] This decline in discount rates does not directly cause an increase in firm-level valuations, because of an offsetting decline in expected cashflow growth.

Despite rising stock-market wealth, investors are still worse off.[7] Consumer welfare depends on the present value of consumption. One might therefore infer from the stock-market boom that welfare has risen, because present values are high. However, in a disaggregated economy, the stock-market claim is not the same as the consumption claim. The stock market is only a claim to current incumbents' profits. The consumption claim, in contrast, also pays out the factor income of workers and the future profits of new entrants. Indeed, stock-market wealth is a very small fraction of total wealth.[8] This distinction is crucial for welfare: all of the gains in the stock market came from a reallocation to incumbents, at the expense of wages and future firms, so the boom in stock-market value did not imply any such boom in the present value of consumption. The value of the consumption claim stagnated with output, implying a disconnect between the stock market and consumer welfare.

---

[6]Farhi and Gourio (2018) and Miller, Paron, and Wachter (2022) focus on the puzzle that interest rates have plummeted since 1980 but the market has not risen commensurately. Farhi and Gourio (2018) argue that the equity premium rose while Miller et al. (2022) argue that the true riskfree rate did not fall as much as real yields on government debt. Both of these stories are qualitatively consistent with my findings. Like Binsbergen (2021), I argue that a decline in growth can help reconcile the performance of stocks and bonds.

[7]Of course, insofar as markets are incomplete, some households may have benefited more than others.

[8]The U.S. stock market is historically only about twice the size of one year of aggregate consumption. Even with extremely high discount rates, the consumption claim would have a much higher multiple than 2.

Related literature This paper advances and bridges the literatures on economic growth, asset pricing, and firm dynamics.

First, I present new, well-identified evidence of declining innovation productivity within firms. To be sure, this is not the first paper to argue that new ideas are getting harder to find. C. I. Jones (1995) points out that stable economic growth, coupled with an exponential increase in the number of researchers, must mean innovation productivity per researcher is falling. In support of this intuition, B. F. Jones (2009, 2010) documents that, over time, innovation has required larger teams of more educated, older researchers. Gordon (2016) presents evidence that ideas have been getting harder to find in the U.S. for decades, and that this explains the recent growth slowdown. Bloom et al. (2020) compile micro-level evidence of declining innovation productivity, demonstrating that the macro-level argument of C. I. Jones (1995) is also true within product markets and firms.

While all of these studies are based on the relationship between research inputs and ex post growth, this paper identifies the decline in innovation productivity from a different source: R&D spending and market valuations. Because R&D and value are both forward-looking and contemporaneous, they can be used to estimate the historical time series of innovation productivity at a relatively high frequency.[9] Moreover, the use of a micro-founded endogenous-growth framework allows me to quantify the effect of this decline on not only economic growth, but also a host of other micro and macro moments.

My findings connect the secular decline in growth with the literature on the secular rise in the value of the stock market (Farhi & Gourio, 2018; Miller et al., 2022; Greenwald, Lettau, & Ludvigson, in press; Atkeson, Heathcote, & Perri, 2024; Eggertsson, Robbins, & Wold, 2021; Cho, Grotteria, Kremens, & Kung, 2024). I make four contributions. First (and foremost), I present new micro-level evidence and a new mechanism to explain this rise. Second, while past studies have taken growth to be an exogenous input, I endogenize both growth and valuations to study their general-equilibrium relationship.[10] Third, I

---

[9]My estimate of a roughly 50% total decline in research productivity since 1975 is smaller in magnitude than, but still reasonably close to, most of the micro and macro estimates in Bloom et al. (2020).

[10]Most of these studies are also interested in understanding the recent decline in capital investment, which is puzzling given rising valuations. I show, in an extension with capital, that this is a natural implication of declining innovation productivity. This corroborates the finding of Crouzet and Eberly (2023) that intangibles (in this case, incumbents' stock of ideas) drove a wedge between "average Q" and "marginal q."

depart from the representative-firm assumption in these papers and instead model the full distribution of individual firms. Because reallocation between firms is key to explaining the rise in valuations, a representative-firm model cannot possibly explain these dimensions of the data. And fourth, because I model individual firms, I distinguish the stock market from the consumption claim. This distinction turns out to have major welfare implications: the reallocation-driven growth in the stock market came at the expense of wages and future firms, implying that the present value of consumption (welfare) stagnated with output.[11]

A separate asset-pricing literature studies the implications of firm heterogeneity and innovation for the aggregate market. One strand of this literature shows how new-firm entry and imperfect competition can increase the market risk premium (Gârleanu, Kogan, & Panageas, 2012; Corhay, Kung, & Schmid, 2020; Loualiche, in press). Another strand shows how firm heterogeneity can increase aggregate return volatility (Cochrane, Longstaff, & Santa-Clara, 2008; Martin, 2013). While this literature has so far focused on high-frequency volatility and risk premia, I show that firm heterogeneity is important even for determining the very-long-run *level* of market valuations. Indeed, all of the findings of this paper remain true even in an economy without any systematic risks or non-standard preferences.

Finally, this work builds on and contributes to the literature on firm dynamics and endogenous growth, the foundations of which are laid in Romer (1990), Grossman and Helpman (1991), and Aghion and Howitt (1992). My framework builds primarily on the Schumpeterian growth model of Klette and Kortum (2004) and Lentz and Mortensen (2008). This class of endogenous-growth model has been extended to study numerous aspects of the economy, including markups and misallocation (Peters, 2020) and optimal innovation policy (Acemoglu, Akcigit, Alp, Bloom, & Kerr, 2018). My model studies a new tradeoff within firms—R&D versus M&A—in an economy with ex ante heterogeneous firm-level productivity (Hopenhayn, 1992; Luttmer, 2011). I use the model to study the economic transition resulting from a new structural change: new ideas getting harder to find.

Several firm-dynamics papers have also studied long-run structural change in the economy, with a particular interest in the fall in growth and the rise in profits and concentration.

---

[11]Greenwald et al. (in press) assume the stock market is a levered consumption claim, thus making the critical distinction between the profit claim and human capital. They do not, however, distinguish between incumbents' profits and future entrants' profits, which I find to be essential.

Peters and Walsh (2022) show how declining population growth reduces firm entry and increases concentration.[12] Aghion, Bergeaud, Boppart, Klenow, and Li (2023) and De Ridder (2024) propose that technological changes enabled some firms to scale up more easily, leading to rising concentration.[13,14] Akcigit and Ates (2023) and Olmstead-Rumsey (2022) argue that a decline in knowledge diffusion between leaders and laggards explains these trends.[15] This literature has by and large not considered asset prices. A notable exception is Liu, Mian, and Sufi (2022), who argue that a decline in discount rates (i.e., a rise in firm valuations) allows leaders to escape competition by laggards, until eventually both firms stop innovating. This could be a complementary force in the long run; however, over the transition, it cannot explain the decline in R&D-to-value and economic growth, or the stability of within-firm valuations, implying that my mechanism is essential for explaining these facts.

Roadmap Section 1 lays out the empirical facts. I describe the model in Section 2 and its solution in Section 3. Section 4 reports the estimation results. Section 5 presents the main results and Section 6 the welfare implications. Finally, Section 7 discusses model extensions. Proofs and additional details are in the Online Appendix.

1 Empirical evidence

This section documents empirical evidence of declining innovation productivity and reallocation of market share to high-valuation firms. I first report and discuss trends in aggregate value, R&D, and M&A. I then examine valuations in firm-level micro data. After laying out these facts, I investigate and rule out alternative explanations.

---

[12]They find that declining population growth initially boosts growth before lowering it. Feeding the population-growth decline into my model would imply both a short- and a long-run growth decline. This could also generate the more extreme prediction of C. I. Jones (2022a) of the "end of economic growth."

[13]In De Ridder (2024), growth falls because of decreasing returns to R&D within firms: rising concentration reduces the productivity of the average researcher, rendering new ideas effectively harder to find.

[14]Aghion et al. (2023) also seek to explain the empirical fact that markups have fallen within firms, yet aggregate markups have risen because of a reallocation to high-markup firms (Kehrig & Vincent, 2021; De Loecker, Eeckhout, & Unger, 2020; Autor, Dorn, Katz, Patterson, & Van Reenen, 2020). My model also generates diverging micro and macro trends, because the high-productivity firms that gain market share are also high-markup firms, but have to compete more with each other as they gain market share.

[15]While declining knowledge diffusion could be a cause of declining innovation productivity—say, if knowledge about competitors' ideas inspires researchers to invent new ones—the key mechanism in these papers is the asymmetric effect between leaders and laggards, which is different from the uniform shift in my paper.

Figure 1: Trends in aggregate value and R&D

A. Aggregate value-to-sales

B. Aggregate R&D-to-value
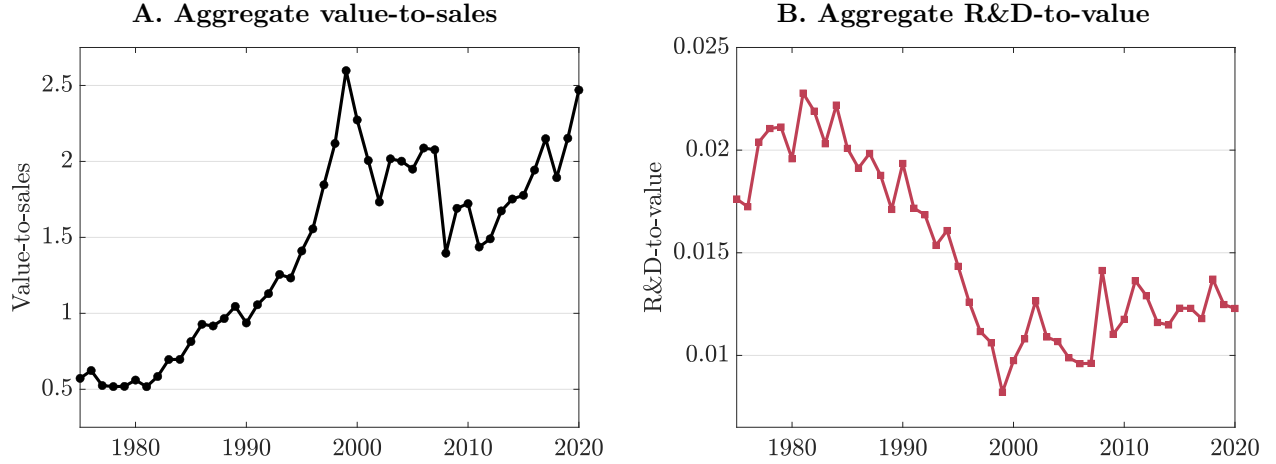
The figure plots the ratio of aggregate market value to aggregate sales (Panel A) and the ratio of aggregate R&D spending to aggregate market value (Panel B) for CRSP/Compustat firms from 1975–2020. Value is defined as market equity plus book debt, less cash.

The main data sources are the annual sample of U.S. CRSP/Compustat firms and the history of M&A transactions from SDC Platinum. Note that 1975 is the first year in which firms were required to report R&D spending. The baseline measure of firm value is market equity plus book debt, less cash; cashflows are the sum of dividend and interest payments. For more details about data sources, selection, and variable construction, see Appendix A.

1.1 Diverging trends in aggregate value and R&D

Panel A of Figure 1 shows that, over the last fifty years, aggregate value has boomed relative to aggregate sales (gross output). The fact that the present value of profits is high relative to output suggests that the expected gains from successful innovation are high, so firms should have a greater incentive to do R&D. Panel B shows that this intuition has been far from true: aggregate R&D spending has fallen substantially relative to aggregate value. Something has driven a wedge between R&D and the present value of profits.[16]

---

[16]In Appendix B.1.1, I show that a similar decline can be seen in the patent value measure of Kogan, Papanikolaou, Seru, and Stoffman (2017). I also discuss how their measure relates to R&D spending.

9

Figure 2: Trends in M&A relative to R&D
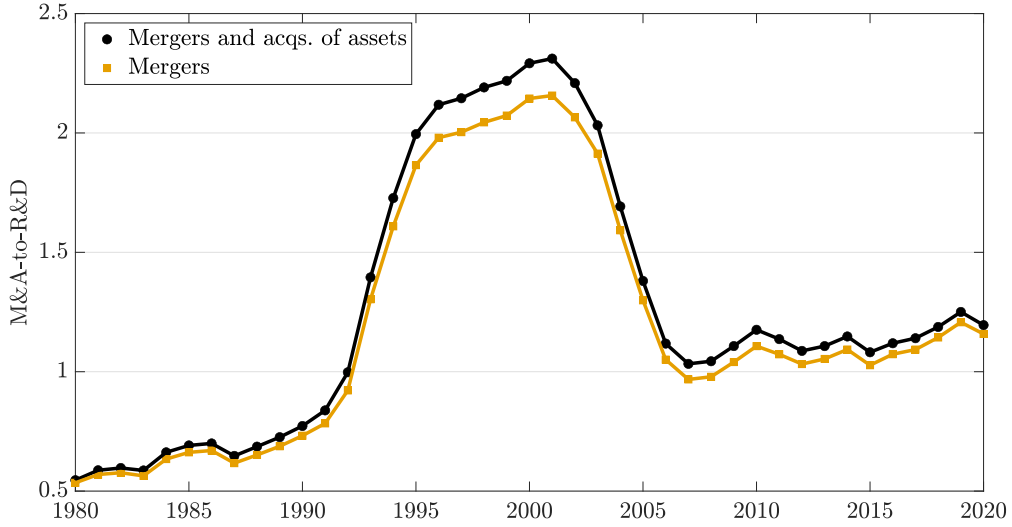
The figure plots the aggregate ratio of M&A spending to R&D spending for public firms in Compustat and SDC Platinum from 1980–2020. Each point is a two-sided moving average with windows of ± 5 years. M&A is defined as the dollar amount spent by acquirers on either mergers (transactions in which all of the target is acquired) or mergers plus acquisitions of assets.

1.2 The rise of M&A over R&D

Figure 2 shows that M&A spending has risen substantially relative to R&D spending. The figure plots the ratio of total M&A spending to total R&D spending from 1980–2020.[17] M&A spending is defined as the total dollar amount paid by the acquirer in the deal. I consider two definitions of M&A, mergers (acquisitions of entire firms) and mergers plus acquisitions of assets, both of which show similar trends.[18]

The M&A-to-R&D ratio effectively doubled over this period. The fact that both M&A and R&D are measured in dollar values (as opposed to, say, number of deals or number of patents) is key for this comparison, because higher rates of spending will reflect expectations of larger value gains to firms. The doubling of M&A-to-R&D over time therefore suggests that firms are growing less and less through innovation and increasingly through reallocation.

---

[17]The series is filtered with a five-year, two-sided moving average to help with visualization and, specifically, to smooth out the M&A wave of the late 1990s. Appendix B.1.2 reports and analyzes the unfiltered series.

[18]Acquisitions of assets may include acquisitions of ideas and product lines—the focus of this paper—or purchases of resources like physical capital. These cannot be disentangled in the SDC data.

1.3 Reallocation accounts for the booming stock market

Reallocation could increase the value of the stock market if it means that high-valuation firms are growing larger. Consider the aggregate value-to-sales ratio V_t/Y_t. This aggregate ratio can be written as a sales-weighted average of firm-level value-to-sales ratios $V_{it}/p_{it}Y_{it}$:

$$\frac{V_t}{Y_t} = \sum_{i=1}^{M_t} \frac{p_{it}Y_{it}}{Y_t} \times \frac{V_{it}}{p_{it}Y_{it}}. \tag{1}$$

The aggregate ratio could rise either because firm-level value-to-sales ratios rise, or because the sales shares of high-valuation firms increase (reallocation).

I decompose these two channels using the following accounting identity. Let $\mathcal{I}$ denote the set of firms (ignoring entry and exit for now). The change in (1) from $t-1$ to t can be written[19]

$$\Delta \frac{V_t}{Y_t} = \underbrace{\sum_{i \in \mathcal{I}} \left(\overline{\frac{p_{it}Y_{it}}{Y_t}} \right) \times \Delta \frac{V_{it}}{p_{it}Y_{it}}}_{\text{within-firm change}} + \underbrace{\sum_{i \in \mathcal{I}} \Delta \frac{p_{it}Y_{it}}{Y_t} \times \left(\overline{\frac{V_{it}}{p_{it}y_{it}}} \right)}_{\text{compositional change}}. \tag{2}$$

The within-firm change measures the effect of changes in firm-level valuations, holding fixed the sales shares of firms; the compositional change measures the effect of changing sales shares, holding fixed valuations. The operator $(\overline{\cdot})$ denotes an index averaging over the two periods: for example,

$$\left(\overline{\frac{V_{it}}{p_{it}Y_{it}}} \right) \equiv \frac{1}{2} \left(\frac{V_{it}}{p_{it}Y_{it}} + \frac{V_{i,t-1}}{p_{i,t-1}Y_{i,t-1}} \right). \tag{3}$$

Using these indices—as opposed to leads or lags—is critical, because changes in sales shares and changes in value-to-sales are mechanically (negatively) correlated in the cross-section. The indices neutralize this correlation. Appendix B.2 lays out this intuition formally.

Because there is entry and exit in the data, we must make two small adjustments to (2). First, for a continuer (a firm operating at both $t-1$ and t), the sales share is defined as the share of total *continuer* sales. Second, for an entering firm (operating at t and not $t-1$), the time-$(t-1)$ sales share is set equal to the time-t share of aggregate sales, while the time-$(t-1)$ value-to-sales ratio is set equal to the time-t value-to-sales ratio of all continuers. Thus, for entrants, the "within-firm change" measures the extent to which the firm has a higher or

---

[19]See Appendix B.2 for a derivation.

Figure 3: Decompositions of aggregate valuations
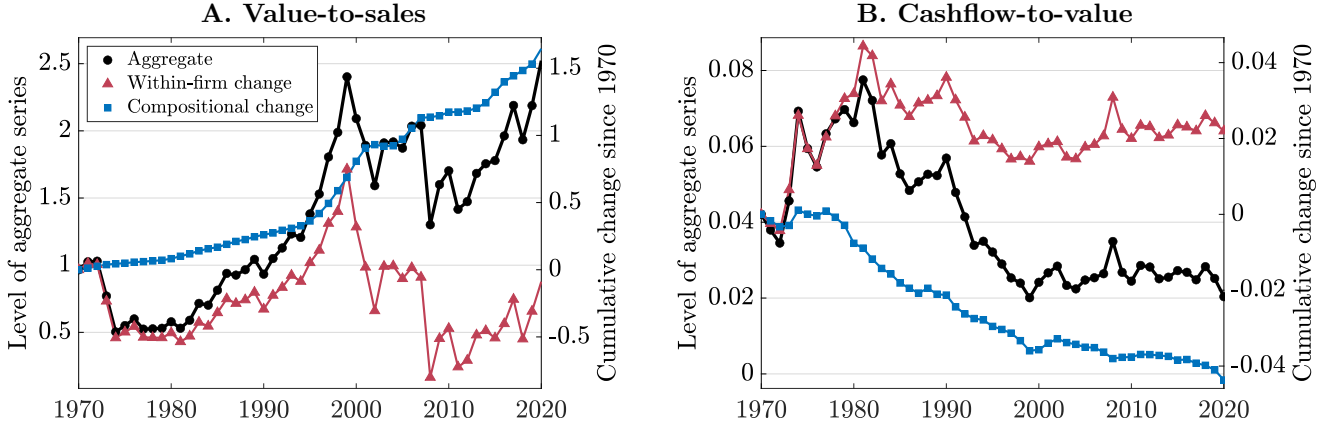
The figure plots the decompositions of changes in aggregate valuations into within-firm and compositional components. The decomposition for value-to-sales is defined in (2) and the decomposition for cashflow-to-value in (4). See the main text for an explanation. Value is defined as market equity plus book debt, less cash; and cashflows are defined as dividends plus interest payments.

lower valuation than the set of continuers. Likewise, for an exiting firm, the time-t sales share equals the time-$(t-1)$ share and the time-t value-to-sales ratio equals the time-$(t-1)$ ratio of all continuers. Empirically, the effects of entry and exit are very small—I show this in Appendix B.3, which separates out these terms. Nevertheless, these adjustments make it so that, under the null hypothesis of a stationary firm distribution (and a constant aggregate value-to-sales ratio), both the within-firm and compositional changes will equal zero. See Appendix B.2 for a detailed explanation and examples.[20]

Panel A of Figure 3 plots the cumulative change in aggregate value-to-sales since 1970 alongside the cumulative changes attributable to each of the channels in (2). While most high-frequency changes in aggregate value-to-sales come from within-firm changes, the cumulative within-firm change has been small. Most of the long-run increase in the aggregate is instead explained by compositional change. In other words, there has been a sizable reallocation of revenues toward high-valuation firms. Reallocation explains the booming market.

Notably, the large and positive compositional change in the late 1990s coincided with the M&A wave (Figure 2). This is consistent with existing evidence that acquiring firms have significantly higher valuation ratios than target firms (Arikan & Stulz, 2016) and that the

---

[20]I also show how the decomposition (2) can be constructed from the Haltiwanger (1997) decomposition that is commonly used in macroeconomics, and how the adjustments I make are important for addressing econometric biases that emerge in this class of decomposition when studying secular change in valuations.

combined value of the target and acquirer rises significantly around merger announcements (Betton, Eckbo, & Thorburn, 2008).[21]

The rise in aggregate value-to-sales is not simply about higher profits; aggregate value has also risen relative to *cashflows*. Panel B of Figure 3 plots the analogous decomposition for aggregate cashflow-to-value D_t/V_t:

$$\Delta \frac{D_t}{V_t} = \underbrace{\sum_{i \in \mathcal{I}_{t-1,t}} \left(\overline{\frac{V_{it}}{V_t}} \right) \times \Delta \frac{D_{it}}{V_{it}}}_{\text{within-firm change}} + \underbrace{\sum_{i \in \mathcal{I}_{t-1,t}} \Delta \frac{V_{it}}{V_t} \times \left(\overline{\frac{D_{it}}{V_{it}}} \right)}_{\text{compositional change}}, \qquad (4)$$

where $\mathcal{I}_{t-1,t}$ is the set of firms operating at $t-1$ or t. I decompose cashflow-to-value, instead of value-to-cashflow, simply because firm-level value-to-cashflow is undefined in periods with zero cashflows. The finding is the same: within firms, valuations have been relatively stable, while high-valuation (low cashflow-to-value) firms have gained substantial market share.

Appendix B.3 shows that these decomposition results are robust to a number of possible concerns. They continue to hold if we separate out entry and exit (which are small), if we consider only equity values and dividends, and if we use alternative definitions of cashflows that include net share repurchases. Moreover, the appendix also studies the implications of the above facts for R&D-to-sales, a common measure of innovation in the literature. High-valuation firms also have high R&D-to-sales ratios, and there has been a corresponding reallocation of sales toward high-R&D-to-sales firms, which has increased the aggregate ratio.

1.4 Examining alternative explanations

The main hypothesis of this paper is that declining innovation productivity can explain the above micro and macro trends. Before evaluating this hypothesis, I consider two alternative explanations for these trends: selection among public firms and changing composition among sectors. Neither one of these alternatives explains the empirical evidence above.

Has R&D been outsourced to private firms? One possible explanation of declining R&D is that R&D has been outsourced to private firms, which then get acquired by public

---

[21]Appendix B.1.3 gives a more detailed survey of the evidence of value creation from the M&A literature.

firms, meaning that the decline in R&D and the rise in M&A are really just driven by selection into public markets. While we are limited in testing this explanation by the lack of valuation data for private firms, there are available data on R&D spending and M&A exits of private firms. I use these data to test two predictions.

First, if R&D has been outsourced to private firms, then we should see that the R&D intensity of private firms has grown by more than that of public firms, offsetting the decline in the aggregate. Figure B.7 plots the time series of R&D-to-output in the U.S., inclusive of all firms, both public and private. The trend in this total R&D intensity is remarkably similar to the ratio of R&D-to-sale among public firms. In fact, the economy-wide measure increased by slightly *less* than the public-firm measure. There is, therefore, no evidence of a shift in R&D activity toward non-listed firms.

A second prediction of outsourcing to private firms is an increase in the share of M&A spending by public firms on acquisitions of private targets. Figure B.7 reports the share of public firms' total M&A spending on public versus private targets over time. These shares exhibit no trend, which is evidence against the prediction that public firms are increasingly growing by buying up private ideas.[22]

Was the reallocation within or between industries?   Another possible explanation for the reallocation-driven stock-market boom is a reallocation of economic activity between industries. For example, perhaps the high-growth tech sector outgrew the lower-growth construction sector due to changing preferences or technologies. If this is the case, then declining innovation productivity within firms could be a separate phenomenon that has nothing to do with the reallocation to high-growth firms.

To test this hypothesis, I re-run the decompositions (2) and (4) at the industry level instead of the firm level. If the reallocation was mainly between industries, then we should see that most of the aggregate change is explained by compositional changes, as in the firm-level decompositions. If instead the reallocation was mainly between firms *within* industries, then most of the changes will come from industry-level changes. Appendix B.5.1 reports

---

[22]This is consistent with the findings of Doidge, Karolyi, and Stulz (2017) and Eckbo and Lithell (2023) that the U.S. listing gap is in part explained by the increasing M&A of public targets by public acquirers. In other words, the "missing firms" in public markets are still operating under the ownership of public acquirers.

the results. In all cases, only a small proportion of the aggregate change is explained by a reallocation between sectors, meaning that the reallocation was mainly *within* sectors.

The fact that the reallocation has been largely within sectors presents an opportunity to test this paper's main mechanism. If declining innovation productivity causes an increase in M&A activity and a reallocation to high-valuation firms, then we should see that industries that experienced a larger decline in R&D-to-value should also have seen a bigger rise in total valuations and M&A, and a larger compositional change in the valuation decompositions. Appendix B.5.2 tests these predictions in the data and finds that they are true.

2 Model

I next present a Schumpeterian growth model in which heterogeneous firms compete in product markets and grow through R&D and M&A. The model is composed of two blocks: households and firms. The household block is standard, serving primarily to define the objectives of firms (including product-market structure) and close the economy. Most of the new and consequential assumptions underlie the firm block.

2.1 Households

The economy is populated by a continuum of individuals of mass L_t with population growth rate $\dot{L}_t/L_t = g_L$. A household is a collection of individuals that consumes, supplies labor, and trades in financial markets.

Intertemporal preferences Individuals have identical log utility over consumption c_t and constant-Frisch disutility over labor hours l_t. The lifetime utility of a single individual takes the form

$$U_t \equiv \mathbb{E}_t \left[\int_0^\infty e^{-\rho\tau} \left(\log c_{t+\tau} - \chi \frac{l_{t+\tau}^{1+1/\zeta}}{1+1/\zeta} \right) d\tau \right], \tag{5}$$

where ρ is the rate of time preference (impatience). Log utility corresponds to an elasticity of intertemporal substitution (EIS) of one. The parameter χ governs the overall disutility of working; the Frisch elasticity ζ governs the sensitivity of labor supply to changes in wages.

Demand for goods (market structure) An individual's total consumption c_t is a bundle of differentiated goods. In particular, firms produce a mass N_t of varieties, indexed j. The individual's final demand for quantities c_{jt} of these goods are then aggregated as

$$c_t \equiv \left(\int_0^{N_t} c_{jt}^{1-1/\eta} dj \right)^{1/(1-1/\eta)}, \tag{6}$$

subject to the budget constraint and taking each good's price as given. The parameter η is the elasticity of substitution between varieties: the higher is η, the more substitutable.

Within variety j, $m_{jt} \in \mathbb{N}$ firms, indexed i, possess a blueprint to produce the good. The household views these blueprints as perfect substitutes, but with potentially heterogeneous qualities, denoted q_{ijt}. Hence, the total consumption of good j is the sum

$$c_{jt} \equiv \sum_{i=1}^{m_{jt}} q_{ijt} c_{ijt}, \tag{7}$$

where c_{ijt} is the quantity of firm i's output consumed by the household. Demand (6) and (7) will ultimately determine the competitive structure of product markets.

Labor supply A fraction $\bar{s}$ of individuals are endowed with skilled labor. They can supply labor hours either to an incumbent firm for a wage w_{xt} or to a private effort to start a new firm (entrepreneurship). Entrepreneurship is detailed below. The household decides the mix of hours between these tasks. The remaining fraction $1 - \bar{s}$ can supply production labor for a wage w_{pt}. While the fraction of laborers $\bar{s}$ dedicated to skilled labor tasks is fixed, it is straightforward to endogenize this separation using a Roy model, as in Luttmer (2011).

Securities market Households trade without constraints in equity and bond markets. Claims to firm profits (the stock market) constitute the total supply of marketable wealth. All other traded assets—in particular, riskfree bonds—are in zero net supply.

Household composition A household is defined as a mass of individuals, a *family*, who jointly make consumption, labor, and portfolio decisions. This implies that each household is comprised of a representative sample of individuals, and so, without loss of generality, we

can consider a single representative household composed of the full population.[23]

2.2 Firms

There is a continuum of firms, indexed i, with mass M_t. Each firm employs labor for production and expansion into new markets in order to maximize its market value.

Firm value Let Π_{it} denote the operating profits of the firm and $w_{xt}X_{it}$ its total expenditure on skilled labor for R&D and M&A. Dividends are the total cashflows paid to owners:

$$D_{it} \equiv \Pi_{it} - w_{xt}X_{it}. \tag{8}$$

The firm's objective is to maximize the present value of its dividends:

$$V_{it} \equiv \mathbb{E}_t \left[\int_0^\infty \frac{\xi_{t+\tau}}{\xi_t} D_{i,t+\tau} d\tau \right], \tag{9}$$

where ξ_t denotes the state-price density of the household. This value-maximization problem has two parts. The first is the static profit maximization, in which the firm produces its current set of products. The second is the dynamic investment decision, in which the firm allocates skilled labor to R&D and M&A.

Production and profits Firm i has blueprints to produce $n_{it} \in \mathbb{N}$ varieties. The firm employs labor to produce variety j according to the linear production function[24]

$$Y_{ijt} = a_i L_{ijt}, \quad j \in \{1, \ldots, n_{it}\}, \tag{10}$$

where the idiosyncratic productivity type a_i can take a low value a_ℓ or a high value a_h. A firm's productivity type is fixed over its life; Section 7 discusses implications of relaxing this

---

[23]Note that this assumption is not the same as a decentralized economy with perfect consumption insurance (complete markets), because entrepreneurial decisions would be distorted by a moral hazard problem from redistributing private gains (C. I. Jones, 2022b). One could consider an economy without consumption insurance, but then the distribution of households would become a state variable. The assumption of a family, in contrast, yields essentially complete markets without distorting private incentives for firm creation.

[24]It is common in endogenous-growth models to put the good-specific product quality q_{ijt} into the production function (10) instead of the demand function (7). These are isomorphic assumptions.

assumption by allowing for stochastic switching between productivity types.

Each blueprint has associated with it a set of characteristics, including its quality q_{ijt} and the qualities and productivities of competitors. Given these characteristics, firms internalize the demand curves of households and engage in Bertrand competition (i.e., choose prices $\{p_{ijt}\}$).[25] All firms pay a common production wage $w_{\mathrm{p}t}$. Hence, a firm with productivity a_i and n_{it} products earns operating profits

$$\Pi_{it} = \max_{\{p_{ijt}\}} \sum_{j=1}^{n_{it}} (p_{ijt} Y_{ijt} - w_{\mathrm{p}t} L_{ijt}), \tag{11}$$

where Y_{ijt} and L_{ijt} depend on prices and product characteristics due to household demand.

Innovation (R&D)   Firms can come up with new ideas in two ways, each corresponding to its own R&D technology (Peters & Walsh, 2022). The first is new-variety creation, whereby the firm invents a blueprint for a new good and becomes a monopolist. The second is a quality improvement on an existing variety, in which case the firm enters the market and competes with incumbents. Firms choose how much labor to allocate to each technology at any given time. Allowing for both types of innovation is essential for two reasons. First, new-variety creation will be necessary for balanced growth with a growing population. Second, the relative frequencies of quality improvements and new varieties will determine the degree of creative destruction between firms, which will feed back into growth rates and valuations.

The R&D technologies combine skilled labor, hired at the wage $w_{\mathrm{x}t}$, with the firm's existing knowledge stock n_{it} to probabilistically yield a new idea. Specifically, an n-product firm that allocates $X_{\mathrm{q}it}$ units of labor to quality innovation will successfully invent a blueprint with Poisson intensity

$$\alpha \varphi X_{\mathrm{q}it}^{\varepsilon} n_{it}^{1-\varepsilon}. \tag{12}$$

Likewise, a firm allocating $X_{\mathrm{n}it}$ units to new-variety creation will succeed with intensity

$$(1-\alpha) \varphi X_{\mathrm{n}it}^{\varepsilon} n_{it}^{1-\varepsilon}. \tag{13}$$

These idea-production functions capture the fact that R&D generates new ideas by building on existing ideas. The parameter $\varepsilon \in (0, 1)$ governs the importance of researchers in this process. Economically, $\varepsilon < 1$ implies decreasing marginal returns to research inputs. The parameter φ represents *innovation productivity*, the total factor productivity of R&D. Higher φ means that it is easier to come up with new ideas in any form. Finally, $\alpha \in (0, 1)$ determines the relative productivities of quality improvements over new varieties. I will also refer to α as the creative destruction parameter, because it increases the extent to which innovation negatively affects incumbent firms' current profits.[26]

Quality improvements are fixed increments $\lambda > 1$ over the current quality of the target firm's blueprint: if firm i improves on the blueprint for good j owned by firm i', then this new blueprint has quality $q_{ijt} = \lambda q_{i'jt}$. The firm identifies an existing blueprint on which to improve by random search, meaning that the probability of improving on a blueprint with a given set of characteristics equals the fraction of blueprints with those characteristics currently in production across the economy.[27] Each new variety has a quality equal to λQ_t, where Q_t is the current aggregate (average) quality of blueprints, defined below.

Acquisition (M&A) Firms can also hire from the pool of skilled laborers in order to identify and acquire blueprints from other firms.[28] In particular, each firm has access to an M&A technology, analogous to the R&D technologies, with which they can hire $X_{\mathrm{m}it}$ units of labor to search for an existing idea. The investment in M&A search successfully finds a target with Poisson intensity

$$\psi X_{\mathrm{m}it}^{\varepsilon} n_{it}^{1-\varepsilon}. \tag{14}$$

This M&A search function is purposely defined to have the same form as the R&D technologies (12) and (13), in order to focus comparison between the relative productivities φ and

---

[26] An alternative assumption would be to have a single R&D technology with undirected innovations ex post: researchers come up with an idea, and then that idea happens to be a quality innovation with probability α (see Peters & Walsh, 2022). Properly calibrated, this alternative would yield similar results.

[27] It would be straightforward to allow firms to direct search in quality innovations by allowing them to target blueprints with specific characteristics. Further, as Section 7 discusses, the assumption that firms only innovate on other firms' blueprints could be relaxed by introducing an own-innovation technology.

[28] I assume that firms acquire individual blueprints instead of entire firms because the latter would significantly increase the dimensionality of the state space. This is of minimal economic consequence, provided that the M&A search function is properly calibrated—having firms acquire single blueprints at a higher frequency has the same implication as having them acquire collections of blueprints at a lower frequency.

ψ. It would, for example, be equivalent to have an M&A search technology that is linear in X_m but subject to an appropriately scaled adjustment cost.

As with quality innovations, acquisition is undirected and hence determined by the current distribution of blueprints in the economy. Once a firm successfully identifies a blueprint target, it then enters into a negotiation with the target firm. Specifically, the two parties engage in a Nash bargaining game in which they decide on a deal price. Letting v_buyer and v_target denote the private values of the idea to the buyer and seller, respectively, the deal price solves

$$v_\mathrm{m} = \underset{v}{\mathrm{argmax}} \left\{ (v_\mathrm{buyer} - v)^\varrho (v - v_\mathrm{target})^{1-\varrho} \right\}. \tag{15}$$

The bargaining power of the acquirer is increasing in the parameter $\varrho \in (0,1)$. If both parties are indifferent between making a deal and walking away, there is no deal, consistent with an infinitesimal fixed cost of transacting.

Entry and exit   Skilled laborers start new firms by creating their own new ideas. I assume that these entrepreneurs have access to the exact same idea-creation technologies (12) and (13) as a single-product incumbent ($n = 1$). Before investing in this technology, the entrepreneur knows the productivity type a_i, which is high with probability $\omega_h = \omega$ or low with probability $\omega_\ell = 1 - \omega$. The opportunity cost of a skilled laborer doing entrepreneurial research is the foregone skilled wage that could be received by working for an incumbent.

If a single-product firm loses its last good, either because of creative destruction or because it was acquired, then that firm ceases to operate. For convenience, I assume that even if a firm ceases operations, its blueprint still exists and incumbents recognize the threat of re-entry. This is not strictly necessary, but it makes the model more tractable because it means we do not need to keep track of those blueprints that are owned by dead firms.

3 Model equilibrium

3.1 Equilibrium definitions

To define an equilibrium, we need a more precise notion of the distribution of ideas in the economy. Each blueprint can be summarized by a specific set of features, including its quality, its producer's productivity, and the characteristics of competitors. For now, let us denote by Z the set of a product's characteristics and $\mathcal{Z}$ the full space of possible Z. For instance, firm i's blueprint for good j can be summarized by some $Z_{ij} \in \mathcal{Z}$.

Let us partition the aggregate blueprint stock across characteristics Z (so $N_t = \sum_{Z \in \mathcal{Z}} N_{Zt}$) and let $f_{Zt} \equiv N_{Zt}/N_t$ represent the proportion of ideas currently in operation with characteristics Z. Thus, the set $\{f_{Zt} : Z \in \mathcal{Z}\}$ represents the full blueprint distribution.[29] With this notation in hand, we can define the concept of equilibrium.[30]

Definition (Equilibrium). *Fix an initial blueprint distribution $\{f_{Zt} : Z \in \mathcal{Z}\}$ and an initial population-to-blueprints ratio L_t/N_t. An* equilibrium *is a time path of goods prices $\{p_{ijt}\}$, wages $\{w_{\mathrm{p}t}, w_{\mathrm{x}t}\}$, state-price densities $\{\xi_t\}$, goods consumption $\{c_{ijt}\}$, labor hours $\{l_{\mathrm{p}t}, l_{\mathrm{x}t}\}$, production-labor allocations $\{L_{ijt}\}$, and skilled-labor allocations $\{X_{\mathrm{n}it}, X_{\mathrm{q}it}, X_{\mathrm{m}it}\}$, such that:*

1. *Optimization: Households and firms optimize according to the setting in Section 2.*

2. *Market clearing: Household labor supply equals firm labor demand, the consumption of each good equals the quantity produced, and household wealth equals total firm value.*

The definition fixes the initial distribution $\{f_{Zt} : Z \in \mathcal{Z}\}$ and ratio L_t/N_t because these are slow-moving states that evolve forward in response to equilibrium policies. We next define a stricter notion of an equilibrium in which these slow-moving states are constant.

---

[29]Note that this distribution is not the same as the distribution across *firms*—which firms own which sets of blueprints—which we will see is sufficient but not necessary for defining and solving the equilibrium.

[30]The equilibrium is well-defined provided that, for every firm i and time t, the expected return on the firm is greater than the expected growth rate of its value: $\mathbb{E}_t[dR_{it}] > \mathbb{E}_t[dV_{it}]/V_{it^-}$. This "$r$ minus g" condition is necessary for all firms to have finite market values; if it is violated, then the valuations of firms with high expected growth rates are infinite/undefined. An example of a parameterization that might violate this condition is a very low value of α and a high a_h/a_ℓ, in which case incumbents face very little creative destruction, so high-productivity firms may grow more quickly than the returns demanded by investors.

Definition (Balanced-growth equilibrium). *A balanced-growth equilibrium (BGE) is an equilibrium in which, for all time t, the blueprint distribution is stationary ($\dot{f}_{Zt} = 0$ for all $Z \in \mathcal{Z}$) and the population-to-blueprint ratio L_t/N_t is constant.*

Intuitively, balanced growth means that all resources are growing at the same rate across firms (a stationary distribution) and that the factor inputs to innovation, L_t and N_t, grow proportionately ($g_N = g_L$). It will turn out that every well-defined equilibrium converges to a BGE over time, so these conditions are natural long-run properties of the economy. The rest of this section characterizes a BGE; subsequent sections generalize to a transition path.

Before proceeding, I impose the parametric restriction that the size of quality innovations is larger than the productivity gap between firm types: $\lambda > a_h/a_\ell$. This means that, if a low-productivity firm innovates on a high-productivity blueprint, it will choose to enter and produce. If $\lambda < a_h/a_\ell$, then low-productivity firms never compete in product markets with high-productivity firms. This restriction is not strictly necessary, as low-productivity firms can still maintain positive market share via new-variety creation. However, $\lambda > a_h/a_\ell$ will prove true in the estimation, so I impose it up front to simplify exposition.

3.2 Equilibrium conditions

I next lay out the equilibrium conditions for a balanced-growth path. Appendix C provides a detailed derivation of all of these conditions for any equilibrium (not just a BGE).

Production and markups   Household consumption allocations $\{c_{ijt}\}$ maximize utility (6) and (7). Aggregating to total consumption (and production) Y_{ijt}, we have demand curves

$$p_{ijt} = q_{ijt} \left(\frac{Y_{jt}}{Y_t} \right)^{-1/\eta} p_t. \tag{16}$$

Aggregate output is the numeraire, so the aggregate price index $p_t = 1$. Note that this demand restriction means that any two firms i and i' producing a version of variety j must equate *quality-adjusted* prices:

$$\frac{p_{ijt}}{q_{ijt}} = \frac{p_{i'jt}}{q_{i'jt}}, \tag{17}$$

consistent with the assumption of perfect substitution within product markets.

Firms choose prices subject to (16) and (17) to maximize operating profits (11). Under Bertrand competition, only the firm with the lowest quality-adjusted marginal cost $w_{\mathrm{p}t}/(a_i q_{ijt})$ will produce variety j. We call this producer the *leader* and index it by i_1; we call the firm with the next lowest quality-adjusted marginal cost the *follower* with index i_2. The leader sets its price either to maximize its monopoly profits or, if necessary, to dissuade the follower from producing. Specifically, the leader's markup over marginal cost equals

$$\mu_{ijt} \equiv \frac{p_{ijt}}{w_{\mathrm{p}t}/a_i} = \min\left\{ \frac{a_{i_1} q_{i_1 jt}}{a_{i_2} q_{i_2 jt}}, \frac{\eta}{\eta - 1} \right\}. \tag{18}$$

The term $\eta/(\eta - 1)$ is the monopoly markup: the optimal markup in the presence of no competitor. The term $(a_{i_1} q_{i_1 jt})/(a_{i_2} q_{i_2 jt})$ is the marginal-cost advantage of the leader: if the marginal-cost advantage is sufficiently small, then the leader must engage in limit pricing by setting its quality-adjusted price equal to the quality-adjusted marginal cost of the follower.

This solution implies three possible market structures, comprised of two leader productivity types $i_1 \in \{h, \ell\}$ and three follower types $i_2 \in \{h, \ell, \varnothing\}$, where $i_2 = \varnothing$ means that there is no follower (a monopoly product). Let $\mu_{i_1 i_2}$ denote the corresponding markup. For a monopoly product ($i_2 = \varnothing$), the marginal-cost advantage is infinite, so $\mu_{i_1 \varnothing} = \eta/(\eta - 1)$. For a market with a follower, the quality advantage is always $q_{i_1 jt}/q_{i_2 jt} = \lambda$, so the markup equals $\mu_{i_1 i_2} = \min\{\lambda a_{i_1}/a_{i_2}, \eta/(\eta - 1)\}$.

Revenues and profits The above conditions together imply that the revenue share of firm i producing variety j equals

$$\frac{p_{ijt} Y_{ijt}}{Y_t} = \frac{1}{N_t} \left(\frac{a_i q_{ijt}}{\bar{a} Q_t} \right)^{\eta - 1} \left(\frac{\mu_{ijt}}{\mathcal{M}_{\eta - 1}} \right)^{-(\eta - 1)}. \tag{19}$$

The aggregates $\{\bar{a}, Q_t, \mathcal{M}_{\eta - 1}\}$ are averages of $\{a_i, q_{ijt}, \mu_{ijt}\}$, defined below in (23), (24), and (C.30). Because $\eta > 1$, a firm generates more revenue for a product than average if it has higher productivity and quality. It earns lower revenues if it charges a higher markup, because high prices reduce demand (16). Operating profits can be expressed in terms of

23

revenues and markups as

$$\Pi_{ijt} = \left(1 - \frac{1}{\mu_{ijt}}\right) p_{ijt} Y_{ijt}. \tag{20}$$

Higher markups evidently raise profits relative to revenues.

Aggregate output and wages Aggregate output (and consumption) equals

$$Y_t = A_t L_{\mathrm{pt}}, \tag{21}$$

where total-factor productivity (TFP, also output per labor hour) is defined as

$$A_t \equiv \underbrace{N_t^{1/(\eta-1)}}_{\text{variety}} \times \underbrace{Q_t}_{\text{quality}} \times \underbrace{\bar{a}}_{\text{productivity}} \times \underbrace{\Omega}_{\text{static misalloc.}}. \tag{22}$$

The first component of TFP is the love-of-variety effect, which is magnified if varieties are less substitutable (η is closer to one). The second term is *aggregate quality*, which is defined as an average of blueprint qualities:[31]

$$Q_t \equiv \left(\frac{1}{N_t} \int_0^{N_t} q_{jt}^{\eta-1} dj\right)^{1/(\eta-1)}. \tag{23}$$

The latter two components capture the allocation of resources across producers. *Aggregate productivity* is defined[32]

$$\bar{a} \equiv \left(\frac{1}{N_t} \int_0^{N_t} \left(\frac{q_{jt}}{Q_t}\right)^{\eta-1} a_{jt}^{\eta-1} dj\right)^{1/(\eta-1)}, \tag{24}$$

a quality-weighted average of firm-level productivity. The final term, *static misallocation* Ω, is a wedge that summarizes the extent to which actual labor allocations L_{ijt}/L_{pt} differ from

---

[31]Note the re-indexing of $\{a_i, q_{ijt}, \mu_{ijt}\}$ to $\{a_{jt}, q_{jt}, \mu_{jt}\}$, because each variety has only one producer.

[32]I call this "aggregate productivity" even though it puts weight on both qualities q_j and productivities a_j. The reason is that, if there is no heterogeneity in productivity ($a_h = a_\ell = 1$), then we get $\bar{a} = 1$ for any quality distribution; but if there is no quality heterogeneity ($q_{jt}/Q_t = 1$, $\forall j$), then $\bar{a}_t \in [a_\ell, a_h]$. In other words, $\bar{a}_t$ is principally a measure of how blueprints are distributed across productivity types a_i.

the optimal allocation under the current blueprint distribution. Specifically, defining

$$A_t^* \equiv \max_{\{L_{ijt}/L_{\mathrm{p}t}\}} \frac{Y_t}{L_{\mathrm{p}t}} \tag{25}$$

as the maximum level of output per labor hour, static misallocation equals the ratio

$$\Omega = \frac{A_t}{A_t^*} \in (0, 1]. \tag{26}$$

As shown in Appendix C.1.5, this wedge attains its maximum ($\Omega = 1$) if and only if there is no cross-product markup heterogeneity ($\mu_{jt} = \mu, \forall j$). As in Peters (2020), heterogeneous market power distorts demand and results in a misallocation of production labor across blueprints. This is why I refer to Ω as static misallocation.[33] Appendix C.1.3 provides a closed-form expression for Ω in terms of the distribution of blueprint characteristics.

The aggregate production wage equals

$$w_{\mathrm{p}t} = \Lambda \frac{Y_t}{L_{\mathrm{p}t}}, \tag{27}$$

where the (production) labor share of output $\Lambda = w_{\mathrm{p}t}L_{\mathrm{p}t}/Y_t$ is a sales-weighted average of product-level labor shares (inverse markups μ_{jt}^{-1}):

$$\Lambda = \int_0^{N_t} \frac{p_{jt}Y_{jt}}{Y_t} \mu_{jt}^{-1} dj. \tag{28}$$

It follows that a natural definition of the aggregate markup is the inverse Λ^{-1}, as in Edmond, Midrigan, and Xu (2023). Recall that there are $(1-\bar{s})L_t$ production laborers who each choose how many labor hours l_{p} to work. Combining the labor demand curve (27) with the labor supply curve implied by utility (5), we get that labor hours per production worker equal

$$l_{\mathrm{p}} = \frac{L_{\mathrm{p}t}}{(1-\bar{s})L_t} = \left(\frac{\Lambda}{\chi(1-\bar{s})}\right)^{\zeta/(1+\zeta)}. \tag{29}$$

---

[33]Note that Ω is only a "static" misallocation measure because it takes as given the current blueprint distribution, which can only be changed over time by R&D and M&A technologies. It also corresponds to the allocation that maximizes output per labor hour ($Y_t/L_{\mathrm{p}t}$), not necessarily output per capita (Y_t/L_t).

Each production worker exerts less effort if the aggregate markup Λ^{-1} is higher, because higher markups reduce wages and therefore disincentivize work.

Rescaled firm's problem A key step in solving the balanced-growth equilibrium is appropriately rescaling the equilibrium conditions. To do this, we can show that profits and firm values grow one-for-one with output per capita Y_t/L_t. From (19) and (20) above, the profits of any given blueprint can be rewritten[34]

$$\Pi_{ijt} = \pi_{i_1 i_2 \hat{q}} \frac{Y_t}{L_t}. \tag{30}$$

where $\pi_{i_1 i_2 \hat{q}}$ is a constant that depends on the current leader $i = i_1 \in \{h, \ell\}$, follower $i_2 \in \{h, \ell, \varnothing\}$, and relative quality $\hat{q} \equiv \log(q_{jt}/Q_t)$. From this fact, it can be shown that a firm's value is additively separable across blueprints:

$$V_{it} = \sum_{j=1}^{n_{it}} V_{ijt}, \quad \text{where} \quad V_{ijt} = v_{i_1 i_2 \hat{q}} \frac{Y_t}{L_t}. \tag{31}$$

The constant $v_{i_1 i_2 \hat{q}}$ denotes the rescaled value of a blueprint with those characteristics. The separability in (31) means that the equilibrium can be solved in terms of blueprint values $\{v_{i_1 i_2 \hat{q}}\}$ instead of firm values, which dramatically reduces the dimensionality of the problem. Before characterizing the solution for these blueprint values, let us derive firms' optimal R&D and M&A policies, taking blueprint values as given.

R&D policies The firm allocates $X_{\mathrm{n}it}$ skilled laborers to new-variety creation (13) and $X_{\mathrm{q}it}$ to quality innovation (12) at the skilled wage $w_{\mathrm{x}t}$. The optimal number of research hours employed for new-variety creation equates the marginal cost of research with its *marginal value product*:

$$w_{\mathrm{x}t} = \underbrace{\varepsilon(1-\alpha)\varphi(n_{it}/X_{\mathrm{n}it})^{1-\varepsilon}}_{\text{marginal product of research}} \times \underbrace{\bar{v}_{\mathrm{n}it}}_{\text{marginal value}} \tag{32}$$

---

[34]There is some abuse of notation here and in what follows: to be precise, we should say that $\Pi_{ijt} = \pi_{i_{1jt} i_{2jt} \hat{q}_{jt}} Y_t/L_t$, where $\{i_{1jt}, i_{2jt}, \hat{q}_{jt}\}$ is the set of variety j's characteristics at time t.

The blueprint values in (31) and the assumption that new varieties have initial quality $q_{jt} = \lambda Q_t$ imply that the marginal value of successful new-variety innovation equals

$$\bar{v}_{\mathrm{n}it} = \bar{v}_{\mathrm{n}i_1}\frac{Y_t}{L_t}, \quad \text{where} \quad \bar{v}_{\mathrm{n}i_1} = v_{i_1\varnothing,\log\lambda}. \tag{33}$$

Likewise, the skilled wage (solved below) grows with per-capita output: $w_{\mathrm{x}t} = \hat{w}_{\mathrm{x}}Y_t/L_t$. These facts imply the policy

$$x_{\mathrm{n}i_1} \equiv \frac{X_{\mathrm{n}it}}{n_{it}} = \left(\frac{\varepsilon(1-\alpha)\varphi\bar{v}_{\mathrm{n}i_1}}{\hat{w}_{\mathrm{x}}}\right)^{1/(1-\varepsilon)}. \tag{34}$$

Fixing the wage $\hat{w}_{\mathrm{x}}$, the firm spends relatively more on new-variety creation if the technology is more productive ($(1-\alpha)\varphi$ is high) or the resulting blueprint is more valuable ($\bar{v}_{\mathrm{n}i_1}$ is high).

By the exact same logic, the optimal allocation to quality innovations is

$$x_{\mathrm{q}i_1} \equiv \frac{X_{\mathrm{q}it}}{n_{it}} = \left(\frac{\varepsilon\alpha\varphi\bar{v}_{\mathrm{q}i_1}}{\hat{w}_{\mathrm{x}}}\right)^{1/(1-\varepsilon)}, \tag{35}$$

where the expected value of a blueprint from successful quality innovation equals

$$\bar{v}_{\mathrm{q}i_1} = \sum_{i'\in\{h,\ell\}}\int_{-\infty}^{\infty} v_{i_1i',\hat{q}+\log\lambda}(f_{i'h\hat{q}} + f_{i'\ell\hat{q}} + f_{i'\varnothing\hat{q}})d\hat{q}. \tag{36}$$

The densities $f_{i_1i_2\hat{q}} \equiv N_{i_1i_2\hat{q}t}/N_t$ denote the share of blueprints with these characteristics; thus, in this context, $f_{i'i_2\hat{q}}$ represents the conditional probability that the firm finds and innovates on a blueprint with current leader $i_1 = i'$, follower i_2, and relative quality $\hat{q}$.

The policies of entrepreneurs—skilled laborers working to come up with new ideas and firms—are identical to those of single-product firms. Specifically, each skilled laborer of prospective type a_i dedicates $\{x_{\mathrm{E}ni}, x_{\mathrm{E}qi}\} = \{x_{\mathrm{n}i}, x_{\mathrm{q}i}\}$ labor hours to entrepreneurial efforts.

M&A policies High-productivity firms are more profitable producers and thus ascribe a higher present value to each blueprint type. It is therefore mutually beneficial for high-productivity firms to buy blueprints from low-productivity firms. Suppose a high-productivity firm identifies a low-productivity target blueprint ($i_1 = \ell$) with follower i_2 and relative quality

$\hat{q}$. The deal price that solves the Nash bargaining problem is

$$v_{\mathrm{m}i_2\hat{q}} = \underset{v}{\operatorname{argmax}} \left\{ (v_{hi_2\hat{q}} - v)^\varrho (v - v_{\ell i_2\hat{q}})^{1-\varrho} \right\} = \varrho v_{\ell i_2\hat{q}} + (1-\varrho) v_{hi_2\hat{q}}. \tag{37}$$

The buyer gets a fraction ϱ of the surplus $v_{hi_2\hat{q}} - v_{\ell i_2\hat{q}}$ and the seller gets the remaining $1 - \varrho$.

In order for this deal to take place, the high-productivity firm must first successfully find a low-productivity blueprint. The skilled wage paid to M&A search is the same as that paid to researchers, so the optimal allocation to the M&A search technology (14) is

$$x_{\mathrm{m}i_1} \equiv \frac{X_{\mathrm{m}it}}{n_{it}} = \left(\frac{\varepsilon \psi \bar{v}_{\mathrm{m}i_1}}{\hat{w}_{\mathrm{x}}} \right)^{1/(1-\varepsilon)}, \tag{38}$$

where $\bar{v}_{\mathrm{m}i_1}$ is the expected value of the acquired blueprint *net* of the deal price,

$$\bar{v}_{\mathrm{m}i_1} = \begin{cases} \varrho f_\ell \sum_{i_2 \in \{h,\ell\}} \int_{-\infty}^{\infty} (v_{hi_2\hat{q}} - v_{\ell i_2\hat{q}}) \frac{f_{\ell i_2\hat{q}}}{f_\ell} d\hat{q} & \text{if} \quad i_1 = h, \\ 0 & \text{if} \quad i_1 = \ell. \end{cases} \tag{39}$$

For low-productivity firms, the expected benefit is zero, because they would never buy any blueprint that they find; these firms therefore invest nothing in the M&A search technology ($x_{\mathrm{m}\ell} = 0$). For high-productivity firms, the expected present value is higher if it has more bargaining power (ϱ) and if it is more likely to find a low-productivity blueprint (f_ℓ).

Skilled wages The skilled wage $w_{\mathrm{x}t}$ clears the labor market by setting the demand for R&D and M&A equal to the supply of skilled labor hours. Appendix C.3.2 shows that $w_{\mathrm{x}t} = \hat{w}_{\mathrm{x}} Y_t / L_t$, where $\hat{w}_{\mathrm{x}}$ is given by (C.127). The rescaled wage $\hat{w}_{\mathrm{x}}$ is increasing in R&D and M&A productivity φ and ψ and in expected blueprint valuations, because these forces increase demand for skilled labor. Higher wages also increase the number of hours worked by each skilled laborer:[35]

$$l_{\mathrm{x}} = \frac{L_{xt}}{\bar{s} L_t} = \left(\frac{\hat{w}_{\mathrm{x}}}{\chi} \right)^\zeta. \tag{40}$$

Thus, even for a fixed research population $\bar{s} L_t$, higher innovation incentives increase the equilibrium supply of research inputs through this intensive margin.

---

[35]Labor hours l_{x} include hours worked for both incumbents and entrepreneurship.

Economic growth   Equilibrium scales with output per capita $y_t = Y_t/L_t$. Before deriving the growth rate, $g_y \equiv \dot{y}_t/y_t$, let us establish some notation. Define a firm's new-variety-creation intensity as

$$\Phi_{\mathrm{n}i} \equiv \frac{\varphi(1-\alpha)X_{\mathrm{n}it}^{\varepsilon}n_{it}^{1-\varepsilon}}{n_{it}} = \varphi(1-\alpha)x_{\mathrm{n}i_1}^{\varepsilon}. \tag{41}$$

Define $\Phi_{\mathrm{q}i}$ and Ψ_i analogously for quality innovation and M&A search, respectively. Lastly, define $\nu_i \equiv \omega_i \bar{s} L_t/N_t$ as the number of potential new type-a_i entrants per existing blueprint (where, again, $\omega_h = \omega$ and $\omega_\ell = 1 - \omega$).

On the balanced-growth path, output per capita grows in proportion to variety $N_t^{1/(\eta-1)}$ and aggregate quality Q_t (see (22)), so its growth rate equals[36]

$$g_y = \frac{1}{\eta - 1}g_N + g_Q. \tag{42}$$

The growth rate of the number of varieties equals

$$g_N \equiv \frac{\dot{N}_t}{N_t} = \Phi_{\mathrm{n}h}(f_h + \nu_h) + \Phi_{\mathrm{n}\ell}(f_\ell + \nu_\ell) = g_L \tag{43}$$

The first equality captures the fact that new-variety creation is higher when there is a higher average new-variety intensity across both incumbents and new entrants. The second equality is a necessary condition for balanced growth: the ratio L_t/N_t must be constant.[37] The growth rate of aggregate quality equals

$$g_Q \equiv \frac{\dot{Q}_t}{Q_t} = \frac{\lambda^{\eta-1} - 1}{\eta - 1}(g_N + \delta), \tag{44}$$

where

$$\delta \equiv \Phi_{\mathrm{q}h}(f_h + \nu_h) + \Phi_{\mathrm{q}\ell}(f_\ell + \nu_\ell) \tag{45}$$

is the rate of *creative destruction*, the intensity with which incumbents and entrants innovate upon existing blueprints. The rate δ is increasing in the creative destruction parameter α. Quality growth (44) is increasing in g_N because new varieties have above-average quality.

---

[36]The constants $\{\bar{a}, \Omega, l_\mathrm{p}\}$ affect the *level* of output, but not the growth rate; they will become relevant for growth only along a transition path in which there is net reallocation.

[37]Intuitively, if $g_{Nt} < g_L$, then L_t/N_t will rise, increasing g_{Nt} until it is equal to g_L.

The larger the quality innovations λ, the faster the rate of quality growth.

Market valuations The policies above depend on and determine the blueprint market values $\{v_{i_1 i_2 \hat{q}}\}$. Using the fact that profits are separable in quality, $\pi_{i_1 i_2 \hat{q}} = \pi_{i_1 i_2} e^{(\eta-1)\hat{q}}$, we can further decompose the blueprint values into two parts:

$$v_{i_1 i_2 \hat{q}} = \underbrace{v_{i_1 i_2}^{\mathrm{P}} e^{(\eta-1)\hat{q}}}_{\text{product value}} + \underbrace{v_{i_1}^{\mathrm{G}}}_{\text{growth value}} \tag{46}$$

The product value represents the present value of profits directly generated by the blueprint; the growth value represents the contribution of the blueprint to generating future profits via R&D and M&A. Because we are in a riskless economy, every firm is discounted by the riskfree rate,

$$r_f = \rho + g_y. \tag{47}$$

The coefficients $\{v_{i_1 i_2}^{\mathrm{P}}, v_{i_1}^{\mathrm{G}}\}$ are characterized by a system of Hamilton-Jacobi-Bellman (HJB) equations. The HJBs for product values are

$$\underbrace{\rho v_{i_1 i_2}^{\mathrm{P}}}_{\text{discount}} = \underbrace{\pi_{i_1 i_2}}_{\text{profits}} + \underbrace{\mathbb{I}_{i_1=\ell} \Psi_h f_h (1-\varrho)(v_{h i_2}^{\mathrm{P}} - v_{\ell i_2}^{\mathrm{P}})}_{\text{M\&A exit}} - \underbrace{\delta v_{i_1 i_2}^{\mathrm{P}}}_{\text{creat. destr.}} - \underbrace{(\eta-1)g_Q v_{i_1 i_2}^{\mathrm{P}}}_{\text{quality depr.}}. \tag{48}$$

Product values are increasing in profits and, for low-productivity firms, the possibility of being acquired at a premium; they are decreasing in the risk of creative destruction δ and in the rate of relative quality depreciation $(\eta-1)g_Q$. The HJBs for growth values are

$$\underbrace{\rho v_{i_1}^{\mathrm{G}}}_{\text{discount}} = \underbrace{\Phi_{\mathrm{n}i_1} \bar{v}_{\mathrm{n}i_1}}_{\text{new variety}} + \underbrace{\Phi_{\mathrm{q}i_1} \bar{v}_{\mathrm{q}i_1}}_{\text{quality innov.}} + \underbrace{\Psi_{i_1} \bar{v}_{\mathrm{m}i_1}}_{\text{acquisition}} - \underbrace{\hat{w}_{\mathrm{x}}(x_{\mathrm{n}i_1} + x_{\mathrm{q}i_1} + x_{\mathrm{m}i_1})}_{\text{expenditures}}$$
$$+ \underbrace{\mathbb{I}_{i_1=\ell} \Psi_h f_h (1-\varrho)(v_h^{\mathrm{G}} - v_\ell^{\mathrm{G}})}_{\text{M\&A exit}} - \underbrace{\delta v_{i_1}^{\mathrm{G}}}_{\text{creat. destr.}} \tag{49}$$

Growth values are higher for firms that do more R&D and M&A—high-productivity firms, in equilibrium. Like product values, growth values are reduced by creative destruction and increased for low-productivity firms if it is more likely that they will get acquired.

Distribution   Like market values, the equilibrium blueprint distribution is summarized over the three product characteristics $\{i_1, i_2, \hat{q}\}$. Recall that $N_{i_1 i_2 \hat{q} t}$ denotes the mass of blueprints with these characteristics and $f_{i_1 i_2 \hat{q}} \equiv N_{i_1 i_2 \hat{q} t}/N_t$ the share. The evolution of this distribution is characterized by a system of Kolmogorov Forward equations (KFEs), which express flows into and out of product markets in terms of firms' R&D and M&A decisions.

The KFEs follow from the fact that the net flow of blueprints equals

$$\dot{N}_{i_1 i_2 \hat{q} t} = N_t \dot{f}_{i_1 i_2 \hat{q} t} + \dot{N}_t f_{i_1 i_2 \hat{q} t} = g_N N_t f_{i_1 i_2 \hat{q}}, \tag{50}$$

where the second equality imposes a stationary distribution. This net flow equals the sum of constituent flows from R&D and M&A. For monopoly products ($i_2 = \varnothing$), the KFEs are

$$\underbrace{g_N f_{i_1 \varnothing \hat{q}}}_{\text{net flow}} = \underbrace{[\partial_{\hat{q}} f_{i_1 \varnothing \hat{q}}] g_Q}_{\text{quality depr.}} + \underbrace{\Phi_{n i_1}(f_{i_1} + \nu_{i_1}) \mathbb{I}_{\hat{q} = \log \lambda}}_{\text{new varieties}} + \underbrace{(\mathbb{I}_{i_1 = h} - \mathbb{I}_{i_1 = \ell}) \Psi_h f_h f_{\ell \varnothing \hat{q}}}_{\text{acquisitions}} - \underbrace{\delta f_{i_1 \varnothing \hat{q}}}_{\text{creat. destr.}}. \tag{51}$$

The first term on the right-hand side captures the change from quality depreciation. The second term is the inflow of blueprints from new-variety creation by both incumbents (f_{i_1}) and new entrants (ν_{i_1}). The third term represents the flows from high-productivity firms buying low-productivity monopoly products—note the asymmetry in i_1. The fourth and final term is the outflow of blueprints from creative destruction by competitors and entrants. The KFEs for competitive markets ($i_2 \in \{h, \ell\}$) are

$$\underbrace{g_N f_{i_1 i_2 \hat{q}}}_{\text{net flow}} = \underbrace{[\partial_{\hat{q}} f_{i_1 i_2 \hat{q}}] g_Q}_{\text{quality depr.}} + \underbrace{\Phi_{q i_1}(f_{i_1} + \nu_{i_1})(f_{i_2 h, \hat{q} - \log \lambda} + f_{i_2 \ell, \hat{q} - \log \lambda} + f_{i_2 \varnothing, \hat{q} - \log \lambda})}_{\text{quality innovations}}$$

$$+ \underbrace{(\mathbb{I}_{i_1 = h} - \mathbb{I}_{i_1 = \ell}) \Psi_h f_h f_{\ell i_2 \hat{q}}}_{\text{acquisitions}} - \underbrace{\delta f_{i_1 i_2 \hat{q}}}_{\text{creat. destr.}}. \tag{52}$$

The only difference from (51) is that, instead of inflows from new varieties, there are inflows from quality innovations on incumbent firms' blueprints.

The total type-a_{i_1} blueprint share $f_{i_1} = N_{i_1 t}/N_t$ is the main determinant of total revenue and value shares in the economy. Aggregating the KFEs across i_2 and $\hat{q}$ implies that the

total type-a_{i_1} blueprint share satisfies

$$\underbrace{g_N f_{i_1}}_{\text{net flow}} = \underbrace{\Phi_{\mathrm{n}i_1}(f_{i_1} + \nu_{i_1})}_{\text{new varieties}} + \underbrace{\Phi_{\mathrm{q}i_1}(f_{i_1} + \nu_{i_1}) - \delta f_{i_1}}_{\text{net quality innovations}} + \underbrace{(\mathbb{I}_{i_1=h} - \mathbb{I}_{i_1=\ell})\Psi_h f_h f_\ell}_{\text{acquisitions}}. \tag{53}$$

This aggregated KFE displays the key forces driving the total high-productivity blueprint share f_h. On the one hand, f_h will be driven up by the fact that these firms do more R&D ($\Phi_h > \Phi_\ell$) and that M&A is asymmetric (the last term of (53)). On the other hand, f_h will be lower if fewer entrepreneurs are high-productivity to begin with ($\omega_h < \omega_\ell$, so $\nu_h < \nu_\ell$).

3.3   Key equilibrium properties

In general, the transition path between BGEs can look different from the initial and final BGEs themselves. This limits the usefulness of comparative statics. With that in mind, there are two key implications of the model that hold unambiguously both across BGEs and along the transition path, the first concerning the consequences of heterogeneous productivity and the second concerning the consequences of changing innovation productivity.

Equilibrium consequences of heterogeneous productivity   In equilibrium, high-productivity firms endogenously arise as high-profit, high-growth, and high-valuation firms.

Proposition 1 (Comparisons between productivity types)**.** *Consider any blueprint with follower of type $a_{i_2} \in \{a_h, a_\ell\}$ and relative quality $\hat{q} \in \mathbb{R}$.*

1. *High-productivity producers earn higher revenues, higher markups, and higher profits:*

$$p_{hi_2\hat{q}} y_{hi_2\hat{q}} \geq p_{\ell i_2\hat{q}} y_{\ell i_2\hat{q}}, \quad \mu_{hi_2} \geq \mu_{\ell i_2}, \quad and \quad \pi_{hi_2\hat{q}} > \pi_{\ell i_2\hat{q}}. \tag{54}$$

2. *Every high-productivity firm spends more per blueprint on R&D and M&A:*

$$w_{\mathrm{x}t} x_{\mathrm{k}h} > w_{\mathrm{x}t} x_{\mathrm{k}\ell} \quad \forall\, \mathrm{k} \in \{\mathrm{n}, \mathrm{q}, \mathrm{m}\}. \tag{55}$$

3. *The product value and growth value of the blueprint are strictly higher for a high-*

productivity producer than for a low-productivity producer:

$$v_{hi_2}^{\mathrm{P}} > v_{\ell i_2}^{\mathrm{P}} \quad and \quad v_h^{\mathrm{G}} > v_\ell^{\mathrm{G}}, \quad so \quad v_{hi_2\hat{q}} > v_{\ell i_2\hat{q}}. \tag{56}$$

See Appendix C.6 for a proof. The first statement simply follows from the fact that high-productivity firms are, by assumption, able to produce the same blueprint at a lower marginal cost. The second and third facts follow from the first. Because high-productivity firms earn higher profits, their future growth opportunities are also more profitable. They hence expect a higher present value from successfully innovating, and choose a higher R&D intensity in equilibrium (recall the first-order condition (32)). High-productivity firms are therefore not just more profitable today; they also have higher expected *growth rates*, and, in turn, tend to have higher valuations relative to sales and cashflows. Put another way, static productivity differences a_h/a_ℓ are compounded in present value terms.

Only high-productivity firms will choose to invest in M&A search. A high-productivity firm ascribes a strictly higher valuation to every given blueprint, and is therefore willing to buy a low-productivity firm's blueprint at a premium relative to its current traded market value. The reverse is not true. This has two macroeconomic implications. First, from the KFE (53), one can see that M&A tends to reallocate market share toward high-productivity (hence high-valuation) firms. In particular, this reallocation matters more for the distribution when M&A is high *relative* to R&D. Second, because M&A transactions create a net surplus, they tend to increase the aggregate valuation of the market.

Equilibrium consequences of changing innovation productivity φ Figure 4 plots comparative statics with respect to φ for three moments of interest. The first, plotted in Panel A, is the ratio of R&D-to-value. As will be explained in detail in the next section, lower innovation productivity decreases the R&D-to-value ratio because it drives a wedge between the present value of profits (which affects both the numerator and denominator) and R&D effort. Panel B shows that, as φ falls, the M&A-to-R&D ratio rises. This is both because R&D falls and because, within high-productivity firms, M&A spending increases.

As just explained, a higher M&A-to-R&D ratio implies more net reallocation toward

33

Figure 4: Comparative statics for innovation productivity across balanced-growth paths
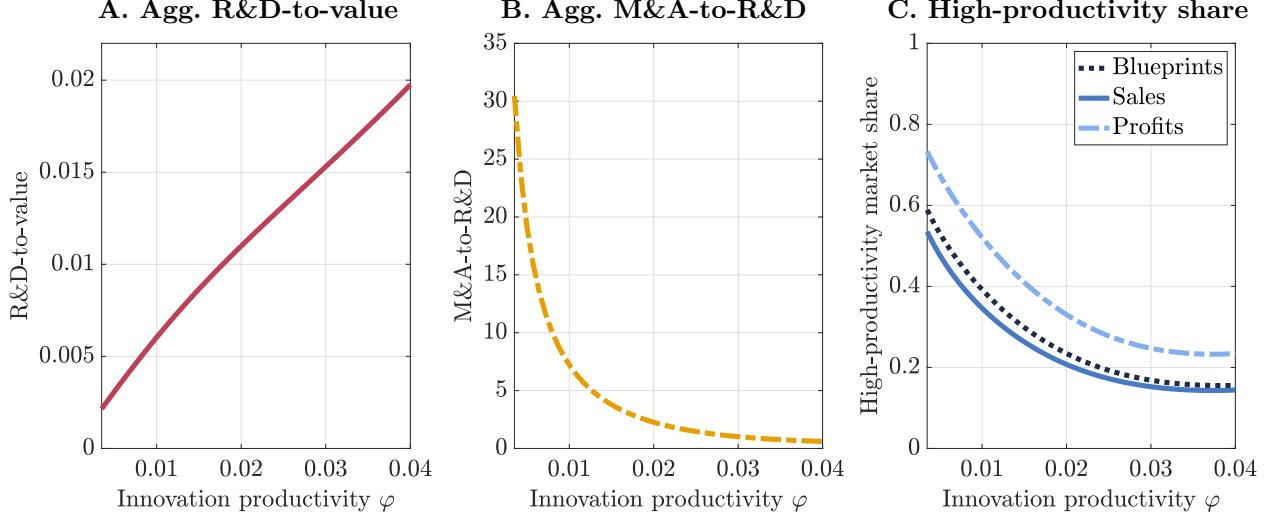
The figure plots comparative statics across balanced-growth equilibria with respect to innovation productivity φ. Panels A and B show the aggregate ratios of R&D-to-value and M&A-to-R&D, respectively. Panel C plots the total share of blueprints, sales, and profits of high-productivity firms. Model parameters take the values estimated in Section 4. See Appendix E.1 for details of the numerical solution method.

high-productivity firms. Panel C of Figure 4 plots the consequence of this for the total high-productivity market share. As innovation productivity φ falls, high-productivity firms gain market share in terms of blueprints, sales, and profits. M&A is the essential mechanism here: were we to assume no M&A ($\psi = 0$), then virtually nothing would happen to high-productivity market shares. One can see how the three predictions in Figure 4 qualitatively align with the motivating empirical facts in Section 1.

4 Estimation of the model and transition path

4.1 Overview and methods

The goal of this section is to quantify the decline in innovation productivity and its consequences for the U.S. economy over the last fifty years. To this end, I first estimate the model parameters using firm-level micro data from before the 1980s, assuming the U.S. was on a balanced-growth path. I then estimate the transition path of innovation productivity, φ_t, using the time series of R&D-to-value. Finally, in Section 5, I feed these estimates into the model and study its equilibrium implications along the transition path.

The numerical methods for solving and simulating the model are laid out in detail in Appendix E, and estimation details are in Appendix F. The solution methods build on the continuous-time iterative methods in Achdou, Han, Lasry, Lions, and Moll (2022) and related work. To solve for the balanced-growth equilibrium, I develop a numerical method that iterates over the distribution $\{f_{i_1 i_2 \hat{q}}\}$ and the ratio L_t/N_t. To solve the transition path, I use a forward-backward iteration scheme between balanced-growth paths. To simulate firms, I develop a new method that exploits analytical results for the distribution of size and age (derived in Appendix D). The estimation is via simulated method of moments (SMM), using the numerical SMM procedure of Catherine, Ebrahimian, Sraer, and Thesmar (2023).

4.2 Estimation of the balanced-growth equilibrium

Externally calibrated parameters Before estimating the model, I calibrate eight parameters externally. Two come directly from data. The population growth rate g_L is 1.25%, the U.S. average from 1955–1985. The share of skilled workers in the labor force is 14.2%, which is the share of managers, engineers, and scientists computed by Acemoglu et al. (2018).

The elasticity of substitution between goods η is set equal to 2.9, the median estimate of Broda and Weinstein (2006). Households have log utility, meaning that they have a unit elasticity of intertemporal substitution (EIS). Because none of the moments of interest will depend on the labor disutility scalar χ, I normalize it to 1. For the Frisch elasticity of labor supply, I assume a value of 0.5, in the range of most empirical estimates from the labor literature (see Keane (2011) and Chetty (2012) for summaries).

The elasticity $\varepsilon \in (0, 1)$, which determines the relative importance of laborers versus existing ideas in the idea-production functions, is common in endogenous-growth models. Following Acemoglu et al. (2018), I set it to 0.5, in line with a large body of evidence from the microeconomic innovation literature (Blundell, Griffith, & Windmeijer, 2002; Hall & Ziedonis, 2001; Hall, 1993; Bloom, Griffith, & Van Reenen, 2002; Wilson, 2009).[38] Finally, I assume buyers and sellers have equal bargaining power in M&A negotiations ($\varrho = 0.5$).[39]

---

[38]For models that assume an adjustment cost instead of an idea-production function, $\varepsilon = 0.5$ corresponds to a quadratic cost function, which is also common (e.g., Klette & Kortum, 2004; Peters, 2020).

[39]One could instead estimate ϱ from the average merger premium in SDC Platinum, which is about 1.8× the target's pre-announcement value. Doing so yields an estimate close to 0.5. The main results are in

Table 1: Externally calibrated model parameters

	Parameter	Value	Source/Target/Assumption
Population			
Population growth	g_L	0.0125	Historical average
Skilled share of population	$\bar{s}$	0.142	Skilled labor force
Preferences			
Substitution elasticity between goods	η	2.9	Broda and Weinstein (2006)
Elasticity of intertemporal substitution		1	Log utility
Labor disutility	χ	1	Normalized
Frisch elasticity	ζ	0.5	Labor literature
Firms			
Production-function elasticity	ε	0.5	Innovation literature
Bargaining power	ϱ	0.5	Equal bargaining power

The table reports the baseline model parameters calibrated prior to the estimation procedure. See the main text for details.

Estimated parameters In the first stage of the estimation, I estimate seven model parameters in the initial balanced-growth path using micro and macro data:

$$\Theta = \left[\underbrace{\rho}_{\text{preferences}} \quad \underbrace{\varphi_0 \quad \lambda \quad \alpha}_{\text{innovation}} \quad \underbrace{\psi}_{\text{acquisition}} \quad \underbrace{a_h/a_\ell \quad \omega}_{\text{heterogeneity}} \right]. \tag{57}$$

Table 2 also lists these parameters with their descriptions. Note that the subscript on innovation productivity φ_0 indicates that this is the constant, pre-transition value.

In the rest of this section, I discuss the estimation of each parameter, focusing on how the parameter is identified by a moment in the data. For further details about moment computation and the SMM procedure, see Appendix F.

Discount factor ρ The rate of time preference ρ discounts all firms' cashflows by determining the growth-adjusted discount rate $r_f - g_y = \rho$. Market valuations are therefore higher (relative to sales and cashflows) when the discount rate ρ is lower. Thus, we can identify ρ using the aggregate value-to-sales ratio V_t/Y_t of the market.[40] In the data, I use the average

---

general not sensitive to ϱ because total surplus, not the split of surplus, is what matters for aggregate value.

[40]In Section 7 below, I investigate what happens when stock returns are risky. In this case, ρ can be replaced by a risk-adjusted discount rate that adds risk premia and subtracts precautionary savings demand. Thus, the estimation is still valid, but one should interpret the estimate for ρ as also including a risk adjustment.

Table 2: Estimated parameters for initial balanced-growth path

	Parameters		Identifying moments		
	Notation	Estimate	Moment	Data	Model
Preferences					
Discount factor	ρ	0.0174 (0.0097)	Agg. value-to-sales	1.0006 (0.0230)	1.0006
Innovation (R&D)					
Innovation productivity	φ_0	0.0412 (0.0036)	Agg. R&D-to-value	0.0204 (0.0008)	0.0204
Quality innovation size	λ	1.1380 (0.0204)	Agg. output growth	0.0248 (0.0040)	0.0248
Relative quality productivity	α	0.7944 (0.0061)	Firm entry rate	0.1191 (0.0048)	0.1191
Acquisition (M&A)					
M&A search productivity	ψ	0.0819 (0.0098)	Agg. M&A-to-value	0.0116 (0.0013)	0.0116
Firm heterogeneity					
Productivity gap	a_h/a_ℓ	1.1271 (0.0284)	St. dev. of value-to-sales	0.5806 (0.0303)	0.5806
High-prod. entrep. share	ω	0.0124 (0.0051)	St. dev. of profit-to-sales	0.5463 (0.0162)	0.5463

The table reports the baseline estimates of parameters in the balanced-growth equilibrium before the transition. The subscript on innovation productivity φ_0 indicates that this is the value on the balanced-growth path, before the transition commences. Cross-sectional standard deviations are weighted by sales and divided by the corresponding aggregate moment. Standard errors are in parentheses. See the main text and Appendix F for further details on the SMM estimation procedure and how each moment is computed.

from 1965–1975 and measure firm value as equity plus debt net of cash.

Initial innovation productivity φ_0 Identification of the initial φ_0 comes from the aggregate R&D-to-value ratio. I discuss the intuition for identification at length in Section 4.3 below when I estimate the transition path of $\{\varphi_t\}$. Because the R&D data only begin in 1975, I use the average ratio from 1975–1985 for this data moment.

Quality innovation step size λ Larger quality-innovation increments $\lambda > 1$ will mean that aggregate quality Q_t grows faster with each new idea. Hence, λ can be identified from the growth rate of aggregate output per capita

$$g_y = \frac{\lambda^{1-\eta}}{\eta - 1} g_L + \frac{\lambda^{1-\eta} - 1}{\eta - 1} \delta. \tag{58}$$

In the data, I use percentage growth in real output per capita from 1950–1985.

Relative quality productivity α The higher is $\alpha \in (0,1)$, the more likely it is that a new idea will be a quality innovation instead of a new variety. Thus, new ideas will tend to displace incumbents' existing ideas, increasing creative destruction δ and the rate of firm exit. Under balanced growth, the rates of firm entry and exit per existing firm are linked by the equation

$$\text{entry rate} = g_L + \text{exit rate}. \tag{59}$$

Hence, both the entry and exit rates can identify α. For the data moment, I use the rate of new-firm entry per existing firm in the Business Dynamics Statistics (BDS) from 1979–1985. Importantly, the BDS data include all firms, not just firms entering public markets.

M&A search productivity ψ Let $V_{\mathrm{m}t}$ denote aggregate M&A deal flow: the total dollar amount spent on completed acquisitions of firms (not on M&A search labor). We can identify ψ from the aggregate M&A-to-value ratio:

$$\frac{V_{\mathrm{m}t}}{V_t} = \underbrace{\Psi_h f_h f_\ell}_{\text{frequency}} \times \underbrace{\frac{\bar{v}_{\mathrm{m}}}{V_t/N_t}}_{\text{values}}. \tag{60}$$

This ratio has two components. The first is the frequency with which targets are found and acquired, which is higher when ψ (and hence Ψ_h) is higher. The second is the ratio of the average value of an acquired blueprint $\bar{v}_{\mathrm{m}}$ to the average value of an existing blueprint V_t/N_t. By scaling by aggregate value, we control for blueprint-level valuation changes (the second component), isolating the frequency component of M&A.

In the data, the numerator in the M&A-to-value ratio is the sum of the values of all mergers in a given year in the SDC Platinum database from 1978–1985. All acquirers are public firms, and acquisition targets include both public and private firms.

Firm heterogeneity parameters a_h/a_ℓ **and** ω The two parameters that govern ex ante firm heterogeneity are the productivity gap a_h/a_ℓ and the high-productivity share of entrepreneurs ω. Identifying the efficiency advantage a_h/a_ℓ is especially important for the

transition path, because this will determine the extent to which reallocation between firms affects macroeconomic aggregates. These parameters must be jointly identified, because both influence the (unobservable) market shares of high-productivity firms.[41]

First, fix a value of ω and consider a_h/a_ℓ. A larger gap has two effects. First, it increases the total market share of high-productivity firms. Second, it increases the cross-sectional difference between high- and low-productivity firms' profits and valuations. One moment that captures these effects is the *sales-weighted* cross-sectional variance of value-to-sales, defined as

$$\text{var}_{py}\left(\frac{V_{it}}{p_{it}y_{it}}\right) \equiv \int_0^{M_t} \frac{p_{it}y_{it}}{Y_t}\left(\frac{V_{it}}{p_{it}y_{it}} - \frac{V_t}{Y_t}\right)^2 di. \tag{61}$$

By the law of total variance, this sales-weighted variance can be decomposed across types as

$$\text{var}_{py}\left(\frac{V_{it}}{p_{it}y_{it}}\right) = \underbrace{\sum_{i'\in\{h,\ell\}} \frac{Y_{i't}}{Y_t}\text{var}_{py}\left(\frac{V_{it}}{p_{it}y_{it}}\bigg| a_i = a_{i'}\right)}_{\text{within-type variance}} + \underbrace{\sum_{i'\in\{h,\ell\}} \frac{Y_{i't}}{Y_t}\left(\frac{V_{i't}}{Y_{i't}} - \frac{V_t}{Y_t}\right)^2}_{\text{between-type variance}}, \tag{62}$$

where Y_{it} is the total sales of all type-a_i firms. The within-type term is largely invariant to a_h/a_ℓ, as within-type dispersion in value-to-sales is all driven by heterogeneous markups and quality. The between-type term is increasing in a_h/a_ℓ, because a_h/a_ℓ increases both the total sales share of high-productivity firms and the gap between total high and low value-to-sales ratios.[42] This between-type variance is the main source of identification.

The challenge to identifying a_h/a_ℓ using the variance of value-to-sales is that this moment is *also* increasing in ω. Specifically, raising ω will increase the total high-productivity sales share, increasing the between-type variance. Thus, in order to separately identify a_h/a_ℓ and ω, one could look for a moment that "controls" for the sales share. To do this, I use the sales-weighted cross-sectional variance of profit-to-sales. Intuitively, this moment depends on sales shares in the same way as the variance of value-to-sales; however, it is comparatively less sensitive to changes in a_h/a_ℓ, because market valuations compound differences in profits (through the present value of growth opportunities).

---

[41]As a robustness check, Appendix F.4 re-estimates the model under alternative identifying moments for a_h/a_ℓ and ω and shows that results are relatively unchanged.

[42]As a_h/a_ℓ becomes very large, this moment will begin to decrease as high-productivity firms accumulate 100% of the revenue share; however, quantitatively, these large values of a_h/a_ℓ are far outside of the relevant parameter space, so this is not a concern for identification. The same intuition holds for ω.

Empirically, these sales-weighted variances are computed across firms in Compustat from 1965–1975. To standardize units, the moments reported in Table 2 are standard deviations, scaled by the respective aggregate ratios (see Appendix F.2 for further discussion).

Discussion of estimation results Table 2 reports the estimation results. The model matches each moment well, and the relatively small standard errors on parameters validate their identification. Importantly, the estimate of φ_0 is very precise, suggesting that R&D-to-value is a good identifying moment (more on this in 4.3). Appendix F.3 presents additional corroborating evidence of identification. The estimated α is about 0.8, meaning that, as one might expect, it is easier to innovate on existing blueprints than to invent entirely new goods. High-productivity firms are estimated to be about 13% more efficient, and the low value of ω implies that they are uncommon among entrepreneurs.[43]

4.3 Estimation of the decline in innovation productivity

In the second stage of the estimation, I estimate the time series of innovation productivity $\{\varphi_t\}$ from 1975 to present, holding all other parameters fixed. I first discuss the intuition for identification from the R&D-to-value ratio, and then present the results.

Identification of innovation productivity φ_t As in the balanced-growth path, innovation productivity is identified by the aggregate R&D-to-value ratio. The intuition is as follows. As we saw in the first-order conditions (34) and (35), a firm's R&D spending is increasing in innovation productivity φ, increasing in the present value of new ideas $\bar{v}_{ni}$ and $\bar{v}_{qi}$, and decreasing in the skilled wage w_{xt}. The main challenge to identifying φ from R&D spending is controlling for valuation and wage changes. The insight for identification is that one can do this by dividing by aggregate value.

It is most straightforward to see this in the case with inelastic labor supply ($\zeta = 0$). Letting $X_{rt} \equiv X_{nt} + X_{qt}$ denote the aggregate quantity of researchers employed by incumbents,

---

[43]A low ω is also consistent with the amount of concentration in the data. If ω is low, the concentration of the firm-size distribution (i.e., the thickness of the Pareto tail) will be higher, because more blueprints will be owned by a smaller number of high-productivity firms. High values of ω will imply too little concentration. Appendix F.4 shows this by re-estimating the model using the Pareto tail coefficient of the employment distribution as an identifying moment; it implies an even lower value of ω and yields similar main results.

the equilibrium R&D-to-value ratio at any point in time equals[44]

$$\frac{w_{\mathrm{x}t}X_{\mathrm{r}t}}{V_t} = \underbrace{\varphi_t}_{\text{prod.}} \times \underbrace{\frac{\bar{v}_{\mathrm{r}t}}{V_t/N_t}}_{\text{values}} \times \underbrace{\varepsilon\left(1 + \frac{N_t}{\bar{s}L_t}\left(1 + \left(\frac{\psi}{\varphi_t}\frac{\bar{v}_{\mathrm{m}t}}{\bar{v}_{\mathrm{r}t}}\right)^{\frac{1}{1-\varepsilon}}\right)\right)^{-\varepsilon}}_{\text{substitution between factors and technologies}}, \qquad (63)$$

where the average present value of a new idea from R&D equals

$$\bar{v}_{\mathrm{r}t} \equiv \left(\sum_{i\in\{h,\ell\}} f_{it}\left([(1-\alpha)\bar{v}_{\mathrm{n}it}]^{\frac{1}{1-\varepsilon}} + [\alpha\bar{v}_{\mathrm{q}it}]^{\frac{1}{1-\varepsilon}}\right)\right)^{1-\varepsilon}. \qquad (64)$$

The ratio (63) moves one-for-one (in percentage terms) with innovation productivity φ_t. The second term, which is the ratio of the average new blueprint value to the average existing blueprint value, is essentially a constant, because parameter changes that increase valuations (e.g., a lower ρ) will affect the numerator and denominator roughly proportionately. This is why scaling R&D by value—and not, say, output—is key to identifying φ_t. The third term combines substitution effects between factors (researchers versus blueprints) and technologies (R&D versus M&A). For the latter tradeoff, if M&A search has comparatively higher returns, then it will increase the skilled wage and reduce R&D-to-value. Under the estimated parameters, this third term is quantitatively insensitive to all parameter changes.

Putting all of this together, we have that the (percent) changes in R&D-to-value and φ_t over any interval $[t, t+\tau]$ are roughly the same:

$$\Delta \log\left(\frac{w_{\mathrm{x},t+\tau}X_{\mathrm{r},t+\tau}}{V_{t+\tau}}\right) \approx \Delta \log \varphi_{t+\tau}. \qquad (65)$$

In the actual case of elastic labor supply $\zeta > 0$, there is also a small effect from the fact that skilled labor supply rises if φ_t or $\bar{v}_{\mathrm{r}t}$ rises. From a firm's perspective, this shows up through a change in the skilled wage, which decreases when labor supply increases.[45]

Appendix F.3 shows quantitatively that other parameter shifts in the model cannot explain the decline in R&D-to-value, verifying the above intuition. I show this by feeding other parameter transitions into the model and computing the resulting path of R&D-to-

---

[44]See Appendix F.3 for a derivation of this and for a discussion of the general case with $\zeta > 0$.

[45]Appendix F.3 explains why these labor-supply effects are negligible in response to changes in φ_t.

Figure 5: Estimated innovation productivity and implied R&D-to-value
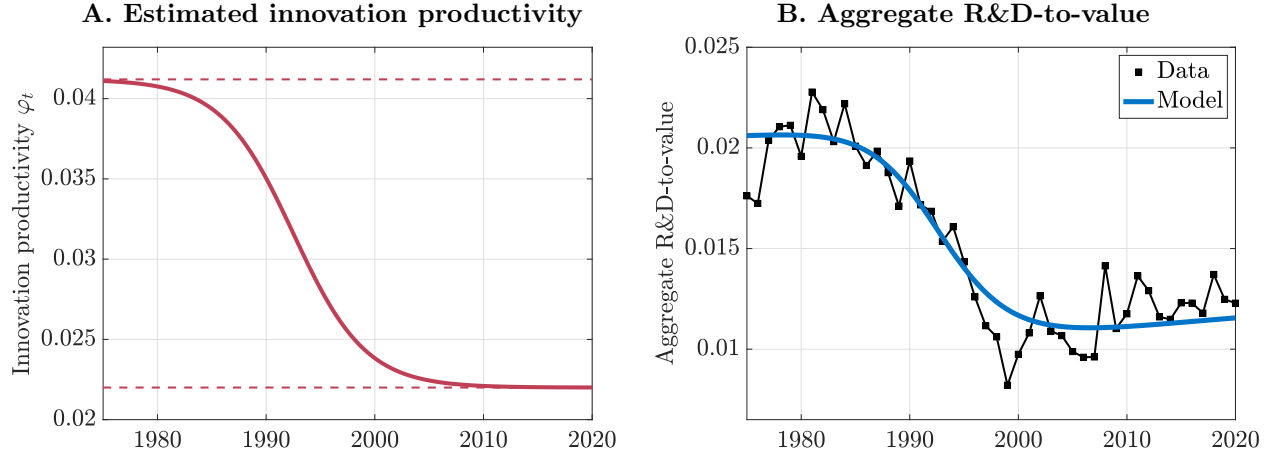
The figure shows the estimated time series of innovation productivity $\{\varphi_t\}$ (Panel A) and the aggregate R&D-to-value ratio in the model and data (Panel B). The function governing the transition of φ_t is given by (66); the horizontal dashed lines in Panel A correspond to the estimated pre-transition value φ_0 and terminal value φ_∞. See the main text and Appendix F for further details on the SMM estimation procedure.

value. In fact, some plausible parameter shifts—like a decline in discount rates ρ, which increases valuations—actually cause R&D-to-value to *rise*, due to the positive labor-supply response. Declining innovation productivity is therefore essential to explaining this moment.

Estimates of innovation productivity φ_t I estimate the time series $\{\varphi_t\}$ to fit the time series of R&D-to-value. To keep this transition parsimonious, and to smooth out higher-frequency movements in R&D-to-value that are potentially unrelated to innovation productivity, I assume that this path takes the functional form

$$\varphi_t = \varphi_0 + \frac{1}{1 + \exp\{\kappa_\varphi(t - t_{\text{mid}})\}}(\varphi_\infty - \varphi_0), \tag{66}$$

where φ_0 is the initial level of innovation productivity (estimated in the first stage), φ_∞ is the terminal value, t_{mid} is the midpoint year of the transition (i.e., when $\varphi_t = \frac{1}{2}(\varphi_0 + \varphi_\infty)$), and $\kappa_\varphi > 0$ governs the speed of the transition. The estimated parameters minimize the sum of squared errors between R&D-to-value in the model and data along the transition.

Panel A of Figure 5 plots the estimated time series of φ_t.[46] Innovation productivity is estimated to have fallen roughly in half since 1980, meaning the same quantity of research inputs in the twenty-first century is expected to yield half the number of new ideas. Panel

---

[46]This time path corresponds to the estimates $\kappa_\varphi = 0.3$, $t_{\text{mid}} = 1992.5$, and $\varphi_\infty = 0.022$.

B plots the R&D-to-value ratio in the data and model, from which this decline is estimated.

5   Main results

I now study the transition-path equilibrium of the model under the estimated decline in innovation productivity φ_t, holding all other parameters fixed.[47] The main finding is that the estimated decline in φ_t explains virtually all of the long-run decline in economic growth and the rise in aggregate value-to-sales. As in the empirical decompositions in Section 1.3, all of the rise in aggregate valuations is from a reallocation of sales to high-valuation firms.

5.1   The decline in economic growth

As one might expect, a decline in innovation productivity implies falling economic growth. Figure 6 plots the rate of per-capita output growth in the data and model. The data line shows a ten-year, two-sided moving average (i.e., ± 10 years) of the annualized growth rate of per-capita real GDP. In the model, this growth rate falls by about 1.1 percentage points, which is virtually all of the long-run decline we see in the data.

Essentially all of this decline comes from the fact that there is less innovation output, but it is worth nothing that there are two other channels coming from reallocation.[48] On the one hand, growth increases because more productive firms are producing more of the products in the economy. On the other hand, because these high-productivity firms tend to have higher markups, the aggregate markup slightly rises, suppressing labor supply and reducing output growth. The net effect of these channels is positive but negligibly small.

C. I. Jones (1995) points out that, in the very long-run, the rate of economic growth is ultimately tied to the rate of population growth, because people are the source of ideas. In other words, very-long-run economic growth is not endogenous, but "semi-endogenous." My model also features semi-endogenous growth, as seen in the fact that the rate of new-variety creation g_N will equal the rate of population growth g_L in any balanced-growth

---

[47]I assume all agents can foresee the transition starting in 1965. Were we to depart from perfect foresight, we would need to make additional assumptions about learning. This would dramatically complicate the numerical solution, because we would need the full *expected* time path of the distribution at each time step.

[48]Changes in static misallocation Ω (markup dispersion) are close to zero.

Figure 6: Economic growth in the model and data
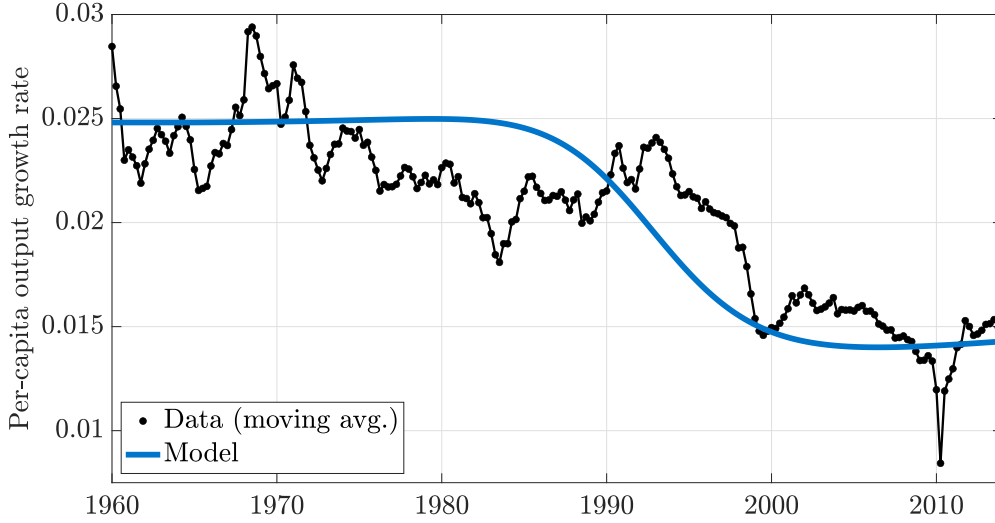
The figure shows the annualized percentage growth rate of real output per capita in the data and model. The data series is a ten-year, two-sided moving average (i.e., an equal-weighted average over a window of ± 10 years) of the annualized quarterly growth rate using real per-capita output from the BEA. The model line is the growth rate of per-capita output: $g_{yt} = \dot{y}_t / y_t$, where $y_t = Y_t / L_t$.

equilibrium (BGE). However, along the transition between BGEs, this is no longer true. When innovation productivity falls, g_{Nt} falls below g_L. This means that the number of researchers slowly begins to outgrow the number of blueprints, so the ratio L_t / N_t begins to increase. Because researchers are an input into blueprint creation, this raises the rate of innovation until g_N again equals g_L. This transition takes centuries to materialize as higher growth, and thus does not matter much over the period in Figure 6.[49] If we assume population growth will also fall in the future (as current U.N. forecasts predict), then the decline in growth will be permanent, as in Peters and Walsh (2022) and C. I. Jones (2022a).

5.2 The reallocation-driven rise in aggregate valuations

The decline in innovation productivity causes a large rise in the aggregate valuation of the stock market. Figure 7 plots the aggregate value-to-sales ratio in the model and data since 1965. The model-implied value-to-sales ratio increases 83%, essentially all of the secular increase from the late 1960s and early 1970s to the twenty-first century. Notably, the decline

---

[49]The output growth rate in the new BGE is not exactly the same as in the initial BGE, because changing the level of innovation productivity also changes the relative allocations of skilled labor between new-variety creation versus quality innovation, which have somewhat different effects on aggregate TFP.

Figure 7: Aggregate value-to-sales in the model and data
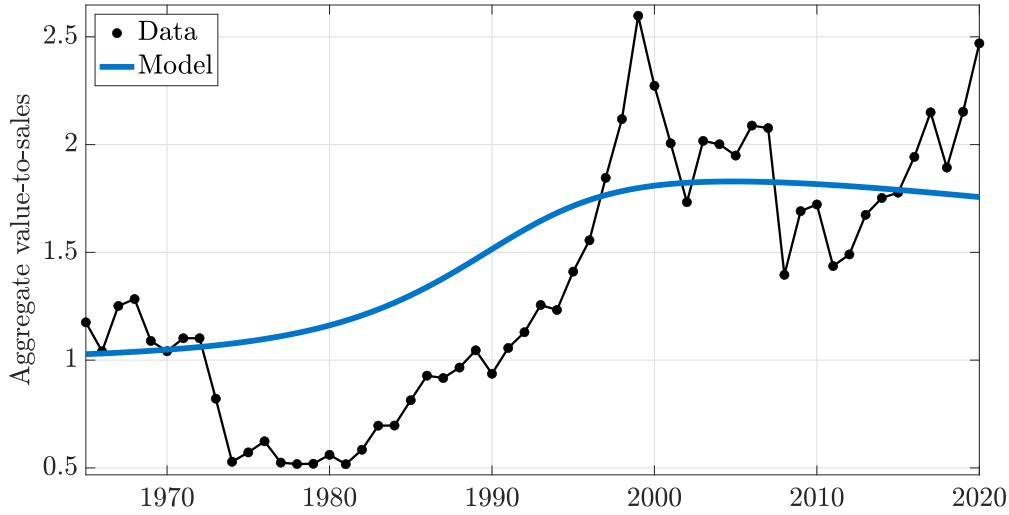
The figure shows the ratio of aggregate value to aggregate sales in the model and data. The data line is the same as the line for value-to-sales in Panel A of Figure 1.

in innovation productivity does not explain the dramatic fall and rise of valuations around 1980. This change coincided with a surge in real interest rates and unusually low profits and markups, neither of which is accounted for by this model.[50]

What explains the stock-market boom? Figure 8 plots the decomposition (2) of the aggregate value-to-sales ratio in the data and model. Recall that the within-firm change measures the cumulative effect of changing firm-level valuations, holding sales shares fixed; while the compositional change measures the effect of changing sales shares, holding fixed firm-level valuations. The data plot (Panel A) repeats Figure 3 and the model plot (Panel B) computes the decomposition on a simulated panel of firms along the model transition. The y-axes in both plots are normalized by the cumulative change in aggregate value-to-sales from 1970–2020, so that each point represents a percentage of this cumulative change.

As in the data, the model implies that all of the cumulative rise in aggregate value-to-sales came from the compositional change, a reallocation of sales shares toward high-valuation firms. Essential to this conclusion are the two main model ingredients—heterogeneous productivity and M&A—without which there would be no reallocation. I will discuss the role of M&A shortly. The within-firm change in Panel B also shows that, in the model, there is an

---

[50]In support of this claim, recall that the empirical decomposition in Figure 3 showed that this episode is explained by a common within-firm change in valuations, not a compositional change in the market.

Figure 8: Decomposition of value-to-sales in the model and data
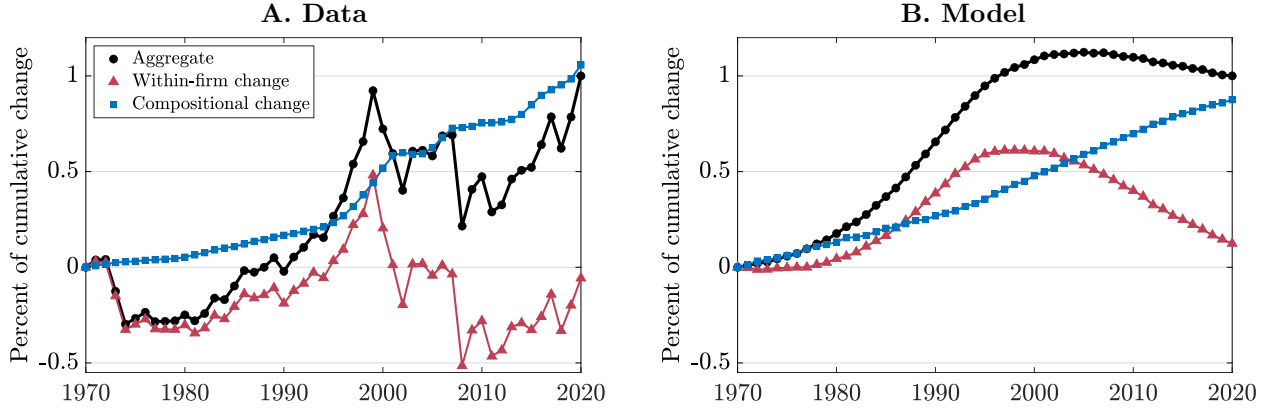
The figure plots the decompositions of changes in aggregate value-to-sales into within-firm and compositional components in the data and model. The decomposition is defined in (2)—see the main text for an explanation. In both panels, the y-axes are rescaled by the cumulative change in the aggregate from 1970–2020. Panel A is the exact same as Panel A of Figure 8 (except with a rescaled y-axis). Panel B is computed from a simulated panel of firms along the transition path in the model.

initial rise in firm-level valuations that subsequently reverts. The reason for this is perfect foresight: investors anticipate the coming rise in M&A and value, and this is priced ex ante.

What explains the stability of within-firm valuations over time? Put another way, how were high-productivity firms able to maintain high growth rates despite falling R&D? There are two main reasons. First is the shift to M&A: high-productivity firms mitigated the effect of lower R&D by reallocating labor toward the (relatively more productive) M&A technology. The second major reason is the decline in creative destruction by competitors and entrants. Because some innovations come in the form of quality improvements, and quality improvements have negative spillovers onto incumbents, a general decline in R&D reduces the likelihood that an incumbent will lose its products. This channel raises the expected growth rates of incumbents, translating into higher (firm-level) valuations relative to both sales and cashflows. This is where the estimated value of α matters for valuations: the larger is α, the larger the decline in creative destruction and firm entry.

5.3 The rise in M&A and reallocation

Figure 9 plots a key implication of the model: the rise in aggregate M&A-to-R&D. The data line is the same as the "mergers and acquisitions of assets" line from Figure 2. M&A-to-R&D rises in the model for two reasons. The first is a denominator effect, which simply follows

Figure 9: Aggregate M&A-to-R&D in the model and data
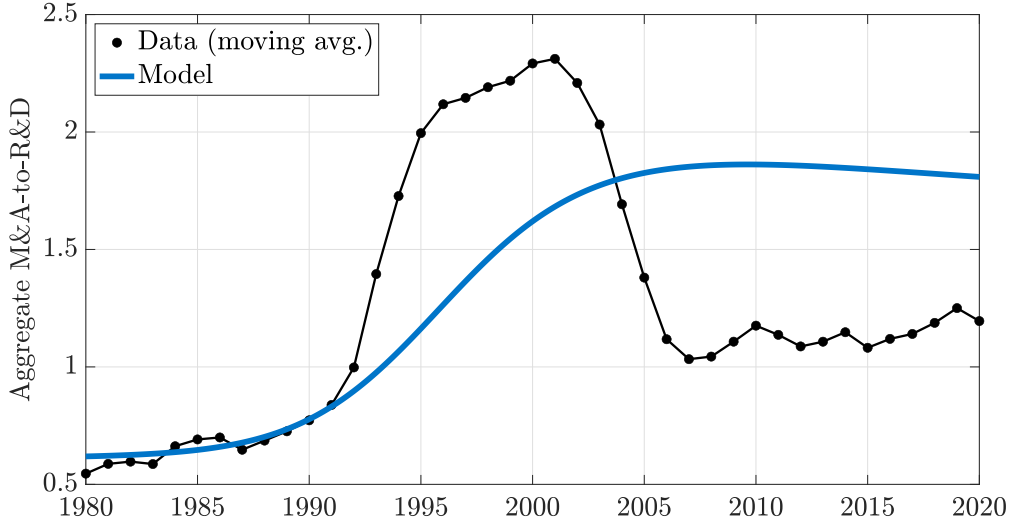
The figure shows the ratio of aggregate M&A spending to aggregate R&D spending in the model and data. M&A spending is the total amount spent by acquirers on targets in completed M&A transactions. The data line is the same as the line for mergers and acquisition of assets in Figure 2.

from the fact that all firms do less R&D. The second is a numerator effect, which follows from the fact that high-productivity firms do more M&A. When R&D gets harder, M&A becomes a relatively more attractive means of growth, so these firms reallocate labor toward M&A. The consequence of rising M&A-to-R&D in the model is an increase in the total market share (share of sales, profits, etc.) of high-productivity firms. Specifically, the high-productivity sales share rises from about 15% to about 30% over this period. This explains the quantitatively large reallocation of sales toward high-valuation firms in Figure 8.

The timing of the model-implied M&A-to-R&D rise in Figure 9 is off: it misses the M&A wave of the late 1990s and overshoots thereafter. However, the *cumulative* amount of M&A activity over the full period from 1980 to 2020 is actually very close in the model and data. Cumulative M&A activity is what ultimately matters when we consider the total amount of reallocation that has taken place across firms in the stock market.

5.4   Implications for other series

We can also use the model to study secular change in other micro and macro series. Here I discuss two implications that have been of interest to the literature. The first is the decline in new-firm entry and the increase in concentration. Figure 10 plots the rate of new-firm entry

Figure 10: New-firm entry rate in the model and data
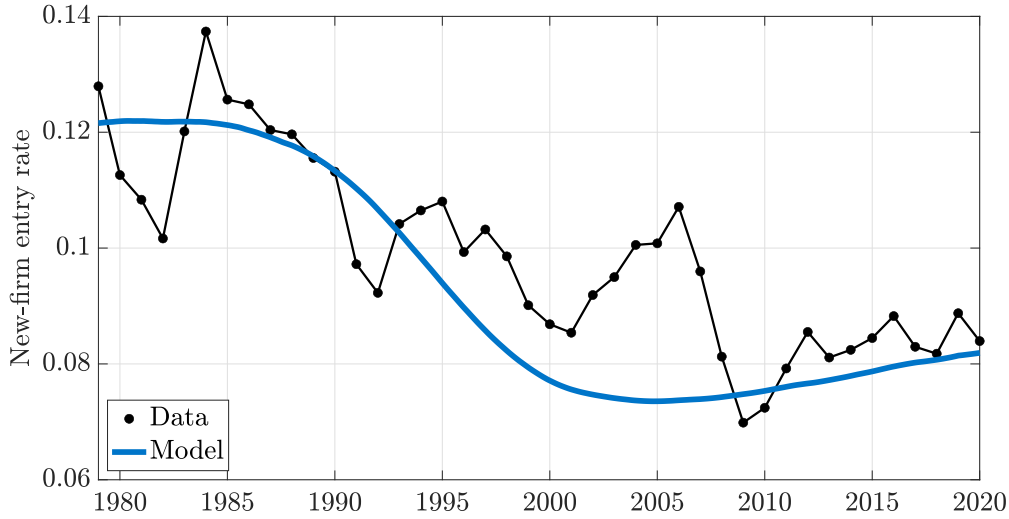
The figure shows the rate of new-firm entry in the model and data, defined as the number of new firms entering the economy per existing firm. The data series is computed from the Business Dynamics Statistics.

in the data (from BDS) and the model, defined as the the number of new firms entering the economy per existing firm. The decline in research productivity explains all of the decline in entry because entrepreneurs have a harder time coming up with new ideas for firms.[51]

As just discussed above, the decline in firm entry means that incumbents face less risk of creative destruction and become entrenched. It also implies a rise in the concentration of blueprints, both because incumbents are able to grow without risk of displacement and because there is less entry of small firms into the economy. A second force that increases concentration is the rise in M&A, which reduces the number of small, low-productivity firms by reallocating their blueprints to larger, high-productivity firms. The increased concentration of blueprints translates into increased concentration of employment, sales, and profits.

A third secular trend that this mechanism can help explain is the long-run secular decline in real interest rates. Figure 11 plots the change in the real interest rate since 1960.[52] I show the change, instead of the level, because the model-implied interest rate does not match

---

[51]In reality, the entry rate is a weighted average of innovative firms (this model) and non-innovative firms (Hurst & Pugsley, 2011). While the model cannot speak directly to the latter, one possibility is that fewer new ideas imply less demand for new non-innovative firms whose services are based around these ideas.

[52]In the data, I infer the real rate from monthly realized returns on short-term bonds, defined as the Fed Funds rate deflated by realized one-month-ahead CPI inflation. In the spirit of Beeler and Campbell (2012), I run time-series regressions of next month's realized return on the current nominal Fed Funds rate, last month's inflation, and last year's inflation. The expected return is then the predicted value from the regression. I use this approach because survey expectations (e.g., the SPF) do not go back far enough.

Figure 11: Change in real interest rates in the model and data
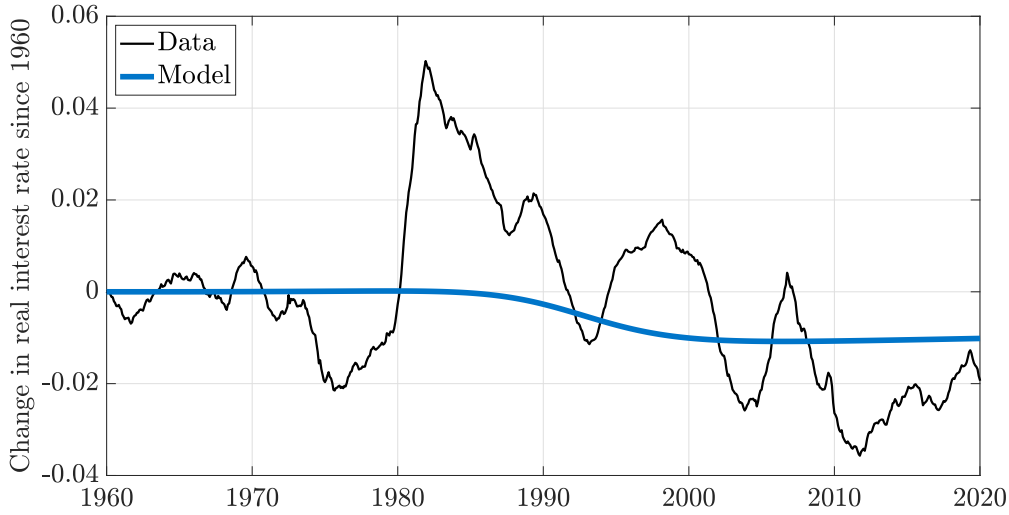
The figure shows the change in real interest rates since 1960 in the data and model. The data series is the annualized expected real return on a nominal bond, where the realized return is measured as the Fed Funds rate deflated by realized CPI inflation. I infer the series of expected real returns from this series using a procedure similar to that of Beeler and Campbell (2012), described in the main text. Note that both the data and model series are shifted to start at zero in 1960, so that this is the change in real rates since 1960.

the data level, as discount rates ρ were estimated from firm valuations, not interest rates— Section 7 discusses how to reconcile this. In the model, the interest rate falls by the same amount as per-capita output growth: about one percentage point. This is about half of the decline since 1960. Clearly, the mechanism does not explain the spike in real rates around 1980, in line with the discussion above about where the model also misses firm valuations.

6 Implications for welfare

What are the welfare implications of the divergence of output growth and stock-market wealth over the last half-century? It might seem obvious that declining output growth should mean welfare stagnates. However, what matters for total welfare is not just the level of consumption, but its *present value*. Thus, the fact that wealth boomed over this period could suggest that other forces have offset this negative growth effect.[53] In this section, I show that welfare did not boom with stock-market wealth. Rather, welfare stagnated with

---

[53]For example, if the valuation boom came from a decline in aggregate risk or increased savings demand from abroad (which would raise the wealth of U.S. investors relative to the rest of the world), then agents might still be better off from a welfare perspective.

output, because all of the gains in the stock market came from a reallocation of output from workers and entrants toward incumbent firms.

To state this formally, note that the value function of the representative household equals

$$U_t = k(t) + \rho^{-1} \log \left(\frac{\overline{W}_t}{L_t} \right), \tag{67}$$

where $k(t)$ is a deterministic function of time and $\overline{W}_t / L_t$ is *total wealth* per person.[54] Total wealth is the present value of per-capita consumption c_t:

$$\frac{\overline{W}_t}{L_t} = \mathbb{E}_t \left[\int_0^\infty \frac{\xi_{t+\tau}}{\xi_t} c_{t+\tau} d\tau \right] = \rho^{-1} c_t. \tag{68}$$

Because consumption equals output in equilibrium, we therefore have

$$\overline{W}_t = \rho^{-1} Y_t. \tag{69}$$

Total wealth is always exactly proportional to output, even along the transition path. Moreover, because output is the sum of all dividends and labor income,

$$Y_t = D_t + w_{\text{p}t} L_{\text{p}t} + w_{\text{x}t} X_t, \tag{70}$$

we can decompose total wealth into three components:

$$\overline{W}_t = V_t + \text{PV}_t^{\text{entrants}} + \text{PV}_t^{\text{labor}}. \tag{71}$$

That is, total wealth includes not only the value V_t of the stock market (the claim to current incumbent firms' dividends), but also the present value $\text{PV}_t^{\text{entrants}}$ of future entrants (firms not yet traded) and the present value $\text{PV}_t^{\text{labor}}$ of labor income paid out to households.

This means that, while the value of the stock market boomed relative to output (see Figure 7), the present value of consumption (69) stagnated with output. Figure 12 plots the cumulative change in the log of each of these wealth measures (V_t and $\overline{W}_t$) since 1975. As output growth fell, total wealth stagnated; at the same time, the value of the stock market

---

[54]$k(t)$ includes the effect of current and future labor disutility on welfare, which I do not consider here.

Figure 12: Growth in stock-market wealth and total wealth
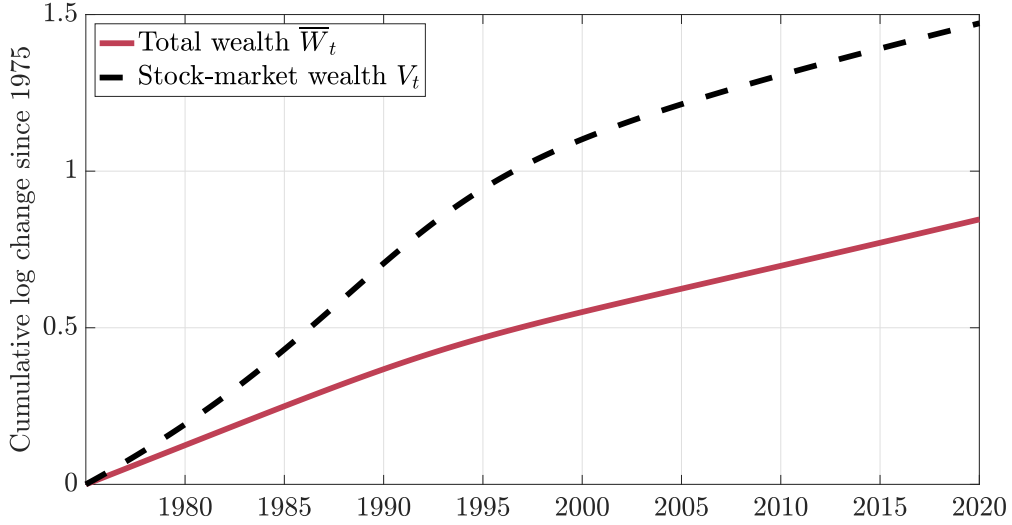
The figure shows cumulative change in log stock-market wealth ($\log V_t$) and log total wealth ($\log \overline{W}_t$) from 1975–2020 in the model. Total wealth is defined as the present value of consumption; see the main text for a definition.

accelerated. The decomposition (71) accounts for this: all of the gains in stock-market wealth came at the expense of labor income and future entrants. The present value of labor income falls mainly because high-productivity firms have higher markups. The present value of entrants falls both because the rate of new-firm entry falls (see Figure 10) and because the average value of incumbent firms (which tend increasingly to be high-productivity) rises relative to the average value of entrants (which tend to be low-productivity).

7   Discussion of model extensions

Here I consider several model extensions and alternative assumptions. Some of the extensions yield new insights about the economic transition. Appendix G gives formal derivations.

Systematic risk and risk aversion   Perhaps the most unrealistic assumption about the baseline model is that there is no systematic risk, so firms are discounted at the riskless interest rate. We can generalize this along two dimension, following Miller et al. (2022) and Farhi and Gourio (2018). First, let us generalize firms' production functions (10) to

$$Y_{ijt} = a_i e^{z_t} L_{ijt}, \tag{72}$$

where z_t is a common, *aggregate* productivity shock, which follows the jump-diffusion

$$dz_t = \kappa_z(\bar{z} - z_t)dt + \sigma_z dB_t - \zeta_z dJ_t, \tag{73}$$

where B_t is a Brownian motion and J_t is a Poisson process with intensity p_z. For small p_z and $\zeta_z > 0$, this jump term represents a rare disaster, and makes it possible to match a sizable equity premium.[55] Second, let us generalize household preferences (5) to the recursive formulation of Duffie and Epstein (1992):

$$u(c_t, l_t, U_t) = \rho(1 - \gamma)U_t \left(\log c_t - \chi \frac{l_t^{1+1/\zeta}}{1 + 1/\zeta} - \frac{\log((1 - \gamma)U_t)}{1 - \gamma} \right). \tag{74}$$

This preference specification continues to assume a unit EIS and a constant Frisch elasticity, but allows for potentially higher risk aversion $\gamma \geq 1$.[56]

Appendix G.1 re-solves the model under these assumptions and shows that the equilibrium conditions are *exactly the same* as in the case without systematic risk, except now the discount factor ρ in the HJBs (48) and (49) is replaced by a *risk-adjusted* discount factor

$$\rho^* \equiv \rho - (\gamma - 1)\frac{\kappa_z}{\kappa_z + \rho}\sigma_z - p_z[e^{-\zeta_z}(1 - e^{-(\gamma-1)\frac{\kappa_z}{\kappa_z+\rho}\zeta_z})(1 - e^{-\zeta_z})]. \tag{75}$$

The risk adjustment incorporates both the equity risk premium, which raises ρ^*, and the precautionary savings motive, which lowers ρ^*. Because the model estimation identified the discount factor using a valuation ratio (not the riskfree rate), the estimate is still appropriate in this setting, but should be reinterpreted as an estimate of ρ^*, not of ρ.

Physical capital Another plausible generalization of the production function (10) is to assume that it takes both labor and physical capital K_t as inputs:

$$Y_{ijt} = a_i K_{ijt}^{\beta} L_{ijt}^{1-\beta}. \tag{76}$$

---

[55]$\kappa_z > 0$ implies that disasters are followed by recoveries (Nakamura, Steinsson, Barro, & Ursúa, 2013).

[56]As Appendix G.1 shows, keeping an EIS of one is essential for tractability with aggregate shocks. If the EIS is not one, then the blueprint distribution is no longer stationary, but depends on the history of z_t.

For simplicity, suppose that the capital stock depreciates at constant rate δ_K and is produced by a competitive capital-production sector that makes total investment I_t and rents capital out to goods producers at an endogenous rental rate r_{Kt}.[57] This implies total capital accumulation

$$\dot{K}_t = I_t - \delta_K K_t. \tag{77}$$

Appendix G.2 derives two results. First, the equilibrium conditions are all the same, except the "labor share" Λ from (28) is reinterpreted as the total factor share, consisting of the labor share $(1 - \beta)\Lambda$ and the capital share $\beta\Lambda$. Because the estimation moments depend only on the profit share $(1 - \Lambda)$, the estimates remain valid. The second implication is that the equilibrium investment-capital ratio is given by

$$\frac{I_t}{K_t} = \delta_K + g_L + g_{yt} + g_{\Lambda t}. \tag{78}$$

Hence, the decline in per-capita output growth g_{yt} and the slight rise in the aggregate markup Λ^{-1} in the model imply a decline in investment-capital. This is a well-documented trend in the data and the subject of several studies (e.g., Farhi & Gourio, 2018; Miller et al., 2022; Crouzet & Eberly, 2023). In the model, investment stagnates because "average Q" booms but the "marginal q" of capital investment remains constant. Intuitively, capital exists to scale up production of new ideas, so fewer new ideas means less need for capital.

Cournot competition The results also remain the same if we assume that firms compete à la Cournot (quantity competition) instead of à la Bertrand (price competition). The only difference in the Cournot model is that firms no longer engage in limit pricing. Consequently, there can arise product markets in which multiple firms choose to produce.

Life-cycle dynamics As Luttmer (2011) explains, assuming that firms remain high-productivity indefinitely implies that the average age of large firms is too high. Akcigit and Kerr (2018) also show that younger firms tend to have higher R&D intensities and growth rates. Luttmer (2011) and Acemoglu et al. (2018) address these facts by assuming

---

[57]Adding capital adjustment costs or assuming goods producers choose investment are straightforward.

that high-productivity firms switch permanently to low-productivity firms at some point in the life cycle, according to an exogenous and constant switching intensity.

Ultimately, allowing for type switching does not change the results above. This is because the original estimation moments do not rely on within-firm life-cycle dynamics (age) to identify parameters. Adding type-switching shocks simply shifts parameters: we would have a higher share of high-productivity entrants and a larger productivity gap, but the cross-sectional firm distribution would be the same (as long as we are not conditioning on age).

Own-quality innovations   A large class of Schumpeterian growth models assumes that, within product markets, firms compete for market share by improving the quality of their existing blueprints. In principle, this could occur in the baseline model if a quality innovation hits its own blueprint; however, this is a probability-zero event. Peters (2020) and Peters and Walsh (2022) propose a tractable way of incorporating own-quality innovations into this framework. As long as the own-quality innovation technology also becomes less productive as new ideas get harder to find, the main results will remain the same.

8   Concluding remarks

How do we reconcile stagnating economic growth with a booming stock market? This paper presents evidence that declining innovation productivity explains both trends. As research gets harder, firms do less R&D and growth stagnates. At the same time, M&A activity rises, which reallocates market share toward high-productivity incumbents. High-productivity firms have high present values, so the aggregate market value booms relative to output.

The mechanism is quantitatively important. From the decline in aggregate R&D-to-value, I estimate that innovation productivity fell in half since 1975. The decline in innovation productivity explains most of the secular decline in growth and the rise in aggregate valuations. It also implies, as I show in the data, that all of the rise in aggregate valuation ratios is explained by a reallocation of market share toward high-valuation firms. The mechanism helps explain other salient trends, including declining firm entry and falling interest rates.

While aggregate stock-market wealth boomed, the present value of consumption—which

is what matters for consumer welfare—did not. Indeed, the present value of consumption stagnated with output. All of the reallocation-driven gains to incumbent firms came at the expense of future labor income and entrants, implying a disconnect between the stock market and the rest of the economy. This finding highlights the importance of studying both growth and the stock-market at a disaggregated level.

References

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

Achdou, Y., Han, J., Lasry, J., Lions, P., & Moll, B. (2022). Income and wealth distribution in macroeconomics: A continuous-time approach. *Review of Economic Studies*, *89*(1), 45–86.

Aghion, P., Bergeaud, A., Boppart, T., Klenow, P. J., & Li, H. (2023). A theory of falling growth and rising rents. *The Review of Economic Studies*, *90*(6), 2675–2702.

Aghion, P., & Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, *60*(2), 323–351.

Akcigit, U., & Ates, S. T. (2023). What happened to us business dynamism? *Journal of Political Economy*, *131*(8), 2059–2124.

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

Arikan, A. M., & Stulz, R. M. (2016). Corporate acquisitions, diversification, and the firm's life cycle. *The Journal of Finance*, *71*(1), 139–193.

Atkeson, A., Heathcote, J., & Perri, F. (2024, May). *The end of privilege: A reexamination of the net foreign asset position of the united states.* (Working paper, UCLA and Federal Reserve Bank of Minneapolis)

Autor, D., Dorn, D., Katz, L. F., Patterson, C., & Van Reenen, J. (2020). The fall of the labor share and the rise of superstar firms. *The Quarterly Journal of Economics*, *135*(2), 645–709.

Baily, M. N., Hulten, C., & Campbell, D. (1992). Productivity dynamics in manufacturing plants. *Brookings Papers on Economic Activity (Microeconomics)*, 187–249.

Bartelsman, E. J., & Doms, M. (2000). Understanding productivity: Lessons from longitudinal microdata. *Journal of Economic Literature*, *38*(3), 569–594.

Beeler, J., & Campbell, J. Y. (2012). The long-run risks model and aggregate asset prices: An empirical assessment. *Critical Finance Review*, *1*, 141–182.

Betton, S., Eckbo, B. E., & Thorburn, K. S. (2008). Corporate takeovers. In B. E. Eckbo (Ed.), *Handbook of empirical corporate finance* (pp. 291–429). Elsevier.

Binsbergen, J. v. (2021, November). *Duration-based stock valuation: Reassessing stock market performance and volatility.* (NBER Working Paper 27367)

Bloom, N., Griffith, R., & Van Reenen, J. (2002). Do R&D tax credits work? evidence from a panel of countries 1979–1997. *Journal of Public Economics*, *85*(1), 1–31.

Bloom, N., Jones, C. I., Van Reenen, J., & Webb, M. (2020). Are ideas getting harder to find? *American Economic Review*, *110*(4), 1104–1144.

Blundell, R., Griffith, R., & Windmeijer, F. (2002). Individual effects and dynamics in count data models. *Journal of Econometrics*, *108*(1), 113–131.

Broda, C., & Weinstein, D. E. (2006). Globalization and the gains from variety. *The Quarterly Journal of Economics*, *121*(2), 541–585.

Catherine, S., Ebrahimian, M., Sraer, D., & Thesmar, D. (2023, January). *Robustness checks in structural analysis.* (Working Paper, University of Pennsylvania, Stockholm School of Economics, UC Berkeley, and MIT)

Chetty, R. (2012). Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply. *Econometrica*, *80*(3), 969–1018.

Cho, T., Grotteria, M., Kremens, L., & Kung, H. (2024, June). *The present value of future market power.* (Working paper, Korea University, London Business School, and University of Washington)

Cocci, M. D., & Plagborg-Møller, M. (in press). Standard errors for calibrated parameters. *Review of Economic Studies.*

Cochrane, J. H., Longstaff, F. A., & Santa-Clara, P. (2008). Two trees. *The Review of Financial Studies*, *21*(1), 347–385.

Corhay, A., Kung, H., & Schmid, L. (2020). Competition, markups, and predictable returns. *The Review of Financial Studies*, *33*(12), 5906–5939.

Crouzet, N., & Eberly, J. (2023). Rents and intangible capital: A q+ framework. *The Journal of Finance*, *78*(4), 1873–1916.

De Loecker, J., Eeckhout, J., & Unger, G. (2020). The rise of market power and the macroeconomic implications. *The Quarterly Journal of Economics*, *135*(2), 561–644.

De Ridder, M. (2024). Market power and innovation in the intangible economy. *American Economic Review*, *114*(1), 199–251.

Doidge, C., Karolyi, G. A., & Stulz, R. M. (2017). The u.s. listing gap. *Journal of Financial Economics*, *123*(3), 464–487.

Duffie, D., & Epstein, L. G. (1992). Asset pricing with stochastic differential utility. *The Review of Financial Studies*, *5*(3), 411–436.

Eckbo, B. E., & Lithell, M. (2023). Merger-driven listing dynamics. *Journal of Financial and Quantitative Analysis*, 1–49.

Edmond, C., Midrigan, V., & Xu, D. Y. (2023). How costly are markups? *Journal of Political Economy*, *131*(7), 1619–1675.

Eggertsson, G. B., Robbins, J. A., & Wold, E. G. (2021). Kaldor and piketty's facts: The rise of monopoly power in the united states. *Journal of Monetary Economics*, *124*, S19–S38.

Ewens, M., Peters, R. H., & Wang, S. (in press). Measuring intangible capital with market prices. *Management Science*.

Farhi, E., & Gourio, F. (2018). Accounting for macro-finance trends: Market power, intangibles, and risk premia. *Brookings Papers on Economic Activity*, 147–250.

Foster, L., Haltiwanger, J. C., & Krizan, C. J. (2001). Aggregate productivity growth: Lessons from microeconomic evidence. In *New developments in productivity analysis*. University of Chicago Press.

Gârleanu, N., Kogan, L., & Panageas, S. (2012). Displacement risk and asset returns. *Journal of Financial Economics*, *105*(3), 491–510.

Gordon, R. J. (2016). *The rise and fall of american growth: The us standard of living since the civil war*. Princeton University Press.

Greenwald, D. L., Lettau, M., & Ludvigson, S. C. (in press). How the wealth was won: Factor shares as market fundamentals. *Journal of Political Economy*.

Griliches, Z., & Regev, H. (1995). Firm productivity in israeli industry 1979–1988. *Journal of Econometrics*, *65*, 175–203.

Grossman, G. M., & Helpman, E. (1991). Quality ladders in the theory of growth. *Review of Economic Studies*, *58*(1), 43–61.

Hall, B. H. (1993). R&D tax policy during the 1980s: Success or failure? *Tax Policy and the Economy*, *7*, 1–35.

Hall, B. H., & Ziedonis, R. H. (2001). The patent paradox revisited: An empirical study of patenting in the u.s. semiconductor industry, 1979-1995. *The RAND Journal of Economics*, *32*(1), 101–128.

Haltiwanger, J. C. (1997). Measuring and analyzing aggregate fluctuations: The importance of building from microeconomic evidence. *Federal Reserve Bank of St. Louis Review*, *79*(3), 55–77.

Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica*, *60*(5), 1127–1150.

Hurst, E., & Pugsley, B. W. (2011). What do small businesses do? *Brookings Papers on Economic Activity*(2), 73–118.

Jones, B. F. (2009). The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *The Review of Economic Studies*, *76*(1), 283–317.

Jones, B. F. (2010). Age and great invention. *The Review of Economics and Statistics*, *92*(1), 1–14.

Jones, C. I. (1995). R&D-based models of economic growth. *Journal of Political Economy*, *103*(4), 759–784.

Jones, C. I. (2022a). The end of economic growth? unintended consequences of a declining population. *American Economic Review*, *112*(11), 3489–3527.

Jones, C. I. (2022b). Taxing top incomes in a world of ideas. *Journal of Political Economy*, *130*(9), 2227–2274.

Judd, K. L. (1998). *Numerical methods in economics*. MIT Press.

Keane, M. P. (2011). Labor supply and taxes: A survey. *Journal of Economic Literature*,

$49(4)$, 961–1075.

Kehrig, M., & Vincent, N. (2021). The micro-level anatomy of the labor share decline. *The Quarterly Journal of Economics*, *136*(2), 1031–1108.

Kendall, D. G. (1948). On the generalized "birth-and-death" process. *The Annals of Mathematical Statistics*, *19*(1), 1–15.

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

Kogan, L., Papanikolaou, D., Seru, A., & Stoffman, N. (2017). Technological innovation, resource allocation, and growth. *The Quarterly Journal of Economics*, *132*(2), 665–712.

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

Liu, E., Mian, A., & Sufi, A. (2022). Low interest rates, market power, and productivity growth. *Econometrica*, *90*(1), 193-221.

Loualiche, E. (in press). Asset pricing with entry and imperfect competition. *The Journal of Finance*.

Luttmer, E. G. (2011). On the mechanics of firm growth. *The Review of Economic Studies*, *78*(3), 1042–1068.

Martin, I. (2013). The lucas orchard. *Econometrica*, *81*(1), 55–111.

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics*, *51*(3), 247–257.

Miller, M., Paron, J. D., & Wachter, J. A. (2022, August). *Sovereign default and the decline in interest rates.* (Jacobs Levy Equity Management Center for Quantitative Financial Research Paper)

Nakamura, E., Steinsson, J., Barro, R., & Ursúa, J. (2013). Crises and recoveries in an empirical model of consumption disasters. *American Economic Journal: Macroeconomics*, *5*(3), 35–74.

Olmstead-Rumsey, J. (2022, September). *Market concentration and the productivity slowdown.* (Working Paper, London School of Economics)

Paron, J. D. (2022, March). *Heterogeneous-agent asset pricing: Timing and pricing id-*

iosyncratic risks. (Jacobs Levy Equity Management Center for Quantitative Financial Research Paper)

Peters, M. (2020). Heterogeneous markups, growth, and endogenous misallocation. *Econometrica*, *88*(5), 2037–2073.

Peters, M., & Walsh, C. (2022, July). *Population growth and firm-product dynamics.* (Working Paper, Yale University and Columbia University)

Phillips, G. M., & Zhdanov, A. (2013). R&d and the incentives from merger and acquisition activity. *The Review of Financial Studies*, *26*(1), 34–78.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, *98*(5), S71–S102.

Wilson, D. J. (2009). Beggar thy neighbor? the in-state, out-of-state, and aggregate effects of R&D tax credits. *The Review of Economics and Statistics*, *91*(2), 431–436.

Online Appendix for

The wealth of stagnation:
Falling growth, rising valuations[*]

James D. Paron[†]

January 4, 2025

Contents

A   Data appendix

This section details the main data sources, selection, and variable definitions used throughout the paper. For details about computation of data moments for the model estimation in Section 4, see Appendix F below.

A.1   CRSP and Compustat

The main data source is the annual CRSP/Compustat Merged Database from 1965 to 2020. I consider firms incorporated in the U.S. and exclude firms in utilities (SIC 2-digit code 49) and public administration (SIC 1-digit code 9), as is standard in the literature. I exclude firm-year observations with nonpositive revenues (`sale`) or market value of equity (defined below). Note that firms have only been required to report R&D expenditures (`xrd`) since 1975, so I do not use observations of this variable in any year prior to 1975.

A firm's equity value is defined as its total market capitalization at the fiscal year end: price per share (`prcc_f`) times number of shares outstanding (`csho`). The main measure of total firm value is then equity value plus book debt (`dlc` + `dltt`), less cash (`che`). Likewise, total dividends to equityholders equal dividends per share times shares (`dvpsx_f` × `csho`), and total firm cashflows are the sum of dividends and interest payments (`xint`). In robustness checks in Appendix B.3 below, I also consider a broader definition of cashflows, which includes net repurchases of common and preferred shares (`prstkc` − `prstkpc` − `scstkc`).

A.2   SDC Platinum

M&A transaction data come from the SDC Platinum database, which contains transaction-level data since 1976. I consider completed deals (`STATUS = "Completed"`) for which the acquirer is based in the U.S. (`ANATIONCODE = "US"`). Mergers (`FORM = "Merger"`) are transactions in which the acquirer purchases 100% of the target firm's stock, or businesses combine; acquisitions of assets (`FORM = "Acq. of Assets"`) are transactions in which the acquirer purchases some of the target's assets. The value of a transaction (i.e., the amount spent by the acquirer) is given by the variable `RANKVAL`, which includes the value of net debt. Because the total M&A value reported in SDC is extremely low in the first two years of data

A.3

(1976–1977), I exclude these years from analyses.

Public and private acquirers are identified by the variable `APUBLIC` $\in$ $\{$`"Public"`,`"Priv."`$\}$, and targets are likewise identified by `TPUBLIC`. To merge SDC into CRSP/Compustat at the acquirer level, I use the acquirer identifier mappings (`agvkey` in SDC matched to `gvkey`) provided by Ewens, Peters, and Wang (in press), who build on the work of Phillips and Zhdanov (2013). The target's market value before a merger announcement, which I use to compute the M&A premium, is given by the variable `MV`.

A.3   Other data sources

For analysis of trends in patent value, I use the estimates from Kogan et al. (2017), updated to 2020. In analyses of public versus private firm R&D intensity, I use the NSF R&D-to-GDP ratio for R&D performed by businesses.[1] The specific series comes from Table 1 of the National Patterns of R&D Resources (annual series) from the National Center for Science and Engineering Statistics. Data on firm entry and the employment distribution come from the Business Dynamics Statistics (BDS). Finally, per-capita real output and population growth data come from the BEA.

B   Empirical evidence: Details and additional results

This section details the data and empirical methods used in Section 1, and presents supplementary results.

B.1   Additional evidence on aggregate R&D and M&A

B.1.1   Trends in patent values

This paper argues that the ratio of R&D spending to value is an ideal moment for identifying innovation productivity. An alternative measure of the expected private value of innovation is the market value of patents. Kogan et al. (2017) estimate such a measure on the dates

---

[1]Note that this measure is based on who performs the R&D, so it includes R&D funded by the federal government if the research is ultimately performed by businesses. The share of R&D performed by the federal government since the 1950s is both small and stable relative to GDP.

at which patents are issued. At the aggregate level, one would expect this measure to also reflect innovation productivity, because, all else equal, declining innovation productivity will reduce the total dollar flow of patents produced.

Figure B.1 plots the ratio of aggregate patent value from Kogan et al. (2017) (the sum of individual patent values each year) to the aggregate value of firms over time. Panel A reports this ratio (in logs) on the dates at which the patents were issued; Panel B reports the ratio on the dates they were filed. The series is much noisier than R&D-to-value, but clearly shows a downward trend over time. Formally, the estimated time trend from a linear regression of the log ratio on the year implies a significant decrease (t-statistics of -3.81 for issue date and -3.39 for filing date). More importantly, the estimated *magnitude* of the decline in patent value implied by these regressions is nearly identical to the decline in the aggregate R&D-to-value ratio over the same period: both fell by about 0.5 log points.

Figure B.1: Trends in patent values
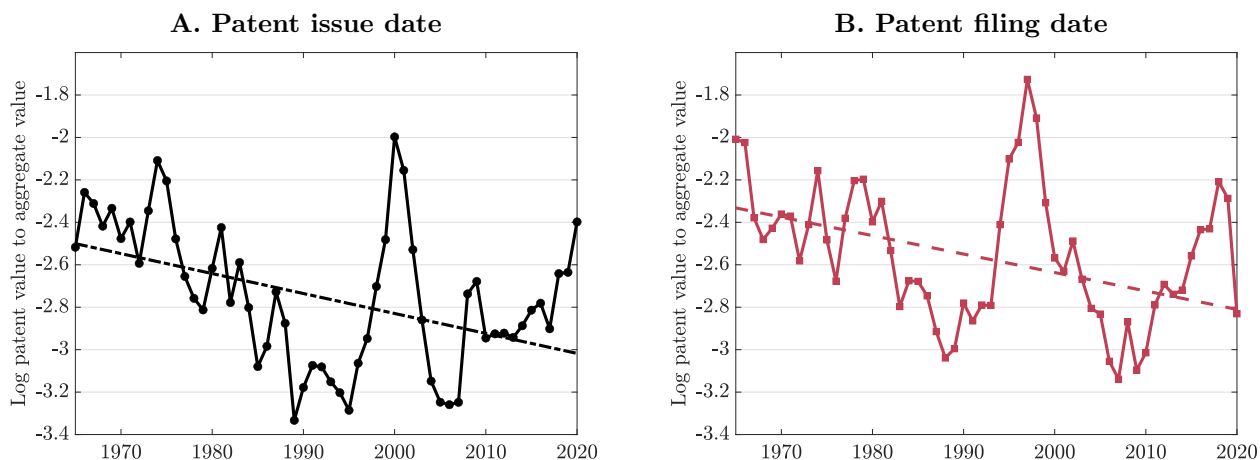
The figure plots the log of the ratio of aggregate patent value from Kogan et al. (2017) to aggregate firm value in Compustat over time. Panel A aggregates patent values based on the year in which the patent was issued, while Panel B aggregates based on the year in which the patent was filed. The lines in each panel are the fitted lines from a regression of each log series on the year.

This result is reassuring, but also raises the question: what is the advantage of using R&D-to-value instead of a patent-value measure to study innovation productivity? The main reason is *timing*.[2] R&D and value are both forward-looking and contemporaneous. Valuations are higher today if the present value of future profits is higher. R&D spending is

---

[2] A second reason is coverage: the set of firms that report R&D is larger than the set of firms that patent. Additionally, within firms, not all new ideas that are generated by R&D necessarily result in patents.

higher today if research is expected to yield a stream of future profits with a higher present value. In contrast, patent valuations are partly backward-looking, in that they reflect a combination of past research efforts and unexpected or unrelated noise. The average time between patent filing and patent issuance in the Kogan et al. (2017) data is 2.8 years (10th percentile 1.2 years, 90th percentile 4.8 years), and the time between R&D spending and patent filing is plausibly even longer. Thus, market conditions at the time at which a patent is issued are plausibly much different from the conditions at the time at which the research was initially conducted. This will lead to a disconnect between expected innovation value and the patent value. R&D spending, on the other hand, is entirely based on the firm's expectations at the time at which research is conducted. This is likely the reason that the series in Figure B.1 is significantly more volatile than the R&D-to-value series in Figure 1. Consequently, R&D-to-value allows for higher-frequency identification of the decline in innovation productivity, which is key to the model estimation.

B.1.2 Aggregate M&A-to-R&D

Figure 2 plotted a two-sided, five-year moving average of the aggregate M&A-to-R&D ratio over time. Figure B.2 plots the underlying, unfiltered time series (in logs). The figure also plots a time trend implied by fitting a linear regression of log M&A-to-R&D on the year. The coefficient estimates are significant (*t*-statistics of 3.34 and 3.36 for mergers and mergers plus acquisitions of assets, respectively), formally confirming the visible increase over time.

Figure B.2: Trends in M&A relative to R&D: Unfiltered series
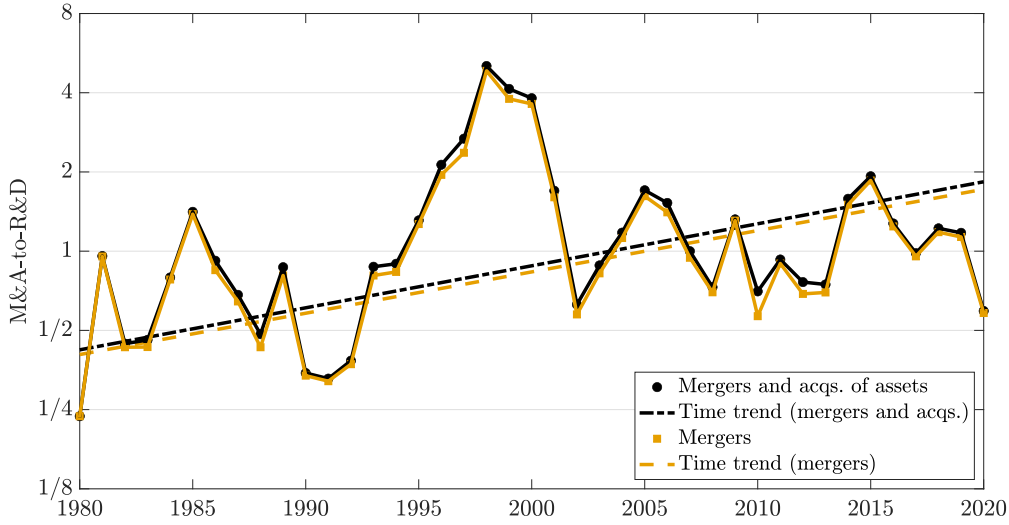
The figure plots the aggregate ratio of M&A spending to R&D spending for public firms in Compustat and SDC Platinum from 1980–2020. Note that the y-axis is transformed onto a log scale. M&A is defined as the dollar amount spent by acquirers on either mergers (transactions in which all of the target is acquired) or mergers plus acquisitions of assets. The lines denoted "Time trend" are the fitted lines from a regression of each log series on the year.

B.1.3 Evidence of value-creation by M&A

The model makes two main predictions about M&A, which link the trends in Figures 2 and 3. The first prediction is that M&A reallocates resources (here, ideas) from less productive to more productive firms. In the corporate finance literature, this has been called the "neoclassical" view, and stands in contrast with agency or behavioral views, which posit that M&A arises from frictions like information asymmetry (mispricing) or overconfidence. The second prediction, which follows from the first, is that M&A creates surplus value for the aggregate market. Here I briefly survey the empirical literature on M&A and discuss the evidence for and against these predictions.

The literature broadly finds that most mergers are consistent with the neoclassical view. Arikan and Stulz (2016) document that acquirers tend to have significantly higher market valuations (measured as Tobin's Q, their proxy for growth opportunities) than targets, which is the key cross-sectional prediction linking the rise in M&A (Figure 2) with the reallocation-driven stock-market boom (Figure 3). Moreover, in support of the claim that M&A creates surplus for the combined firm, Betton et al. (2008) show that average cumulative abnormal

A.7

returns (CARs) for the combined acquirer and target are large (1.8%) and significantly positive over both the run-up and announcement.[3] The fact that joint CARs are large and significant is supportive, but one should note that this is a limited measure of the value created through M&A. The model predicts that most of anticipated gains from M&A should be priced into both acquirers and targets before any news of a transaction occurs; thus, announcement returns capture only the (potentially small) part of the wealth gain that is not already priced.

From a macro perspective, what matters for measuring total surplus is the value created for both the acquirer and the target. Much of the literature is instead concerned with the value created for owners of the acquiring firm. These studies tend to find that most, if not all, of the positive joint returns accrue to owners of the target firm. This is consistent both with the fact that acquirers tend to pay extremely large premia relative to the target's pre-announcement value and with the fact that announcement CARs are substantially larger for targets than for bidders. Betton et al. (2008) and Arikan and Stulz (2016) find that acquirer CARs tend not to be significantly positive, and may even be negative conditional on some characteristics (e.g., all-stock acquisitions of public targets by old acquirers).[4] Note, however, that an insignificant abnormal return for an acquirer does not necessarily mean the deal did not create net value. First, it could simply mean that most of the surplus went to the owners of the target firm. This could be explained either by differential bargaining power or by overpayment. And second, as alluded to above, near-zero announcement returns for large acquirers are actually a prediction of the model in this paper, as M&A becomes increasingly predictable for large firms and hence becomes priced before announcements.[5]

---

[3]Specifically, they consider a value-weighted average of the CARs for the target and acquirer, meaning that this is the CAR for an investor who held both firms before the announcement, as in the model.

[4]These empirical results should be interpreted with caution for two major reasons. First, the gains from M&A must be measured relative to the counterfactual in which the transaction did not take place. Second, as Betton et al. (2008) point out, most of the large-bidder acquisitions with negative returns that drive the average results are concentrated in a few firms in 1999 and 2000, which may indicate a temporary period of overvaluation and also confounds acquirer-size and time fixed effects in empirical studies.

[5]Formally, if Δn_{m} represents the number of successful M&A searches (acquisitions) over a small time interval Δt, then the variance of acquisitions per blueprint behaves like $\mathrm{var}(\Delta n_{\mathrm{m}}/n) \propto (\psi X_{\mathrm{m}}^{\varepsilon} n^{1-\varepsilon} \Delta t)/n^2 \propto 1/n$. For large n, this percentage variance tends to zero, so M&A announcements yield little news about future firm cashflows; in contrast, for small n, news of a successful M&A deal is almost entirely unanticipated, so returns should jump up upon announcement. Recall that firm value V grows with n on average, so the unexpected return upon announcement, $(\Delta V - \mathbb{E}\Delta V)/V$, behaves like $(\Delta n_{\mathrm{m}} - \mathbb{E}\Delta n_{\mathrm{m}})/n$ on average.

B.2 Micro-level decompositions: Derivation and properties

B.2.1 Derivation of empirical decomposition

To derive the decompositions (2) and (4), let us consider an arbitrary aggregate variable

$$Z_t \equiv \sum_{i=1}^{M_t} s_{it} Z_{it}, \tag{B.1}$$

where Z_{it} is some firm-level variable (e.g., value-to-sales) and s_{it} are some arbitrary weights that sum to one (e.g., sales shares). Note that the weights s_{it} need not have any relation to Z_{it}; we only assume that they are nonnegative and sum to one.

Before deriving the decomposition, let us define some useful notation. Let $\mathcal{I}_t$ represent the set of firms operating at time t. Let $\mathcal{I}_t^{\text{ent}}$ denote the set of firms that entered the sample at time t, $\mathcal{I}_t^{\text{cont}}$ the set of firms that continued to operate in both periods $t-1$ and t, and $\mathcal{I}_t^{\text{exit}}$ the set of firms that exited in period t. Let the operator Δ denote the change from one period to the next (e.g., $\Delta Z_t \equiv Z_t - Z_{t-1}$). Define $s_t^{\text{ent}} \equiv \sum_{i \in \mathcal{I}_t^{\text{ent}}} s_{it}$ to be the total entrant share at t, and $s_{t-1}^{\text{exit}} \equiv \sum_{i \in \mathcal{I}_t^{\text{exit}}} s_{i,t-1}$ the total exiter share at $t-1$. Note that this implies $s_t^{\text{cont}} = 1 - s_t^{\text{ent}}$ and $s_{t-1}^{\text{cont}} = 1 - s_{t-1}^{\text{exit}}$. Finally, define the conditional continuer shares as $s_{it}^{\text{cont}} \equiv s_{it}/s_t^{\text{cont}}$ and $s_{i,t-1}^{\text{cont}} \equiv s_{i,t-1}/s_{t-1}^{\text{cont}}$, implying conditional aggregates $Z_t^{\text{cont}} \equiv \sum_{i \in \mathcal{I}_t^{\text{cont}}} s_{it}^{\text{cont}} Z_{it}$.

Under this notation, we have $\mathcal{I}_t = \mathcal{I}_t^{\text{ent}} \cup \mathcal{I}_t^{\text{cont}}$, so the time-$t$ aggregate value is

$$Z_t = \sum_{i \in \mathcal{I}_t^{\text{ent}}} s_{it} Z_{it} + \sum_{i \in \mathcal{I}_t^{\text{cont}}} s_{it} Z_{it}. \tag{B.2}$$

Likewise, $\mathcal{I}_{t-1} = \mathcal{I}_t^{\text{exit}} \cup \mathcal{I}_t^{\text{cont}}$, so the lagged aggregate can be written

$$Z_{t-1} = \sum_{i \in \mathcal{I}_t^{\text{exit}}} s_{i,t-1} Z_{i,t-1} + \sum_{i \in \mathcal{I}_t^{\text{cont}}} s_{i,t-1} Z_{i,t-1}. \tag{B.3}$$

Differencing these equations, we have that the total change equals

$$\Delta Z_t = \sum_{i \in \mathcal{I}_t^{\text{ent}}} s_{it} Z_{it} - \sum_{i \in \mathcal{I}_t^{\text{exit}}} s_{i,t-1} Z_{i,t-1} + \sum_{i \in \mathcal{I}_t^{\text{cont}}} \left(s_{it} Z_{it} - s_{i,t-1} Z_{i,t-1} \right). \tag{B.4}$$

Adding and subtracting $s_t^{\text{ent}} Z_t^{\text{cont}}$ and $s_{t-1}^{\text{exit}} Z_{t-1}^{\text{cont}}$ on the right-hand side (and recalling that $s_t^{\text{ent}} = 1 - s_t^{\text{cont}}$ and $s_{t-1}^{\text{exit}} = 1 - s_{t-1}^{\text{cont}}$) then implies

$$\Delta Z_t = \sum_{i \in \mathcal{I}_t^{\text{ent}}} s_{it}(Z_{it} - Z_t^{\text{cont}}) - \sum_{i \in \mathcal{I}_t^{\text{exit}}} s_{i,t-1}(Z_{i,t-1} - Z_{t-1}^{\text{cont}}) + \Delta Z_t^{\text{cont}}. \tag{B.5}$$

The last term can be decomposed as

$$\Delta Z_t^{\text{cont}} = \sum_{i \in \mathcal{I}_t^{\text{cont}}} (s_{it}^{\text{cont}} Z_{it} - s_{i,t-1}^{\text{cont}} Z_{i,t-1}), \tag{B.6}$$

$$= \sum_{i \in \mathcal{I}_t^{\text{cont}}} \frac{1}{2}(s_{it}^{\text{cont}} + s_{i,t-1}^{\text{cont}})\Delta Z_{it} + \sum_{i \in \mathcal{I}_t^{\text{cont}}} \Delta s_{it}^{\text{cont}} \frac{1}{2}(Z_{it} + Z_{i,t-1}). \tag{B.7}$$

Using the index notation $(\overline{\,\cdot\,})$ from (3), we therefore have the total decomposition

$$\Delta Z_t = \underbrace{\sum_{i \in \mathcal{I}_t^{\text{ent}}} s_{it}(Z_{it} - Z_t^{\text{cont}}) - \sum_{i \in \mathcal{I}_t^{\text{exit}}} s_{i,t-1}(Z_{i,t-1} - Z_{t-1}^{\text{cont}}) + \sum_{i \in \mathcal{I}_t^{\text{cont}}} (\overline{s_{it}^{\text{cont}}})\Delta Z_{it}}_{\text{within-firm change}}$$

$$+ \underbrace{\sum_{i \in \mathcal{I}_t^{\text{cont}}} \Delta s_{it}^{\text{cont}}(\overline{Z_{it}})}_{\text{compositional change}}. \tag{B.8}$$

If we redefine the lags and leads of entrants as described in the main text, then this decomposition can be written more concisely in the form (2) and (4).

B.2.2 Relation to other macro decompositions

The decomposition I construct is closely related to a larger class of macroeconomic decompositions. In the notation above, the Haltiwanger (1997) decomposition has components

$$\Delta Z_t = \underbrace{\sum_{i \in \mathcal{I}_t^{\mathrm{cont}}} s_{i,t-1} \Delta Z_{it}}_{\mathrm{within_H}} + \underbrace{\sum_{i \in \mathcal{I}_t^{\mathrm{cont}}} \Delta s_{it} \Delta Z_{it}}_{\mathrm{cross_H}} + \underbrace{\sum_{i \in \mathcal{I}_t^{\mathrm{cont}}} \Delta s_{it} (Z_{i,t-1} - Z_{t-1})}_{\mathrm{between_H}}$$

$$+ \underbrace{\sum_{i \in \mathcal{I}_t^{\mathrm{ent}}} s_{it} (Z_{it} - Z_{t-1})}_{\mathrm{entry_H}} - \underbrace{\sum_{i \in \mathcal{I}_t^{\mathrm{exit}}} s_{i,t-1} (Z_{i,t-1} - Z_{t-1})}_{\mathrm{exit_H}}. \quad \text{(B.9)}$$

Haltiwanger (1997) and Foster, Haltiwanger, and Krizan (2001) use this decomposition as a slight adjustment to that of Baily, Hulten, and Campbell (1992). Conventionally, these decompositions have been used to understand the components of productivity growth from plant-level data (see, e.g., Foster et al., 2001; Bartelsman & Doms, 2000; Lentz & Mortensen, 2008). More recently, the same decomposition has been used to understand the drivers of secular macro trends. For example, De Loecker et al. (2020) use it to decompose the rise in the sales-weighted aggregate markup.

While this decomposition may be ideal for studying the drivers of productivity growth, it is not suitable for studying secular change in aggregate valuations for two reasons. First, the cross term is mechanically very negative, because transitory shocks to sales shares tend to be highly negatively correlated with transitory shocks to value-to-sales ratios (through the denominator).[6] The solution to this is to use the indices $(\bar{\cdot})$ for continuers. Without entry and exit, this indexing is equivalent to adding half of the cross term to the within effect and

---

[6] This is not a problem for De Loecker et al. (2020), who show that the cross term is close to zero for their markup series.

the other half to the between effect:[7]

$$\sum_{i=1}^{M} \overline{s_{it}} \Delta Z_{it} = \underbrace{\sum_{i=1}^{M} s_{i,t-1} \Delta Z_{it}}_{\text{within}_{\text{H}}} + \underbrace{\frac{1}{2} \sum_{i=1}^{M} \Delta s_{it} \Delta Z_{it}}_{\text{cross}_{\text{H}}}, \tag{B.10}$$

$$\sum_{i=1}^{M} \Delta s_{it} \overline{Z_{it}} = \underbrace{\sum_{i=1}^{M} \Delta s_{it} Z_{i,t-1}}_{\text{between}_{\text{H}}} + \underbrace{\frac{1}{2} \sum_{i=1}^{M} \Delta s_{it} \Delta Z_{it}}_{\text{cross}_{\text{H}}}. \tag{B.11}$$

When we have entry and exit, these will no longer be quite the same, because I adjust shares s_{it} to s_{it}^{cont}. The second reason we cannot use (B.9) to study secular change in valuations is that there may be life-cycle dynamics that bias the within, entry, and exit terms. For example, suppose firms tend to enter with high valuations and then these valuations tend to fall with age. The within term will be negative, the entry term positive, and the two will exactly offset if the distribution is stationary. We therefore (1) compare entrant valuations to time-t continuer valuations (not time-$(t-1)$) and (2) sum the within and net entry terms. Next, I illustrate how these adjustments neutralize biases.

B.2.3   Illustrative examples

To illustrate the properties of the above decomposition, I consider three stylized economies. The first two show the advantage of using the indices for continuers in the presence of transitory and persistent idiosyncratic shocks, while the third illustrates why entry and exit are included in the within-firm change. I consider a continuum of firms in each case in order to abstract from sample noise.

Example 1: Transitory idiosyncratic shocks   Suppose there is no entry or exit, and the set of firms are ex ante homogeneous. Firm i has value V_{it} and sales Y_{it}, where

$$Y_{it} = \bar{Y} \epsilon_{\text{Y}it} \quad \text{and} \quad V_{it} = \bar{V} \epsilon_{\text{V}it}, \tag{B.12}$$

---

[7]Griliches and Regev (1995) propose a similar accounting for TFP. Foster et al. (2001) argue that the Haltiwanger (1997) decomposition is more ideal for TFP accounting, because the cross term is meaningful in that context.

where the shocks $\epsilon_{\mathrm{Y}it}$ and $\epsilon_{\mathrm{V}it}$ are independent over time and of each other—that is, they are completely transitory. Without loss of generality, assume that $\mathbb{E}[\epsilon_{\mathrm{Y}it}] = \mathbb{E}[\epsilon_{\mathrm{V}it}] = 1$.

The aggregate value-to-sales ratio is then a constant

$$\frac{V_t}{Y_t} = \frac{\bar{V}}{\bar{Y}}. \tag{B.13}$$

Now let us consider what happens to the decomposition terms. The valuation ratio and sales share of firm i are

$$\frac{V_{it}}{Y_{it}} = \frac{\bar{V}}{\bar{Y}} \frac{\epsilon_{\mathrm{V}it}}{\epsilon_{\mathrm{Y}it}} \quad \text{and} \quad \frac{Y_{it}}{Y_t} = \epsilon_{\mathrm{Y}it}. \tag{B.14}$$

Hence, the Haltiwanger terms from $t-1$ to t equal

$$\mathrm{within_H} \equiv \frac{\bar{V}}{\bar{Y}} \int_0^1 \epsilon_{\mathrm{Y}i,t-1} \left(\frac{\epsilon_{\mathrm{V}it}}{\epsilon_{\mathrm{Y}it}} - \frac{\epsilon_{\mathrm{V}i,t-1}}{\epsilon_{\mathrm{Y}i,t-1}} \right) di \tag{B.15}$$

$$\mathrm{between_H} \equiv \frac{\bar{V}}{\bar{Y}} \int_0^1 (\epsilon_{\mathrm{Y}it} - \epsilon_{\mathrm{Y}i,t-1}) \frac{\epsilon_{\mathrm{V}i,t-1}}{\epsilon_{\mathrm{Y}i,t-1}} di \tag{B.16}$$

$$\mathrm{cross_H} \equiv \frac{\bar{V}}{\bar{Y}} \int_0^1 (\epsilon_{\mathrm{Y}it} - \epsilon_{\mathrm{Y}i,t-1}) \left(\frac{\epsilon_{\mathrm{V}it}}{\epsilon_{\mathrm{Y}it}} - \frac{\epsilon_{\mathrm{V}i,t-1}}{\epsilon_{\mathrm{Y}i,t-1}} \right) di. \tag{B.17}$$

Evaluating these expectations and using the fact that all shocks are uncorrelated implies

$$\mathrm{within_H} = \frac{\bar{V}}{\bar{Y}} \left(\mathbb{E}[\epsilon_{\mathrm{Y}i}^{-1}] - 1 \right) > 0, \tag{B.18}$$

$$\mathrm{between_H} = \frac{\bar{V}}{\bar{Y}} \left(\mathbb{E}[\epsilon_{\mathrm{Y}i}^{-1}] - 1 \right) > 0, \tag{B.19}$$

$$\mathrm{cross_H} = \frac{\bar{V}}{\bar{Y}} 2 \left(1 - \mathbb{E}[\epsilon_{\mathrm{Y}i}^{-1}] \right) < 0. \tag{B.20}$$

($\mathbb{E}[\epsilon_{\mathrm{Y}i}^{-1}] > 1$ follows from Jensen's inequality.) Intuitively, the within term is positive because lagged sales shares are on average going to mean revert in the next period, creating a positive covariance between lagged sales share and next-period valuation—for example, firms for which sales shares are unusually high at $t-1$ will tend to have unusually low valuations at $t-1$, but these will on average mean revert at t, raising the valuation. Similar intuition holds for the between term.

A.13

The indexing adjustment corrects this bias:

$$\text{within}_\text{H} + \frac{1}{2}\text{cross}_\text{H} = \text{between}_\text{H} + \frac{1}{2}\text{cross}_\text{H} = 0. \tag{B.21}$$

Consistent with a stationary firm distribution, both the within-firm change and compositional change are now zero in this economy.

Example 2: Persistent idiosyncratic shocks Consider the setting in the previous example, but now, instead of independent transitory shocks, the numerator and denominator are subject to (potentially correlated) persistent shocks. Specifically, the log sales of firm i, denoted $y_{it} \equiv \log Y_{it}$, follow the autoregression

$$y_{it} = \log \bar{Y} + \phi_y y_{i,t-1} + \epsilon_{yit}, \tag{B.22}$$

where $\phi_y \leq 1$ is the persistence of the shocks. Similarly, the log value-to-sales ratio, $v_{it} \equiv \log(V_{it}/Y_{it})$, follows the autoregression[8]

$$v_{it} = \bar{v} + \phi_v v_{i,t-1} + \epsilon_{vit}. \tag{B.23}$$

The shocks $[\epsilon_{yit}, \epsilon_{vit}]$ are iid over time and, at each time t, jointly normal with variances σ_y^2 and σ_x^2 and covariance σ_{xy}. It is straightforward to compute the aggregate value-to-sales ratio V_t/Y_t and see that it is a constant.

The compositional change (corrected with the index) equals

$$\text{between}_\text{H} + \frac{1}{2}\text{cross}_\text{H} = \frac{1}{2}\int_0^1 (e^{y_{it}} - e^{y_{i,t-1}})(e^{v_{it}} + e^{v_{i,t-1}})di \tag{B.24}$$

$$= \frac{1}{2}\mathbb{E}\left[e^{y_{it}+v_{i,t-1}} - e^{y_{i,t-1}+v_{it}}\right]. \tag{B.25}$$

Substituting in $y_{it} = \bar{y} + \sum_{j=0}^{\infty} \phi_y^j \epsilon_{yi,t-j}$ and $v_{it} = \bar{v} + \sum_{j=0}^{\infty} \phi_v^j \epsilon_{vi,t-j}$, the first term in the

---

[8]Specifying stationary dynamics for firm-level value-to-sales ratios ensures a stationary distribution.

A.14

expectation equals

$$\mathbb{E}\left[e^{y_{it}+v_{i,t-1}}\right] = \exp\left\{\frac{1}{2}\left(\text{var}(y_{it}) + \text{var}(v_{i,t-1}) + 2\text{cov}(y_{it}, v_{i,t-1}))\right\} \tag{B.26}$$

$$= \exp\left\{\frac{1}{2}\left(\sum_{j=0}^{\infty}\phi_y^{2j}\sigma_y^2 + \sum_{j=0}^{\infty}\phi_v^{2j}\sigma_v^2 + 2\sum_{j=1}^{\infty}(\phi_y\phi_v)^j\sigma_{yv}\right)\right\} \tag{B.27}$$

$$= \exp\left\{\frac{1}{2}\left(\frac{\sigma_y^2}{1-\phi_y^2} + \frac{\sigma_v^2}{1-\phi_v^2} + 2\frac{\phi_y\phi_v}{1-\phi_y\phi_v}\sigma_{yv}\right)\right\}. \tag{B.28}$$

By the same algebra, one can see that $\mathbb{E}\left[e^{y_{i,t-1}+v_{it}}\right]$ is the exact same. Hence, the total compositional change (and, in turn, the within-firm change) is zero:

$$\text{between}_{\text{H}} + \frac{1}{2}\text{cross}_{\text{H}} = \text{within}_{\text{H}} + \frac{1}{2}\text{cross}_{\text{H}} = 0. \tag{B.29}$$

Again, the indexing corrects the bias that would be present in the individual terms.

Example 3: Life-cycle trends and entry This final example demonstrates how adding net entry to the within-firm change corrects for life-cycle trends in firm valuations. Consider an economy with a mass of firms producing aggregate output Y_t. At each time t, a mass of new entrants begins producing; the total output of this cohort is proportional to last period's total output:

$$Y_t^{(0)} = \theta Y_{t-1}. \tag{B.30}$$

All other firms continue producing the same amount as before, so it follows that

$$Y_t = (1 + \theta)Y_{t-1}. \tag{B.31}$$

This means that the total sales share of firms aged τ equals

$$\frac{Y_t^{(\tau)}}{Y_t} = \frac{\theta}{1+\theta}(1+\theta)^{-\tau}. \tag{B.32}$$

Furthermore, assume that every entering firm has the same value-to-sales ratio

$$\frac{V_{it}}{Y_{it}} = v^{(0)} = \bar{v}. \tag{B.33}$$

As firms age, their value-to-sales ratios fall by a factor of $(1 + \delta)^{-1}$: for a firm of age τ, the valuation is therefore

$$\frac{V_{it}}{Y_{it}} = v^{(\tau)} = \bar{v}(1 + \delta)^{-\tau}. \tag{B.34}$$

All of these facts together imply that the aggregate value-to-sales ratio equals the constant

$$\frac{V_t}{Y_t} = \sum_{\tau=0}^{\infty} \frac{Y_t^{(\tau)}}{Y_t} v^{(\tau)} = \bar{v}\frac{\theta(1 + \delta)}{\theta(1 + \delta) + \delta}. \tag{B.35}$$

The entrants offset the declining valuations of aging firms, yielding a stationary distribution.

Consider first the Haltiwanger decomposition terms from $t - 1$ to t:

$$\text{entry}_{\text{H}} = \frac{Y_t^{(0)}}{Y_t}\left(\bar{v} - \frac{V_t}{Y_t}\right) > 0, \tag{B.36}$$

$$\text{within}_{\text{H}} = \sum_{\tau=0}^{\infty} \frac{Y_{t-1}^{(\tau)}}{Y_{t-1}}(v^{(\tau+1)} - v^{(\tau)}) < 0, \tag{B.37}$$

$$\text{between}_{\text{H}} = \sum_{\tau=0}^{\infty} \left(\frac{Y_t^{(\tau+1)}}{Y_t} - \frac{Y_{t-1}^{(\tau)}}{Y_{t-1}}\right) v^{(\tau)} < 0, \tag{B.38}$$

$$\text{cross}_{\text{H}} = \sum_{\tau=0}^{\infty} \left(\frac{Y_t^{(\tau+1)}}{Y_t} - \frac{Y_{t-1}^{(\tau)}}{Y_{t-1}}\right) (v^{(\tau+1)} - v^{(\tau)}) > 0. \tag{B.39}$$

The entry term is positive because entrants have above-average valuation ratios. The within term is negative because, conditional on being a continuer, valuations tend to fall over time. The between term is negative because sales shares of continuers fall as entrants crowd them out. The cross term is positive because of the interaction of the within and between effects.

The adjusted decomposition makes two important adjustments to the above terms. First, the sales shares among continuers are defined relative to total continuer output ($Y_t^{\text{cont}} = \sum_{\tau=1} Y_t^{(\tau)}$), not aggregate output ($Y_t = Y_t^{\text{cont}} + Y_t^{(0)}$). Consequently, the "between" term

above is no longer negative, but zero:

$$\text{between}_{\text{H}} = \sum_{\tau=0}^{\infty} \left(\frac{Y_t^{(\tau+1)}}{Y_t^{\text{cont}}} - \frac{Y_{t-1}^{(\tau)}}{Y_{t-1}} \right) v^{(\tau)} = 0, \tag{B.40}$$

and similarly for the "cross" term. Second, by adding the entry term to the within-firm term, we offset the deterministic life-cycle trend in valuations, so that it is now also zero:

$$\text{entry}_{\text{H}} + \text{within}_{\text{H}} = 0. \tag{B.41}$$

This example works similarly if we allow firm death and add the resulting exit term to the within-firm component.

B.2.4 Details about computation in the data

In the examples above, we assumed a continuum of firms, which allowed us to apply a law of large numbers. In the data, there is a finite number of firms, and thus these decompositions can be influenced by extreme outliers. To mitigate this, I do two things. First, I winsorize firm-level ratios at the 10th and 90th percentiles within each year before conducting the decomposition. Winsorizing gives a more conservative estimate of the total amount of reallocation that took place over this period; the fact that there is still substantial reallocation confirms that the decomposition results are not driven by extremely high value-to-sales firms.[9] Second, to stabilize changes in firm-level ratios, I smooth sales (the denominator) as an average of its current and lagged value. This has a minimal effect in the long-run, but helps reduce outlier-driven noise in the short-run.

---

[9]Specifically, choosing smaller cutoffs (e.g., the 1st and 99th percentile) or larger cutoffs yields the same conclusion: most of the increase in the aggregate value-to-sales ratio comes from a compositional change. Under smaller cutoffs, the positive composition effect in the decomposition is even larger, which is why I say that this choice is "conservative."

B.3 Micro-level decompositions: Additional results

B.3.1 Separating out firm entry and exit

The main decompositions add the entry and exit margins into the within-firm change, for the reasons laid out above in Appendix B.2. Figure B.3 separates these two margins out and plots the cumulative contribution of the four decomposition terms for value-to-sales and cashflow-to-value. In both cases, both the entry and exit margins are very small, implying that the vast majority of aggregate variation is coming from continuers.

Figure B.3: Decompositions of aggregate valuations with entry and exit components
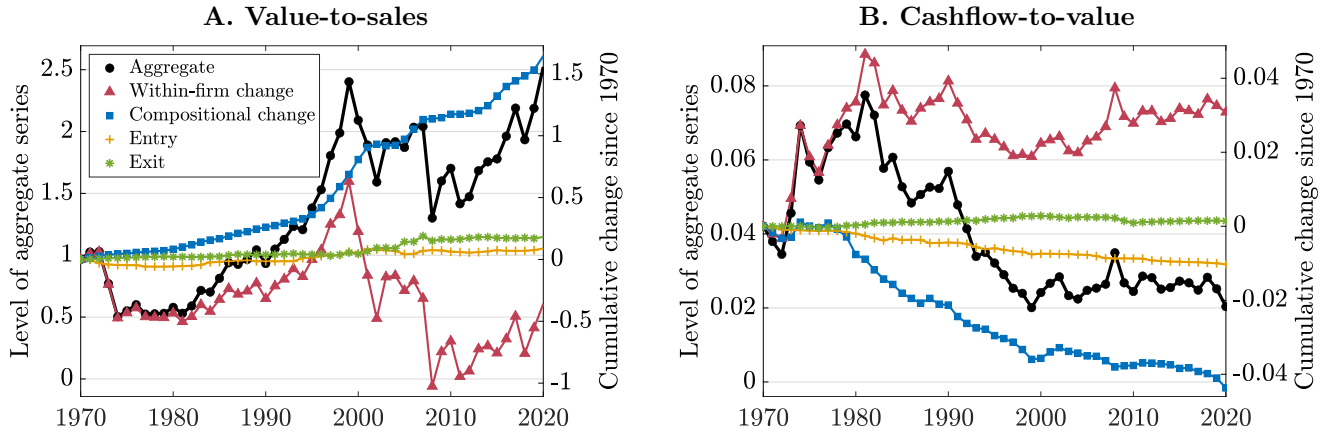
The figure plots the decompositions of changes in aggregate valuations into within-firm, compositional, and entry and exit components. The decompositions for value-to-sales and cashflow-to-value are as in (2) and (4), respectively, except that now the entry and exit components are not included in the within-firm change. See the text for details. Value is defined as market equity plus book debt, less cash; and cashflows are defined as dividends plus interest payments.

B.3.2 Equity value and alternative cashflow measures

The baseline measure of valuations and cashflows uses total firm value (equity value plus net debt) and total cashflows (dividends plus interest payments). One might wonder whether these decomposition results also hold for the equity component of firms. Figure B.4 plots the decompositions from the main text for equity only—that is, defining value as the market value of equity and cashflows as dividends. The results are virtually unchanged.

Figure B.4: Decompositions of aggregate equity valuations

A. Equity value to sales

B. Dividend yield on equity
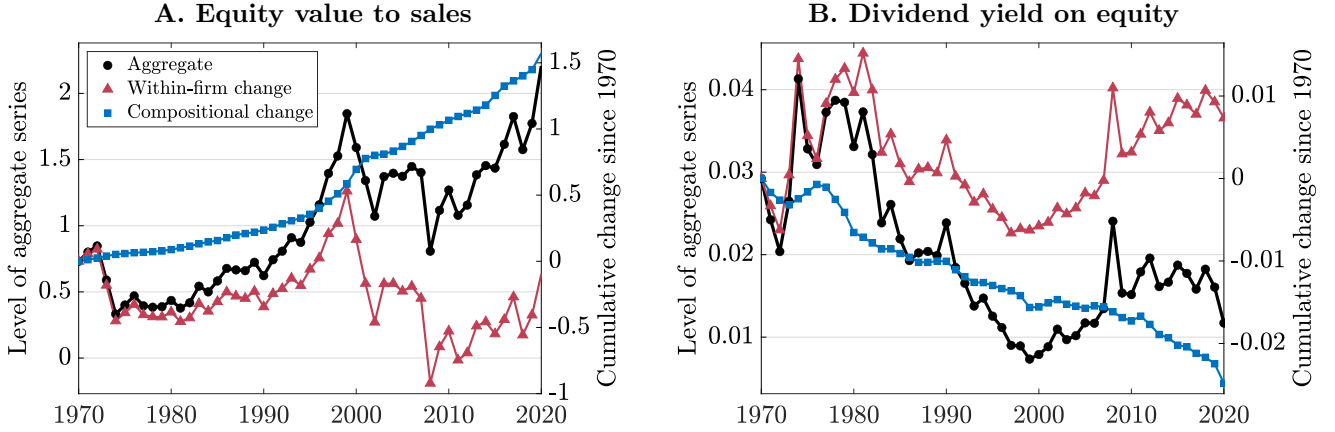
The figure plots the decompositions of changes in aggregate equity valuations into within-firm and compositional components. The decomposition for value-to-sales is defined in (2) and the decomposition for cashflow-to-value in (4). See the main text for an explanation. Value is defined as market equity and cashflows are defined as dividends.

Another question that might arise is whether the cashflow-to-value results are sensitive to the definition of payouts to investors. One could argue that, from an investor's perspective, what matters for valuation is the total payout, inclusive of net share repurchases. Figure B.5 plots the decompositions of cashflow-to-value after adding net repurchases to the numerator. Panel A corresponds to total firm value and cashflows (i.e., including debt) and Panel B corresponds to equity only. While adding net buybacks does attenuate the total increase in aggregate valuations, the decomposition results remain the same: there has been a large reallocation of market share toward high-valuation (low cashflow-to-value) firms over time.

Figure B.5: Decompositions of aggregate cashflow-to-value including share buybacks
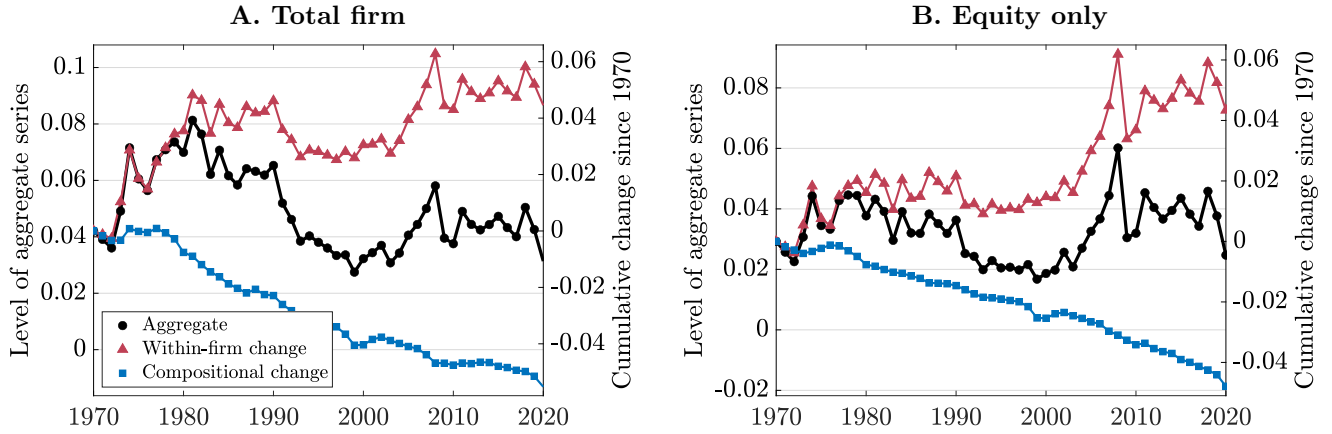
The figure plots the decompositions of changes in aggregate cashflow-to value ratios where cashflows include net buybacks of equity. The within-firm and compositional components of the decompositions are defined in (4)—see the main text for an explanation. In Panel A, value is defined as market equity plus book debt, less cash; and cashflows are defined as dividends and interest payments plus net equity buybacks. In Panel B, value is only market equity while cashflows are dividends and net buybacks.

B.3.3 Implications for R&D-to-sales

An alternative measure of innovation effort that is sometimes used in the literature is the ratio of aggregate R&D to aggregate sales (or output). The black line in Figure B.6 shows that this aggregate ratio has increased since 1975. A main insight of this paper is that the ratio of R&D to *value*, not sales, is the ideal moment for identifying innovation productivity, because, holding innovation productivity fixed, higher present values will increase R&D spending relative to output. Still it is informative to understand what happened to R&D-to-sales at the micro and macro levels.

The first fact that is key to understanding trends in R&D-to-sales is that high-valuation firms tend to be high-R&D-to-sales firms. Table B.1 reports estimates of cross-sectional regressions of valuations relative to sales and cashflows on R&D-to-sales ratios. There is a significant positive relationship between valuations and R&D intensities within years (specifications (2) and (5)) and even within narrowly defined industries (specifications (3) and (6)). Perhaps most importantly, the R^2 values from these regressions are high (especially for value-to-sales), implying that R&D-to-sales explains a large proportion of the variation in firm-level valuations. This is intuitive, because firms with high valuations have a greater incentive to innovate, and firms with high rates of innovation will tend to have higher future

A.20

cashflows. Put simply, high-valuation firms are high-growth firms.

Table B.1: R&D intensity and valuation ratios in the cross-section of firms

Dependent variable:	$\log(\text{value}_{it}/\text{sales}_{it})$			$\log(\text{value}_{it}/\text{cashflow}_{it})$		
	(1)	(2)	(3)	(4)	(5)	(6)
$\log(\text{R\&D}_{it}/\text{sales}_{it})$	0.508***	0.478***	0.511***	0.164***	0.133***	0.096***
	(0.005)	(0.006)	(0.008)	(0.009)	(0.009)	(0.018)
Year FE		Y			Y	
Year $\times$ NAICS-6 FE			Y			Y
R^2	0.457	0.492	0.569	0.079	0.252	0.524
Within R^2	0.457	0.421	0.335	0.079	0.062	0.018

The table reports cross-sectional regressions of firm-level valuation ratios (value to sales and value-to-cashflows) on firm-level R&D-to-sales ratios. Specifications (2) and (5) control for year fixed effects; while (3) and (6) control for year×industry fixed effects, where the industry is defined at the NAICS 6-digit level. The within R^2 values represent the R^2 values after demeaning by fixed effects (i.e., the variation explained by R&D-to-sales alone). Standard errors are clustered at the firm level.

Second, because high-valuation firms have high R&D intensities, we should expect a similar reallocation toward firms with high R&D-to-sales ratios. Figure B.6 decomposes aggregate R&D-to-sales into within-firm changes and composition changes (i.e., reallocation of sales), analogous to the valuation decompositions. On average, within-firm changes in R&D-to-sales have actually been negative. The rise in the aggregate is explained by a reallocation toward high-R&D-to-sales firms. The decline in R&D-to-sales within firms is consistent with relatively stable within-firm valuations (meaning incentives to innovate have not risen so much at the firm level) coupled with declining innovation productivity.

Figure B.6: Decomposition of aggregate R&D-to-sales
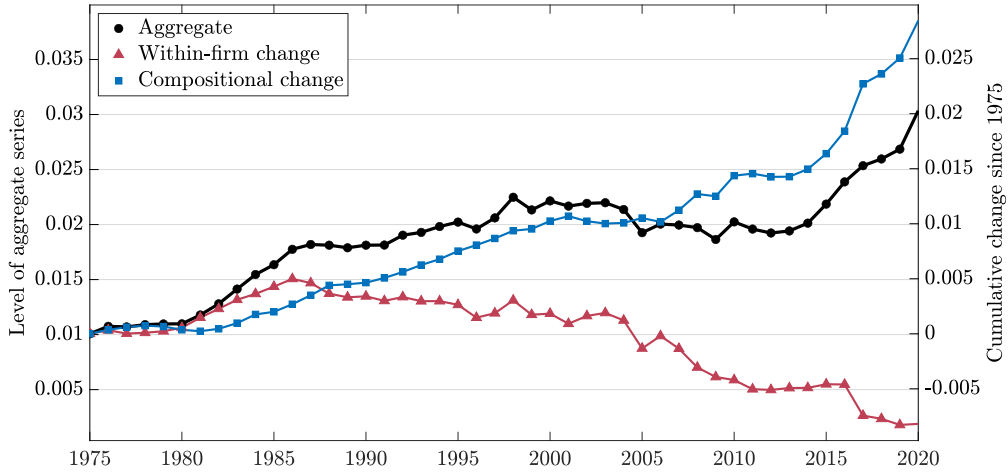
The figure plots the decompositions of changes in the aggregate R&D-to sales ratios. See the main text for an explanation of the within-firm and compositional components.

B.4   Additional evidence on public versus private firms

As explained in the main text, I establish two main facts about public versus private firms' R&D and M&A. The first, plotted in Panel A of Figure B.7, is that there has been no detectable increase in private-firm R&D intensity compared to public-firm R&D intensity. In particular, the black line plots the aggregate R&D-to-sales ratio among public firms (i.e., in Compustat), while the green line plots the aggregate R&D-to-GDP ratio reported by the NSF, which includes both public and private firms. If the R&D intensity of private firms rose dramatically relative to that of public firms, then we would expect the latter measure to rise by much more than the former; however, the R&D-to-GDP ratio of all firms is even more stable than R&D-to-sales of public firms.

Figure B.7: Public versus private R&D and M&A

A. Public versus private R&D intensity

B. Public versus private acquisition shares
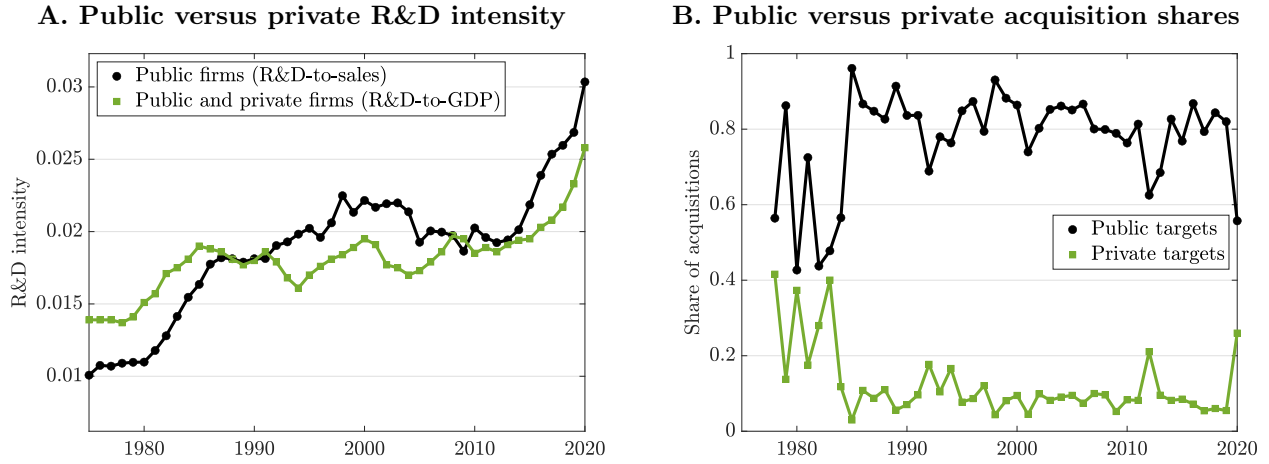
The figure compares R&D and M&A across public and private firms. Panel A plots the R&D intensities of public and private firms over time. Public firm R&D intensity is the aggregate ratio of R&D to sales across all firms in Compustat. Public and private R&D intensity is the ratio of total R&D spending to GDP in the U.S., as reported by the NSF. Panel B plots the share of M&A spending by public acquirers on public versus private targets over time. M&A is defined here as mergers (purchases of the entire target). Note that these lines do not necessarily add to one, because there are other categories of targets (e.g., government).

The second fact, plotted in Panel B of Figure B.7, is that there has been no increase in the share of M&A by public acquirers of private targets. Specifically, the figure plots the percent of all public firms' M&A spending (in dollars) on private targets and public targets in mergers (deals in which they acquired 100% of the target).[10] If public firms were increasingly buying private targets to grow, the share of spending on private targets would rise over time. In the data, there is no such increase—in fact, there is a very slight decrease.

B.5 Additional evidence across industries

B.5.1 Decompositions across industries

Figure B.8 plots the decompositions (2) and (4) at the industry level instead of the firm level. Panels A and B use NAICS 2-digit industry definitions and C and D use NAICS 3-digit industries. In all plots, the amount of reallocation is much smaller than at the firm level, suggesting that most of the reallocation was a within-industry phenomenon.

---

[10]The two lines do not necessarily add to one, because there are alternative categories of targets (e.g., government entities).

Figure B.8: Decompositions of aggregate valuations across industries
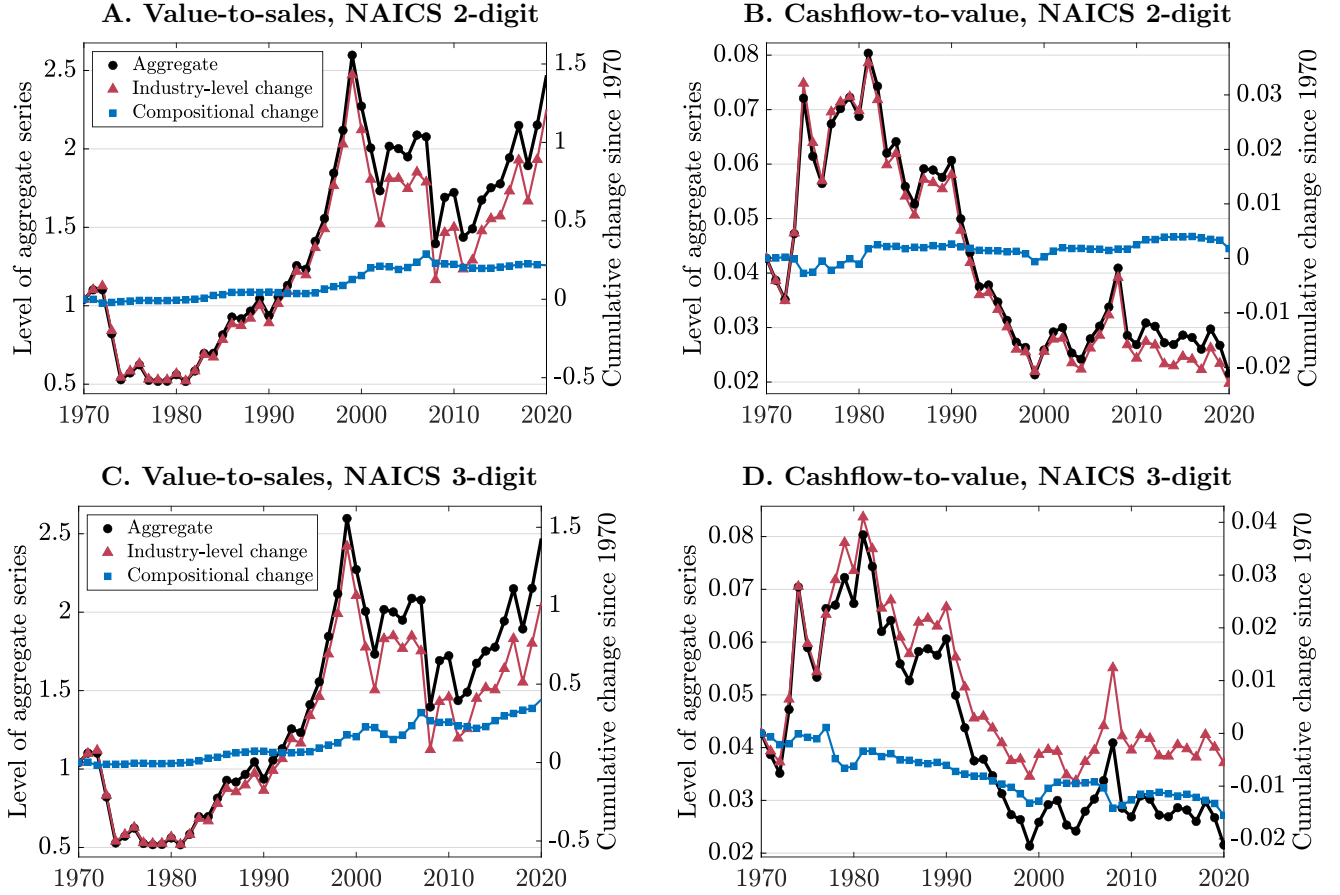
The figure plots the decompositions of changes in aggregate valuations into industry-level and compositional changes. The decompositions for value-to-sales and cashflow-to-value are as in (2) and (4), respectively, at the industry level (instead of the firm level). Panels A and B use NAICS 2-digit industry classifications while Panels C and D use NAICS 3-digit classifications. Value is defined as market equity plus book debt, less cash; and cashflows are defined as dividends plus interest payments. Industry-level valuation ratios are winsorized at the 1% and 99% levels.

B.5.2 Cross-industry regressions

As discussed in the main text, we can exploit cross-industry variation over time to test whether the key macro facts are related in the cross-section. The first prediction I test is whether those industries with the largest declines in R&D-to-value had the largest increase in market valuations and M&A intensity. Table B.2 reports the results for four regressions for NAICS-5 industries. The first specification regresses the total (log) change in an industry value-to-sales ratio from 1980 to 2020 on the total change in the industry's R&D-to-value ratio over the same period. The negative coefficient estimate implies that those sectors with

the largest decline in R&D-to-value had the largest increase in total value-to-sales ratios. Similarly, the other three regressions imply that falling R&D-to-value is associated with rising value-to-cashflows, rising M&A-to-R&D, and rising M&A-to-value. These findings corroborate the first prediction.

Table B.2: Industry regressions: Changes in valuations and M&A versus R&D, 1980–2020

Dependent variable:	$\Delta \log \left(\frac{\text{value}_k}{\text{sales}_k} \right)$	$\Delta \log \left(\frac{\text{value}_k}{\text{cashflow}_k} \right)$	$\Delta \log \left(\frac{\text{M\&A}_k}{\text{R\&D}_k} \right)$	$\Delta \log \left(\frac{\text{M\&A}_k}{\text{value}_k} \right)$
	(1)	(2)	(3)	(4)
$\Delta \log(\text{R\&D}_k/\text{value}_k)$	-0.189^{***}	-0.130^{***}	-0.424^{***}	-0.184^{***}
	(0.031)	(0.033)	(0.097)	(0.086)
Observations	515	515	236	279
R^2	0.069	0.030	0.076	0.016

The table reports the coefficient estimates for regressions of the cumulative change in industry-level valuations ((1) value-to-sales and (2) value-to-cashflow) and M&A intensity ((3) M&A-to-R&D and (4) M&A-to-value) on the cumulative change in industry-level R&D-to-value. Industries are NAICS 5-digit sectors, excluding any sector for which there are years with missing observations (i.e., zero firms) or for which there are not enough observations to compute the regression variables (i.e., insufficient consecutive R&D or M&A observations). Industry-level variables (value, sales, cashflow, M&A, and R&D) are defined as the sums across firms within the industry—for example, if sector k is composed of M_{kt} firms $i \in \{1, \ldots, M_{kt}\}$, then value$_{kt} = \sum_{i=1}^{M_{kt}}$ value$_{it}$. The regression variables are then the cumulative log change in each industry-level variable from 1980 to 2020 (so there is one observation per industry). M&A spending includes both mergers and acquisitions of assets. In all specifications, variables are winsorized at the 1% and 99% level.

The second prediction I test is that those industries in which R&D-to-value fell the most are also the industries that saw the most reallocation toward high-valuation firms. To measure this latter effect, I run the decompositions (2) and (4) *within* each industry k and sum up the cumulative compositional change from 1980 to 2020. Table B.3 reports the results of these regressions. Specification (1) shows that declining R&D-to-value is associated with a larger positive compositional change in value-to-sales (i.e., more reallocation to high-valuation firms); while specification (3) shows that it is associated with more reallocation to low cashflow-to-value firms. Both validate the prediction. To see whether this result is sensitive to the sign of the R&D-to-value change, specifications (2) and (4) condition on only those industries for which the cumulative change in R&D-to-value was negative. The results are the same.

Table B.3: Industry regressions: Changes in composition versus R&D, 1980–2020

Dependent variable:	Compositional change in valuation			
	$\text{value}_k/\text{sales}_k$		$\text{cashflow}_k/\text{value}_k$	
	(1)	(2)	(3)	(4)
$\Delta \log(\text{R\&D}_k/\text{value}_k)$	-0.1038^{***}	-0.3160^{***}	0.0020^{***}	0.0021^{**}
	(0.0519)	(0.0980)	(0.0006)	(0.0013)
Subsample with $\Delta \log(\text{R\&D}_k/\text{value}_k) < 0$		Y		Y
Observations	515	307	515	307
R^2	0.008	0.033	0.019	0.009

The table reports the coefficient estimates for regressions of the cumulative compositional change in industry-level valuations (value-to-sales and cashflow-to-value) on the cumulative change in industry-level R&D-to-value. Industries are NAICS 5-digit sectors, excluding any sector for which there are years with missing observations (i.e., zero firms) or for which there are not enough observations to compute the regression variables (i.e., insufficient consecutive R&D observations). Industry-level variables (value, sales, cashflow, and R&D) are defined as the sums across firms within the industry—for example, if sector k is composed of M_{kt} firms $i \in \{1, \ldots, M_{kt}\}$, then $\text{value}_{kt} = \sum_{i=1}^{M_{kt}} \text{value}_{it}$. For value-to-sales, the compositional change for sector k from time $t-1$ to time t is defined as $\sum_{i=1}^{M_{kt}} \Delta \left(\frac{\text{sales}_{it}}{\text{sales}_{kt}^{\text{cont}}} \right) \times \frac{1}{2} \left(\frac{\text{value}_{it}}{\text{sales}_{it}} + \frac{\text{value}_{i,t-1}}{\text{sales}_{i,t-1}} \right)$, where $\text{sales}_{kt}^{\text{cont}}$ is the total sales of firms in sector k that continued operating from $t-1$ and t. The cumulative change is then the sum over time (from 1980 to 2020) of all of these changes (so there is one observation per industry). The compositional change for the cashflow-to-value is similar, but uses value weights instead of sales weights. Specifications (1) and (3) consider all sectors, while (2) and (4) consider only sectors for which the cumulative change in R&D-to-value was negative. In all specifications, variables are winsorized at the 5% and 95% level.

C Derivation of equilibrium conditions

C.1 Profits and production

C.1.1 Household final-good demand

Consider any individual household in the economy. The household chooses quantities $\{y_{ijt}\}$ of each good in order to maximize the value of its total consumption (6) and (7). The household takes the prices $\{p_{ijt}\}$ of these goods as given. Let W_{ct} denote the total expenditure of the household on all consumption goods (to be determined in the household's dynamic consumption-saving problem). Its budget constraint is then

$$W_{ct} \geq \int_0^{N_t} \sum_{i=1}^{m_{jt}} p_{ijt} y_{ijt} dj. \tag{C.1}$$

If we define $\hat{y}_{ij} \equiv q_{ij}y_{ij}$ and $\hat{p}_{ij} \equiv p_{ij}/q_{ij}$, then this optimization solves the static Lagrangian

$$\mathcal{L}_{ct} = \max_{\{\hat{y}_{ijt}\}} \left\{ c(\{\hat{y}_{ijt}\}) + \lambda_{ct} \left(W_{ct} - \int_0^1 \sum_{i=1}^{m_{jt}} \hat{p}_{ijt}\hat{y}_{ijt}dj \right) \right\}, \tag{C.2}$$

where $c(\{\hat{y}_{ijt}\})$ nests (6) and (7). Stated this way, the problem is of the same form as the standard CES optimization without quality differences. The first-order conditions are

$$\left(\frac{y_{jt}}{c_t} \right)^{-1/\eta} = \lambda_{ct}\hat{p}_{ijt}. \tag{C.3}$$

This implies that, within product market j, firms must equate quality-adjusted prices:

$$\hat{p}_{ijt} = \hat{p}_{i'jt}, \tag{C.4}$$

for all i and i' producing in the market. Denote by $p_j = \hat{p}_{ij}$ the price index for the whole product market, and note that $p_j y_j = \sum_{i=1}^{m_j} p_{ij}y_{ij}$.

Given this price index, the remainder of the solution follows standard CES algebra. First, the first-order condition (C.3) implies that, for any goods j and j',

$$p_{jt}^{\eta}y_{jt} = p_{j't}^{\eta}y_{j't}. \tag{C.5}$$

This means that the household's total expense on consumption is

$$\int_0^{N_t} p_{j't}y_{j't}dj' = p_{jt}^{\eta}y_{jt} \int_0^{N_t} p_{j't}^{1-\eta}dj'. \tag{C.6}$$

Now define the aggregate price index p_t such that $p_t c_t = \sum_j p_{jt}y_{jt}$. Substituting this into the left-hand side of the last equation, then aggregating the y_{jt} up to c_t implies

$$p_t = \left(\int_0^{N_t} p_{jt}^{1-\eta}dj \right)^{1/(1-\eta)}. \tag{C.7}$$

This implies that good j's price satisfies

$$p_{jt} = \left(\frac{y_{jt}}{c_t}\right)^{-1/\eta} p_t, \tag{C.8}$$

and hence

$$p_{ijt} = q_{ijt} \left(\frac{y_{jt}}{c_t}\right)^{-1/\eta} p_t. \tag{C.9}$$

for every firm i in variety j.

This is the solution for a single household; the final step is to aggregate across households. Note that the first-order condition (C.3) aggregates across households to

$$\left(\frac{Y_{jt}}{Y_t}\right)^{-1/\eta} = \lambda_{ct}\hat{p}_{ijt}, \tag{C.10}$$

where Y_{jt} is aggregate consumption of good j and Y_t is aggregate consumption of all goods. Thus, all of the steps above hold in terms of aggregates, implying the aggregate demand curve

$$p_{ijt} = q_{ijt} \left(\frac{Y_{jt}}{Y_t}\right)^{-1/\eta} p_t. \tag{C.11}$$

Aggregate consumption is the numeraire, so we normalize $p_t = 1$.

C.1.2   Firm profit maximization

The firm's profits, and hence its pricing and production decisions, are separable across products, so it suffices to consider the profit maximization for a single product. Firms compete à la Bertrand, taking as given the demand curves (C.9) to determine their output. A standard result in Bertrand competition is that firms may be at a corner solution if competition is sufficiently intense. Substituting the demand curve into firm i's profits for product j implies profits

$$\Pi_{ijt} = \max_{p_{ijt}} \left(p_{ijt} - \frac{w_{\mathrm{p}t}}{a_i}\right) Y_{ijt} \tag{C.12}$$

$$= \max_{p_{jt}} \left(p_{jt} - \frac{w_{\mathrm{p}t}}{a_i q_{ijt}}\right) p_{jt}^{-\eta}\left(Y_{jt} - \sum_{-i \neq i} q_{-ijt} Y_{-ijt}\right) \tag{C.13}$$

A.28

The second equality imposes the constraint (C.4). The firm chooses its (quality-adjusted) price $p_{jt} = p_{ijt}/q_{ijt}$ given the qualities and quantities $\{q_{-ijt}, Y_{-ijt}\}$ of competitors.

First, let us consider interior solutions: the first-order condition with respect to p_{jt} is

$$(1 - \eta)p_{jt}^{-\eta} + \eta \frac{w_{\text{p}t}}{a_i q_{ijt}} p_{jt}^{-\eta-1} = 0. \tag{C.14}$$

(Note that $Y_t - \sum_{-i \neq i} q_{-ijt} Y_{-ijt} = q_{ijt} Y_{ijt} > 0$ in an interior solution). This implies

$$p_{ijt} = q_{ijt} p_{jt} = \frac{\eta}{\eta - 1} \frac{w_{\text{p}t}}{a_i}. \tag{C.15}$$

Defining the firm's markup as its price relative to marginal cost, we have

$$\mu_{ijt} \equiv \frac{p_{ijt}}{w_{\text{p}t}/a_i} = \frac{\eta}{\eta - 1}, \tag{C.16}$$

the usual CES monopoly markup.

This interior solution will not be a Nash equilibrium if it is possible for a competitor to lower its price while still making a profit. Define a firm's quality-adjusted marginal cost as

$$mc_{ijt} \equiv \frac{w_{\text{p}t}}{a_i q_{ijt}}. \tag{C.17}$$

Now consider the two firms with the lowest marginal costs; denote by i_1 the firm with the lowest marginal cost (the "leader") and by i_2 that with the second lowest (the "follower"). Because quality-adjusted prices are equated through $p_j = p_{ij}/q_{ij}$, the markup of firm i_2 can be equivalently expressed as:

$$\mu_{i_2jt} = \frac{p_{jt}}{mc_{i_2jt}} = \frac{p_{i_1jt}}{q_{i_1jt} mc_{i_2jt}}. \tag{C.18}$$

The lowest price at which firm i_1 can guarantee firm i_2 will be unwilling to further reduce prices is that for which $\mu_{i_2jt} = 1$, or

$$p_{i_1jt} = q_{i_1jt} mc_{i_2jt}. \tag{C.19}$$

At this price, the follower earns zero profits and the markup of the leader is

$$\mu_{i_1jt} = \frac{q_{i_1jt}mc_{i_2jt}}{w_{\mathrm{p}t}/a_i} = \frac{mc_{i_2jt}}{mc_{i_1jt}} = \frac{a_{i_1}q_{i_1jt}}{a_{i_2}q_{i_2jt}}. \tag{C.20}$$

Combining these two solutions, the leader's markup equals

$$\mu_{ijt} = \min\left\{\frac{a_{i_1}q_{i_1jt}}{a_{i_2}q_{i_2jt}}, \frac{\eta}{\eta-1}\right\} \tag{C.21}$$

and the follower always has a markup of one. Because we have assumed that firms earning zero markups choose not to produce (Section 2), product j is only produced by the leader.

The last piece is to determine the set of possible market structures and resulting markups. If a firm creates a new variety, it obviously has no follower ($i_2 = \varnothing$). Alternatively, a firm could enter an existing market by innovating on an existing blueprint or by buying the leader's blueprint. In either of these cases, the quality gap is $q_{i_1}/q_{i_2} = \lambda$. Hence, we have six potential market structures, which can be summarized by the pair $\{i_1, i_2\} \in \{h, \ell\} \times \{h, \ell, \varnothing\}$. Denoting by $\mu_{i_1i_2}$ the markup given this pair, the results above imply that monopolists get the CES markup,

$$\mu_{h\varnothing} = \mu_{\ell\varnothing} = \frac{\eta}{\eta-1}, \tag{C.22}$$

same-type leaders earn (at least) the quality advantage,

$$\mu_{hh} = \mu_{\ell\ell} = \min\left\{\lambda, \frac{\eta}{\eta-1}\right\}, \tag{C.23}$$

high types leading low types earn both the quality and productivity advantage,

$$\mu_{h\ell} = \min\left\{\lambda\frac{a_h}{a_\ell}, \frac{\eta}{\eta-1}\right\}, \tag{C.24}$$

and low types leading high types earn

$$\mu_{\ell h} = \min\left\{\lambda\frac{a_\ell}{a_h}, \frac{\eta}{\eta-1}\right\}, \tag{C.25}$$

We have assumed from the outset that $\lambda > a_h/a_\ell$, so $\mu_{\ell h} > 1$ and this market structure is

A.30

indeed consistent with firm profit incentives. If instead we had $\lambda < a_h/a_\ell$, then this market structure would not exist and we would be left with only the other five.[11]

C.1.3 Aggregation

Substituting the demand curve, $p_{jt} = (Y_{jt}/Y_t)^{-1/\eta}$, and fact that only one firm produces each distinct product, $Y_{jt} = q_{ijt}Y_{ijt}$, into the definition of the markup (C.17), we have

$$w_{\text{pt}} = \left(\frac{L_{ijt}}{Y_t} \right)^{-1/\eta} (a_i q_{ijt})^{1-1/\eta} \mu_{ijt}^{-1}. \tag{C.26}$$

Because there is only one firm producing each variety j, we can replace a_i with a_{jt} and drop all other i indices (e.g., $q_{ijt} = q_{jt}$). Before aggregating this expression, let us define some useful macroeconomic aggregates. Aggregate quality is defined as an average of blueprint qualities:

$$Q_t \equiv \left(\frac{1}{N_t} \int_0^{N_t} q_{jt}^{\eta-1} dj \right)^{1/(\eta-1)}. \tag{C.27}$$

Aggregate productivity is defined as[12]

$$\bar{a}_t \equiv \left(\frac{1}{N_t} \int_0^{N_t} \left(\frac{q_{jt}}{Q_t} \right)^{\eta-1} a_{jt}^{\eta-1} dj \right)^{1/(\eta-1)}, \tag{C.28}$$

a quality-weighted average of firm-level productivity. Finally, we define two aggregates of markups,

$$\mathcal{M}_{\eta t} \equiv \left(\frac{1}{N_t} \int_0^{N_t} \left(\frac{a_{jt} q_{jt}}{\bar{a}_t Q_t} \right)^{\eta-1} \mu_{jt}^{-\eta} dj \right)^{-1/\eta} \tag{C.29}$$

and

$$\mathcal{M}_{\eta-1,t} \equiv \left(\frac{1}{N_t} \int_0^{N_t} \left(\frac{a_{jt} q_{jt}}{\bar{a}_t Q_t} \right)^{\eta-1} \mu_{jt}^{-(\eta-1)} dj \right)^{-1/(\eta-1)}. \tag{C.30}$$

---

[11]If $\lambda < a_h/a_\ell$, then we could, in principle, have markets in which a high type with a quality *disadvantage* leads a low type because of its productivity gap. However, this would never emerge in equilibrium because entry by R&D and M&A always implies that the new firm replaces the leader, not the follower.

[12]I call this "aggregate productivity" even though it puts weight on both qualities q_j and productivities a_j. The reason is that, if there is no heterogeneity in productivity ($a_h = a_\ell = 1$), then we get $\bar{a} = 1$ for any quality distribution; but if there is no quality heterogeneity ($q_{jt}/Q_t = 1$, $\forall j$), then $\bar{a}_t \in [a_\ell, a_h]$. In other words, $\bar{a}_t$ is principally a measure of how blueprints are distributed across productivity types a_i.

These are averages of markups across products, with higher weight placed on blueprints with higher productivity and quality.

With these objects defined, we can use the expression (C.26) to derive two aggregate equations. First, integrating (C.26) over j and imposing market clearing $L_{\text{p}t} = \int_0^{N_t} L_{jt}dj$, we get

$$w_{\text{p}t} = \left(\frac{L_{\text{p}t}}{N_t Y_t}\right)^{-1/\eta} (\bar{a}_t Q_t)^{1-1/\eta} \mathcal{M}_{\eta,t}^{-1}. \tag{C.31}$$

Second, solving for Y_{jt} in (C.26) and substituting it into the definition of total output Y_t (6), we get

$$w_{\text{p}t} = N_t^{1/(\eta-1)} \bar{a}_t Q_t \mathcal{M}_{\eta-1,t}^{-1}. \tag{C.32}$$

Now equating the right-hand sides of these expressions and solving for output, we have

$$Y_t = N_t^{1/(\eta-1)} \bar{a}_t Q_t \Omega_t L_{\text{p}t}, \tag{C.33}$$

where the wedge term Ω_t is defined

$$\Omega_t \equiv \frac{\mathcal{M}_{\eta t}^{\eta}}{\mathcal{M}_{\eta-1,t}^{\eta}}. \tag{C.34}$$

As shown in Section C.1.5 below, this wedge takes a value in $(0,1]$ and attains its maximum ($\Omega = 1$) if and only if there is no cross-product markup heterogeneity ($\mu_{jt} = \mu_t$, $\forall j$). This is because heterogeneous market power distorts demand and results in a misallocation of production labor across blueprints, as in Peters (2020). I therefore refer to Ω as *static misallocation*.

To solve for aggregate wages and the labor share of output, note that combining the expression (C.33) with either condition (C.31) or condition (C.32) implies

$$w_{\text{p}t} = \Lambda_t \frac{Y_t}{L_{\text{p}t}}, \tag{C.35}$$

where the (production) labor share of output $\Lambda_t = w_{\text{p}t} L_{\text{p}t}/Y_t$ has solution

$$\Lambda_t \equiv \frac{\mathcal{M}_{\eta-1,t}^{\eta-1}}{\mathcal{M}_{\eta t}^{\eta}}. \tag{C.36}$$

Notice that, from the household demand curve, the sales share of good j in total sales is

$$\frac{p_{jt}Y_{jt}}{Y_t} = p_{jt}^{1-\eta} = \left(\frac{\mu_{jt}w_{\mathrm{pt}}}{a_i q_{jt}}\right)^{1-\eta}, \tag{C.37}$$

implying that

$$\Lambda_t = \int_0^{N_t} \frac{p_{jt}Y_{jt}}{Y_t}\mu_{jt}^{-1}dj \tag{C.38}$$

The aggregate labor share is a sales-weighted average of product-level labor shares (inverse markups μ_{jt}^{-1}). A natural definition of the aggregate markup is therefore Λ_t^{-1}.

C.1.4 Firm labor allocations, revenues, and profits

Let us bring back the firm-product-level indices $\{i,j\}$. The conditions (C.26) and (C.31) together imply labor allocations

$$\frac{L_{ijt}}{L_{\mathrm{pt}}} = \frac{1}{N_t}\left(\frac{a_i q_{ijt}}{\bar{a}_t Q_t}\right)^{\eta-1}\left(\frac{\mu_{ijt}}{\mathcal{M}_{\eta t}}\right)^{-\eta}. \tag{C.39}$$

A product receives a higher labor allocation if it has higher quality, its producer is more productive, or its markup is lower.[13] Now using the fact that revenue $p_{ijt}Y_{ijt} = \mu_{ijt}w_{\mathrm{pt}}L_{ijt}$ and the wage $w_{\mathrm{pt}} = \Lambda_t Y_t / L_{\mathrm{pt}}$, it follows that the firm receives revenues

$$p_{ijt}Y_{ijt} = \mu_{ijt}\left(\frac{L_{ijt}}{L_{\mathrm{pt}}}\right)\Lambda_t Y_t \tag{C.40}$$

$$= \frac{1}{N_t}\left(\frac{a_i q_{ijt}}{\bar{a}_t Q_t}\right)^{\eta-1}\left(\frac{\mu_{ijt}}{\mathcal{M}_{\eta-1,t}}\right)^{-(\eta-1)}Y_t \tag{C.41}$$

Fixing relative allocations, the scale of production grows with aggregate output per product Y. The corresponding profits are then

$$\Pi_{ijt} = \left(1 - \frac{1}{\mu_{ijt}}\right)p_{ijt}Y_{ijt}. \tag{C.42}$$

---

[13]The reason (relative) markups reduce labor allocations is the household demand curve: when firms charge a higher price (markup), households demand less of their output.

C.1.5 Quantification of static misallocation

Before proceeding, we must verify the claim that the wedge term Ω_t, defined in (C.34), is indeed in $(0, 1]$ and equals one if and only if the labor allocations L_{ijt}/L_{pt} maximize aggregate TFP (output per worker $A_t = Y_t/L_{\mathrm{pt}}$). The maximum level of TFP can be written

$$A_t^* \equiv \max_{\{L_{ijt}/L_{\mathrm{pt}}\}} \frac{Y_t}{L_{\mathrm{pt}}} = \max_{\{\hat{L}_{ijt}\}} \left(\int_0^{N_t} \left(\sum_{i=1}^{m_{jt}} a_i q_{ijt} \hat{L}_{ijt} \right)^{1-1/\eta} dj \right)^{1/(1-1/\eta)}, \tag{C.43}$$

where $\hat{L}_{ijt} \equiv L_{ijt}/L_{\mathrm{pt}}$ denotes the share of production labor. Within variety j, it is clearly the case that all labor should be allocated to the blueprint with the highest value of $a_i q_{ijt}$, just as occurs the decentralized equilibrium. Letting $\{a_{jt}, q_{jt}\}$ denote the values for this blueprint, this means that

$$A_t^* = \max_{\{\hat{L}_{jt}\}} \left(\int_0^{N_t} (a_{jt} q_{jt} \hat{L}_{jt})^{1-1/\eta} dj \right)^{1/(1-1/\eta)}, \tag{C.44}$$

subject to the constraint that

$$1 \geq \int_0^{N_t} \hat{L}_{jt} dj. \tag{C.45}$$

Letting λ_{Lt} denote the Lagrange multiplier on this constraint, the first-order condition for this maximization is

$$Y_t^{1/\eta} (q_{jt} a_{jt})^{1-1/\eta} (\hat{L}_{jt}^*)^{-1/\eta} = \lambda_{Lt}. \tag{C.46}$$

Aggregating this in the same was as we did the firms' first-order conditions, we have

$$Y_t^{1/\eta} N_t^{1/\eta} (\bar{a}_t Q_t)^{1-1/\eta} = \lambda_{Lt}. \tag{C.47}$$

where $\bar{a}_t$ and Q_t are identical to their values in the decentralized equilibrium. Dividing these two conditions implies

$$\hat{L}_{jt}^* = \frac{1}{N_t} \left(\frac{a_{jt} q_{jt}}{\bar{a}_t Q_t} \right)^{\eta-1}. \tag{C.48}$$

Substituting this solution back into the demand aggregator, we get

$$A_t^* = N_t^{1/(\eta-1)} Q_t \bar{a}_t. \tag{C.49}$$

Because $\{N_t, Q_t, \bar{a}_t\}$ are the same as in the decentralized equilibrium, the ratio of actual (decentralized) aggregate TFP to this maximal TFP is therefore

$$\frac{A_t}{A_t^*} = \Omega_t. \tag{C.50}$$

We will necessarily have $\Omega_t \in (0, 1]$, because $A_t^* \geq A_t > 0$ by definition.

To understand what Ω_t represents, note that $\Omega_t = 1$ if and only if there is zero markup dispersion ($\mu_{jt} = \mu_t$, $\forall j$). This can be seen by noticing that the ratio of the actual labor allocation (C.39) to the maximizing allocation (C.48) equals

$$\frac{\hat{L}_{jt}}{\hat{L}_{jt}^*} = \left(\frac{\mu_{jt}}{\mathcal{M}_{\eta t}} \right)^{-\eta}. \tag{C.51}$$

This ratio equals one for all j if and only if $\mu_{jt} = \mathcal{M}_{\eta t}$ for all j — that is, if and only if all products earn the same markup.

C.2   Dynamic firm problem

This appendix derives the equilibrium conditions for the dynamic firm and household problems, taking as given the path of the blueprint distribution $f_{\mathcal{Z}t}$ and the ratio L_t/N_t. In what follows, statements that an equilibrium object is a "function of time" are more precisely interpreted as those objects depending on the full distribution $f_{\mathcal{Z}t}$ and the ratio L_t/N_t, which themselves will be deterministic functions of time.

C.2.1   Rescaling the firm problem

Firm i maximizes the present value V_{it} of its cash flows D_{it}, as defined in (9). Because the profit maximization at the blueprint level is static (see Section C.1.4), this objective amounts to choosing R&D and M&A policies given the optimal profits (C.42) of each product.

To solve this problem, it is helpful to first establish some facts about the state-price density ξ_t which prices firms.[14] The absence of arbitrage opportunities implies that the riskfree rate r_{ft} equals

$$r_{ft} = -\mathbb{E}_t \left[\frac{d\xi_t}{\xi_t} \right] \frac{1}{dt}. \tag{C.52}$$

and the expected return on any firm i equals

$$r_{it} \equiv \mathbb{E}_t[dR_{it}] \frac{1}{dt} = r_{ft} - \mathbb{E}_t \left[\frac{d[\xi, V_i]_t}{\xi_t V_{it^-}} \right] \frac{1}{dt}. \tag{C.53}$$

Because we have assumed that markets are essentially complete, we will ultimately find that (i) ξ_t is unique and given by the marginal utility of a representative household, and (ii) all expected returns r_{it} equal the riskfree rate r_{ft}. But these intuitive results must be proven; for now, we work with the more general expected returns r_{it}.

We can rescale the firm's problem to de-trend its growth relative to the aggregate. Everything will be rescaled by per-capita aggregate output $y_t = Y_t/L_t$. Denote rescaled variables with hats: $\hat{Z}_t \equiv Z_t/y_t$. The rescaled dividend $\hat{D}_{it}$ equals rescaled profits,

$$\hat{\Pi}_{it} = \sum_{j=1}^{n_{it}} \hat{\Pi}_{ijt} = \sum_{j=1}^{n_{it}} (1 - \mu_{ijt}^{-1}) \frac{L_t}{N_t} \left(\frac{a_i q_{ijt}}{\bar{a}_t Q_t} \right)^{\eta-1} \left(\frac{\mu_{ijt}}{\mathcal{M}_{\eta-1,t}} \right)^{-(\eta-1)}, \tag{C.54}$$

minus rescaled R&D and M&A expenditures,

$$\hat{w}_{\mathrm{x}t} X_{it} = \hat{w}_{\mathrm{x}t} (X_{\mathrm{n}it} + X_{\mathrm{q}it} + X_{\mathrm{m}it}). \tag{C.55}$$

The reason for scaling by per-capita output, as opposed to total output, is that, in a balanced-growth equilibrium, the number of firms M_t and blueprints N_t will grow with the population, so the distribution of rescaled firm sizes (e.g., of $\hat{D}_{it}$) will be stationary. The choice to rescale the skilled wage in (C.55), instead of the number of laborers X_{it}, has the same rationale.

It follows that the rescaled value of the firm equals

$$\hat{V}_{it} = \sup_{\{X_{\mathrm{n}it}, X_{\mathrm{q}it}, X_{\mathrm{m}it}\}} \mathbb{E}_t \left[\int_0^\infty \frac{\xi_{t+\tau}}{\xi_t} \frac{y_{t+\tau}}{y_t} \hat{D}_{i,t+\tau} d\tau \right]. \tag{C.56}$$

---

[14]At this point, $\{\xi_t\}$ may or may not be unique; but market completeness will imply that it is.

Let us derive an HJB equation for this value and express it in terms of scaled variables. First, note that the unscaled firm value can be written recursively as

$$V_{it} = \mathbb{E}_t \left[\int_t^T \frac{\xi_s}{\xi_t} D_{is} ds \right] + \mathbb{E}_t \left[\frac{\xi_T}{\xi_t} V_{iT} \right], \tag{C.57}$$

for any $T > t$. Next, multiplying both sides of this expression by ξ_t and adding $\int_{t_0}^t \xi_s D_{is} ds$, where $t_0 < t$, implies:

$$\int_{t_0}^t \xi_s D_{is} ds + \xi_t V_{it} = \mathbb{E}_t \left[\int_{t_0}^T \xi_s D_{is} ds + \xi_T V_{iT} \right]. \tag{C.58}$$

Hence, the expression on the left-hand size is a martingale, and so it is also specifically a local martingale:

$$\mathbb{E}_t \left[d \left(\int_{t_0}^t \xi_s D_{is} ds + \xi_t V_{it} \right) \right] = 0. \tag{C.59}$$

Evaluating this differential and using the expected return definition (C.53) above, we have the unscaled HJB

$$r_{it} V_{it} = D_{it} + \mathbb{E}_t[dV_{it}] \frac{1}{dt}. \tag{C.60}$$

In order to rescale this, note that aggregate output growth, $\dot{y}_t / y_t = g_{yt}$, is deterministic, so the covariation $d[y, \hat{V}_i]_t = 0$; thus, we have the decomposition

$$dV_{it} = d(y_t \hat{V}_{it}) = g_{yt} y_t \hat{V}_{it} dt + y_t d\hat{V}_{it}. \tag{C.61}$$

Substituting in this decomposition and dividing everything by y_t, we get the rescaled HJB

$$(r_{it} - g_{yt}) \hat{V}_{it} = \sup_{\{X_{\mathrm{n}it}, X_{\mathrm{q}it}, X_{\mathrm{m}it}\}} \left\{ \hat{D}_{it} + \mathbb{E}_t[d\hat{V}_{it}] \frac{1}{dt} \right\}. \tag{C.62}$$

Henceforth, let $\rho_{it} \equiv r_{it} - g_{yt}$, foreseeing that in equilibrium ρ_{it} will equal the constant preference parameter ρ.

C.2.2 Solving the firm problem

To solve the firm's problem (C.62), I first note two facts that help to dramatically simplify the problem. First, note that the rescaled profits of the firm are additively separable across products, and each product's profits depends only on (i) deterministic macroeconomic aggregates (i.e., time) and (ii) the set of blueprint-specific characteristic $Z_{ij} = \{i_1, i_2, \hat{q}\}$, where $\hat{q} \equiv \log(q/Q)$ represents *relative quality*. Letting $\mathcal{Z}_{it} \equiv \{Z_{ijt}\}_{j=1}^{n_{it}}$ denote the collection of blueprint characteristics, firm profits can be written

$$\hat{\Pi}_{it} = \sum_{j=1}^{n_{it}} \hat{\Pi}_{ijt} \equiv \sum_{\{i_1, i_2, \hat{q}\} \in \mathcal{Z}_{it}} \pi_{i_1 i_2 \hat{q} t}. \tag{C.63}$$

The significance of this is that the profits $\pi_{i_1 i_2 \hat{q} t}$ are independent of each other and of firm characteristics (except i_1). The second useful fact, which will be proven below, is that the optimal R&D and M&A policies $\{X_{\mathrm{n}it}, X_{\mathrm{q}it}, X_{\mathrm{m}it}\}$ will be linear in the number of blueprints n and, given n, will depend only on the firm's type $a_i = a_{i_1}$. Hence, we can define the allocations per blueprint as $x \equiv X/n$ and note that the R&D and M&A production functions can be written as linear functions of n. For example, for new-variety creation, we will have $x_{\mathrm{n}it} \equiv X_{\mathrm{n}it}/n_{it}$ and

$$(1 - \alpha)\varphi X_{\mathrm{n}it}^{\varepsilon} n^{1-\varepsilon} = \Phi_{\mathrm{n}}(x_{\mathrm{n}it})n_{it}, \quad \text{where} \quad \Phi_{\mathrm{n}}(x) \equiv (1 - \alpha)\varphi x^{\varepsilon}. \tag{C.64}$$

Define $\{x_{\mathrm{q}it}, x_{\mathrm{m}it}\}$ and $\{\Phi_{\mathrm{q}}(x), \Psi(x)\}$ analogously for quality innovation and M&A search. A consequence of linearity in n is that total expenditures are linear in n:

$$\hat{w}_{\mathrm{x}t} X_{it} = \hat{w}_{\mathrm{x}t} x_{it} n_{it}. \tag{C.65}$$

Combining these two facts implies that the rescaled dividend can be written

$$\hat{D}_{it} = \sum_{\{i_1, i_2, \hat{q}\} \in \mathcal{Z}_{it}} \left(\pi_{i_1 i_2 \hat{q} t} - \hat{w}_{\mathrm{x}t} x_{i_1 t} \right). \tag{C.66}$$

A.38

Like firm profits, the total firm dividend is separable across blueprint characteristics. Thus, a reasonable conjecture is that the value of the firm is also separable across blueprints:

$$\hat{V}_{it} = \sum_{\{i_1, i_2, \hat{q}\} \in \mathcal{Z}_{it}} v_{i_1 i_2 \hat{q} t}, \tag{C.67}$$

where $v_{i_1 i_2 \hat{q} t}$ represents the value of a single blueprint with a leader of type a_{i_1}, a follower of type a_{i_2}, and relative quality $\hat{q}$.

The separability of the value function (C.67) means that, instead of solving a system of HJB equations across firms, we can instead solve a system of HJBs across blueprint types $\{i_1, i_2, \hat{q}\}$. Let us derive this from the firm-level HJBs (C.62). Growth in the (rescaled) firm value equals the sum of six components:

$$\mathbb{E}_t[d\hat{V}_{it}]\frac{1}{dt} = \underbrace{\Phi_{\mathrm{n}}(x_{\mathrm{n}it})n_{it}\bar{v}_{\mathrm{n}it}}_{\text{new variety}} + \underbrace{\Phi_{\mathrm{q}}(x_{\mathrm{q}it})n_{it}\bar{v}_{\mathrm{q}it}}_{\text{quality innov.}} + \underbrace{\Psi(x_{\mathrm{m}it})n_{it}\bar{v}_{\mathrm{m}it}}_{\text{acquisition}}$$

$$- \underbrace{\sum_{j=1}^{n_{it}} \delta_{ijt}\left(\hat{V}_{it} - \hat{V}_{it}^{(-j)}\right)}_{\text{creative destruction}} + \underbrace{\sum_{j=1}^{n_{it}} \delta_{ijt}^{\psi}\left(\bar{v}_{\mathrm{m}ijt} - \left(\hat{V}_{it} - \hat{V}_{it}^{(-j)}\right)\right)}_{\text{M\&A exit}} - \underbrace{\sum_{j=1}^{n_{it}} \frac{\partial \hat{V}_{it}}{\partial \hat{q}_{ijt}} g_{Qt}}_{\text{quality depr.}} + \underbrace{\frac{\partial \hat{V}_{it}}{\partial t}}_{\text{time path}}. \tag{C.68}$$

The first is the expected increase in firm value from new-variety creation. If the firm succeeds, which occurs with intensity $\Phi_{\mathrm{q}}(x_{\mathrm{q}it})n_{it}dt$, then it expects its market value to increase by $\bar{v}_{\mathrm{n}it}$. The second terms is analogous for quality innovation. The third term is the expected increase in value from finding and acquiring an M&A target. If the firm successfully finds a blueprint, then $\bar{v}_{\mathrm{m}it}$ is the surplus it expects: the value of the acquired blueprint net of the deal price it must pay. Note that the expectation $\bar{v}_{\mathrm{m}it}$ includes the set of blueprints that do not result in a deal, for which the surplus is zero. The fourth term is the expected loss in value from creative destruction. If a competitor improves on the quality of blueprint j, then the firm's value falls to some $\hat{V}_{it}^{(-j)} < \hat{V}_{it}$; each δ_{ijt} denotes the Poisson intensity with which a blueprint could be lost to creative destruction. The fifth term is the expected gain in firm value from selling a blueprint to an acquirer at a premium: the firm loses the blueprint, but receives a payout $\bar{v}_{\mathrm{m}ijt}$ determined in the Nash bargaining game. This occurs with intensity δ_{ijt}^{ψ}. The sixth and final term is the depreciation in the relative quality of existing blueprints.

Aggregate quality grows at a rate $\dot{Q}_t/Q_t = g_{Qt}$, so the relative quality of existing blueprints depreciates by $\dot{\hat{q}}_{ijt} = -g_{Qt}$.

Substituting (C.68) into the firm-level HJB (C.62) implies the first-order condition for new-variety R&D

$$\hat{w}_{\mathrm{x}t} = \varepsilon(1 - \alpha)\varphi \left(\frac{X_{\mathrm{n}it}}{n_{it}} \right)^{\varepsilon-1} \bar{v}_{\mathrm{n}it}. \tag{C.69}$$

Analogous first-order conditions hold for quality R&D and M&A inputs. These conditions imply the optimal policies

$$x_{\mathrm{n}it} \equiv \frac{X_{\mathrm{n}it}}{n_{it}} = \left(\frac{\varepsilon(1 - \alpha)\varphi\bar{v}_{\mathrm{n}it}}{\hat{w}_{\mathrm{x}t}} \right)^{1/(1-\varepsilon)}, \tag{C.70}$$

$$x_{\mathrm{q}it} \equiv \frac{X_{\mathrm{q}it}}{n_{it}} = \left(\frac{\varepsilon\alpha\varphi\bar{v}_{\mathrm{q}it}}{\hat{w}_{\mathrm{x}t}} \right)^{1/(1-\varepsilon)}, \tag{C.71}$$

$$x_{\mathrm{m}it} \equiv \frac{X_{\mathrm{m}it}}{n_{it}} = \left(\frac{\varepsilon\psi\bar{v}_{\mathrm{m}it}}{\hat{w}_{\mathrm{x}t}} \right)^{1/(1-\varepsilon)}. \tag{C.72}$$

Provided that the expected market values $\{\bar{v}_{\mathrm{n}it}, \bar{v}_{\mathrm{q}it}, \bar{v}_{\mathrm{m}it}\}$ are invariant to size n_{it}, these are independent of n_{it}, as conjectured.

With these results in hand, we can implement and prove the conjecture (C.67). Recall that, conditional on creating a new variety, the blueprint quality is drawn from the distribution $\Gamma_{\hat{q}}$. Thus, under (C.67), the expected value of a new variety is

$$\bar{v}_{\mathrm{n}it} = \int_{-\infty}^{\infty} v_{i\varnothing\hat{q}t}\Gamma_{\hat{q}}d\hat{q}. \tag{C.73}$$

Recall also that quality innovations are undirected (i.e., target blueprints are chosen at random). This means that the expected present value a new blueprint through quality innovation is

$$\bar{v}_{\mathrm{q}it} = \int_{-\infty}^{\infty} [v_{i\ell,\hat{q}+\log\lambda,t}f_{\ell\hat{q}t} + v_{ih,\hat{q}+\log\lambda,t}f_{h\hat{q}t}]d\hat{q}, \tag{C.74}$$

where $f_{i\hat{q}t} \equiv \sum_{i_2} f_{ii_2\hat{q}t}$. In words, the firm finds a blueprint with current relative quality $\hat{q}$ and creates a new one with relative quality $\hat{q} + \log\lambda$. Finally, to solve for the expected value of M&A, we need to pin down when a deal will take place and, if it does, the deal price. If a firm with type i purchases a blueprint of type $\{i_1 = i', i_2, \hat{q}\}$, then the Nash bargaining

solution (should it exist) is

$$v_{\mathrm{m},ii'i_2\hat{q}t} = \underset{v_\mathrm{m}}{\mathrm{argmax}} \left\{ (v_{ii_2\hat{q}t} - v_\mathrm{m})^\varrho (v_\mathrm{m} - v_{i'i_2\hat{q}t})^{1-\varrho} \right\} = \varrho v_{i'i_2\hat{q}t} + (1-\varrho) v_{ii_2\hat{q}t}. \tag{C.75}$$

The deal is clearly mutually beneficial if and only if $v_{ii_2\hat{q}t} > v_{i'i_2\hat{q}t}$ (if they are equal, no deal occurs, consistent with an infinitesimal fixed cost of negotiating). This will ultimately be the case if and only if there is a high-type buyer ($i = h$) and a low-type target ($i' = \ell$), because high-productivity firms obtain strictly higher profits ($\pi_{hi_2\hat{q}t} > \pi_{\ell i_2\hat{q}t}$) and will face the same risk of creative destruction ($\delta_{hi_2\hat{q}t} = \delta_{\ell i_2\hat{q}t}$, shown below).[15] Consequently, the expected value of a new blueprint for the acquirer (net of deal price) is

$$\bar{v}_{\mathrm{m}it} = \varrho \int_{-\infty}^{\infty} \sum_{i' \in \{h,\ell\}} \sum_{i_2 \in \{h,\ell,\varnothing\}} \mathbb{I}_{v_{ii_2\hat{q}t} > v_{i'i_2\hat{q}t}} (v_{ii_2\hat{q}t} - v_{i'i_2\hat{q}t}) f_{i'i_2\hat{q}t} d\hat{q} \tag{C.76}$$

$$= \mathbb{I}_{i=h} \varrho f_{\ell t} \int_{-\infty}^{\infty} \sum_{i_2 \in \{h,\ell,\varnothing\}} (v_{hi_2\hat{q}t} - v_{\ell i_2\hat{q}t}) \frac{f_{\ell i_2\hat{q}t}}{f_{\ell t}} d\hat{q}. \tag{C.77}$$

There is zero expected value to M&A if the acquirer is low-productivity, because no deal will ever go through. This implies $X_{\mathrm{m}\ell t} = 0$. For a high-productivity firm, a deal will only take place if M&A search identifies a low-productivity blueprint (hence the multiplication by $f_{\ell t} = N_{\ell t}/N_t$); conditional on finding such a target, the acquirer gets $1 - \varrho$ of the (expected) surplus.

The blueprint-loss intensities $\delta_{ijt} = \delta_{i_1 i_2 \hat{q}t}$ in (C.68) are pinned down by the fact that the flows of blueprints out of existing markets via quality innovation (creative destruction) must equal flows in:

$$\delta_{i_1 i_2 \hat{q}t} N_{i_1 i_2 \hat{q}t} = \sum_{i \in \{h,\ell\}} \Big(\underbrace{\Phi_\mathrm{q}(x_{\mathrm{q}it}) N_{it} f_{i_1 i_2 \hat{q}t}}_{\text{incumbents}} + \underbrace{\Phi_\mathrm{q}(x_{\mathrm{Eq}it}) \omega_i \bar{s} L_t f_{i_1 i_2 \hat{q}t}}_{\text{new firms}} \Big), \tag{C.78}$$

where the policies $x_{\mathrm{Eq}it}$ are the amount of labor each type-a_i entrepreneur dedicates to quality innovation. In the next section, we will prove that these policies are identical to the incumbent firm policies $x_{\mathrm{q}it}$. Both terms on the right-hand side are multiplied by the densities $f_{i_1 i_2 \hat{q}t} = N_{i_1 i_2 \hat{q}t}/N_t$ because entry is random across blueprints and therefore proportional to

---

[15]The formal proof of this is stated in the proof of Proposition 1 in Appendix C.6

the number of existing blueprints with those characteristics. It follows that the solution is

$$\delta_{i_1 i_2 \hat{q} t} = \delta_t = \sum_{i \in \{h, \ell\}} \left(f_{it} + \omega_i \frac{\bar{s} L_t}{N_t} \right) \Phi_{\mathrm{q}i}(x_{\mathrm{q}it}), \tag{C.79}$$

which is independent of blueprint characteristics. Only incumbents can do M&A, and only high-productivity incumbents will choose to do so, so similar logic implies the intensity of M&A exit satisfies

$$\delta^{\psi}_{i_1 i_2 \hat{q} t} N_{i_1 i_2 \hat{q} t} = \mathbb{I}_{i_1 = \ell} \Psi(x_{\mathrm{m}ht}) N_{ht} f_{i_1 i_2 \hat{q} t}, \tag{C.80}$$

yielding the solution

$$\delta^{\psi}_{i_1 i_2 \hat{q} t} = \delta^{\psi}_{i_1 t} = \mathbb{I}_{i_1 = \ell} \Psi(x_{\mathrm{m}ht}) f_h, \tag{C.81}$$

which depends on the producer's type (only low types will choose to sell blueprints), but is invariant to the other blueprint characteristics i_2 and $\hat{q}$.

Using these intensities along with the values (C.73) through (C.77), the HJB (C.62) with (C.68) can be separated into a system of n_{it} HJB equations (with $i = i_1$) of the form

$$\underbrace{\rho_{it} v_{ii_2 \hat{q} t}}_{\text{discount}} = \underbrace{\pi_{ii_2 \hat{q} t} - \hat{w}_{\mathrm{x}t} x_{it}}_{\text{cash flow}} + \underbrace{\Phi_{\mathrm{n}}(x_{\mathrm{n}it}) \bar{v}_{\mathrm{n}it}}_{\text{new variety}} + \underbrace{\Phi_{\mathrm{q}}(x_{\mathrm{q}it}) \bar{v}_{\mathrm{q}it}}_{\text{quality innov.}} + \underbrace{\Psi(x_{\mathrm{m}it}) \bar{v}_{\mathrm{m}it}}_{\text{acquisition}}$$

$$- \underbrace{\delta_t v_{ii_2 \hat{q} t}}_{\text{creative destr.}} + \underbrace{\mathbb{I}_{i = \ell} \Psi(x_{\mathrm{m}ht}) f_{ht} (1 - \varrho)(v_{h i_2 \hat{q} t} - v_{\ell i_2 \hat{q} t})}_{\text{M\&A exit}} - \underbrace{\frac{\partial v_{ii_2 \hat{q} t}}{\partial \hat{q}} g_{Qt}}_{\text{quality depr.}} + \underbrace{\frac{\partial v_{ii_2 \hat{q} t}}{\partial t}}_{\text{time path}} . \tag{C.82}$$

Fixing time t, the whole system of HJBs for blueprint values therefore constitutes an exactly identified system across $\{i_1, i_2, \hat{q}\} \in \{h, \ell\} \times \{h, \ell, \varnothing\} \times \mathbb{R}$. This cross-section is still infinite-dimensional because the quality distribution spans $\mathbb{R}$. However, we can reduce it to a finite-dimensional system by using the fact that the profit coefficients $\pi_{i_1 i_2 \hat{q} t}$ are multiplicatively separable across market structure and quality:

$$\pi_{i_1 i_2 \hat{q} t} = \pi_{i_1 i_2 t} e^{(\eta - 1)\hat{q}}. \tag{C.83}$$

A.42

Thus, a reasonable conjecture is that the value of a blueprint takes the separable form

$$v_{i_1 i_2 \hat{q} t} = v^{\mathrm{P}}_{i_1 i_2 t} e^{(\eta-1)\hat{q}} + v^{\mathrm{G}}_{i_1 t}, \tag{C.84}$$

where $v^{\mathrm{P}}_{i_1 i_2 t} e^{(\eta-1)\hat{q}}$ is naturally interpreted as the present value of the blueprint's future profits and $v^{\mathrm{G}}_{i_1 t}$ the present value of growth opportunities. Substituting (C.84) into (C.82) and collecting terms in $e^{(\eta-1)\hat{q}}$ implies, for each pair $\{i_1, i_2\}$, the two sets of equations

$$(\rho_{it} + \delta_t + (\eta-1)g_{Qt})v^{\mathrm{P}}_{i_1 i_2 t} = \pi_{i_1 i_2 t} + \mathbb{I}_{i_1 = \ell}\Psi(x_{\mathrm{m}ht})f_{ht}(1-\varrho)(v^{\mathrm{P}}_{h i_2 t} - v^{\mathrm{P}}_{\ell i_2 t}) + \frac{\partial v^{\mathrm{P}}_{i_1 i_2 t}}{\partial t} \tag{C.85}$$

and, using the first-order conditions for R&D and M&A expenditures,

$$(\rho_{it} + \delta_t)v^{\mathrm{G}}_{i_1 t} = (1-\varepsilon)(\Phi_{\mathrm{n}}(x_{\mathrm{n}it})\bar{v}_{\mathrm{n}it} + \Phi_{\mathrm{q}}(x_{\mathrm{q}it})\bar{v}_{\mathrm{q}it} + \Psi(x_{\mathrm{m}it})\bar{v}_{\mathrm{m}it})$$
$$+ \mathbb{I}_{i=\ell}\Psi(x_{\mathrm{m}ht})f_{ht}(1-\varrho)(v^{\mathrm{G}}_{ht} - v^{\mathrm{G}}_{\ell t}) + \frac{\partial v^{\mathrm{G}}_{i_1 t}}{\partial t}. \tag{C.86}$$

The expected blueprint values from R&D and M&A can also be separated as

$$\bar{v}_{\mathrm{n}it} = \bar{v}^{\mathrm{P}}_{\mathrm{n}it} + v^{\mathrm{G}}_{it}, \qquad\qquad \bar{v}^{\mathrm{P}}_{\mathrm{n}it} \equiv v^{\mathrm{P}}_{i\varnothing t}\int_{-\infty}^{\infty} e^{(\eta-1)\hat{q}}\Gamma_{\hat{q}}d\hat{q}, \tag{C.87}$$

$$\bar{v}_{\mathrm{q}it} = \bar{v}^{\mathrm{P}}_{\mathrm{q}it} + v^{\mathrm{G}}_{it}, \qquad\qquad \bar{v}^{\mathrm{P}}_{\mathrm{q}it} \equiv \sum_{i'\in\{h,\ell\}} v^{\mathrm{P}}_{ii't}\lambda^{\eta-1}\int_{-\infty}^{\infty} e^{(\eta-1)\hat{q}}f_{i'\hat{q}t}d\hat{q}, \tag{C.88}$$

$$\bar{v}_{\mathrm{m}it} = \mathbb{I}_{i=h}\varrho f_{\ell t}(\bar{v}^{\mathrm{P}}_{\mathrm{m}ht} - \bar{v}^{\mathrm{P}}_{\mathrm{m}\ell t} + v^{\mathrm{G}}_{ht} - v^{\mathrm{G}}_{\ell t}), \qquad \bar{v}^{\mathrm{P}}_{\mathrm{m}it} \equiv \sum_{i'\in\{h,\ell,\varnothing\}} v^{\mathrm{P}}_{ii't}\int_{-\infty}^{\infty} e^{(\eta-1)\hat{q}}\frac{f_{\ell i'\hat{q}t}}{f_{\ell t}}d\hat{q}. \tag{C.89}$$

At any time t, this reduced system now constitutes eight equations, determining the six product values $\{v^{\mathrm{P}}_{i_1 i_2 t}\}$ and the two growth values $\{v^{\mathrm{G}}_{i_1 t}\}$. The only additional conditions we need in order to solve this are the terminal conditions $\lim_{t\to\infty} v^{\mathrm{P}}_{i_1 i_2 \hat{q} t}$ and $\lim_{t\to\infty} v^{\mathrm{G}}_{i_1 t}$, which we will see below are the solutions to the balanced-growth equilibrium. This system verifies the conjecture.

C.2.3 Aggregate economic growth

We can combine firms' policies with the expression for aggregate output (C.33) to compute aggregate economic growth and its components. The growth rate of aggregate output per capita equals

$$g_{yt} = \frac{1}{\eta - 1} g_{Nt} + g_{Qt} + g_{\bar{a}t} + g_{\Omega t} + g_{l_p t}. \tag{C.90}$$

It will turn out that only the first two terms, the growth rates of the number of varieties ($g_{Nt} = \dot{N}_t/N_t$) and aggregate quality ($g_{Qt} = \dot{Q}_t/Q_t$), are needed to characterize the equilibrium; thus, I defer discussion of the last three terms to Section **??** on growth accounting. Growth in new varieties is given by

$$g_{Nt} = \sum_{i \in \{h,\ell\}} \Phi_n(x_{nit}) \left(f_{it} + \omega_i \bar{s} \frac{L_t}{N_t} \right). \tag{C.91}$$

To compute growth in aggregate quality, note that the change in aggregate quality comes from two sources: new varieties, which have initial quality λQ_t; and quality innovations, which improve the current quality of each target blueprint by a factor of λ. By the chain rule, the change in aggregate quality is

$$dQ_t = \frac{1}{\eta - 1} Q_t^{2-\eta} d \left(\frac{1}{N_t} \int_0^{N_t} q_{jt}^{\eta-1} dj \right), \tag{C.92}$$

where

$$d \left(\frac{1}{N_t} \int_0^{N_t} q_{jt}^{\eta-1} dj \right) = \underbrace{\frac{1}{N_t} Q_t^{\eta-1} \left(\lambda^{\eta-1} - 1 \right) dN_t}_{\text{new varieties}} + \underbrace{\frac{1}{N_t} \int_0^{N_t} d \left(q_{jt}^{\eta-1} \right) dj}_{\text{quality innov.}}. \tag{C.93}$$

The second term can be simplified using a law of large numbers: with intensity $\delta_t dt$, an individual blueprint's quality jumps ($d(q_{jt}^{\eta-1}) = (\lambda^{\eta-1} - 1)q_{jt}^{\eta-1}$), so we have

$$\frac{1}{N_t} \int_0^{N_t} d \left(q_{jt}^{\eta-1} \right) dj = \delta_t Q_t^{\eta-1} \left(\lambda^{\eta-1} - 1 \right) dt. \tag{C.94}$$

Putting all of this together, we have the growth rate

$$g_{Qt} = \frac{\lambda^{\eta-1} - 1}{\eta - 1}(g_{Nt} + \delta_t). \tag{C.95}$$

C.3 Dynamic household problem

C.3.1 Definition of a family

I formalize the assumption that individuals within a household make decisions as a "family" as follows. A family is a random collection of individuals on a continuum of arbitrary mass $\varpi_t \in (0, L_t]$. Each newborn individual enters an existing family, and every individual is permanently a member of one family. Thus, family size grows with the population, $\dot{\varpi}_t/\varpi_t = g_L$, and the density of families L_t/ϖ_t has a constant mass.

Let us index the individuals within a family by k. At any time t, we can partition the family into those individuals who are currently alive ($k \in [0, \varpi_t]$) and those that will be born in a future period ($k \in (\varpi_t, \infty)$). The flow utility of the family is the average utility per living member at time t:

$$u_t \equiv \frac{1}{\varpi_t} \int_0^{\varpi_t} u(c_{kt}, l_{kt})dk, \tag{C.96}$$

where the individual utility functions are given by (5). The family is a dynasty, valuing the utility of future generations and discounting it at the same rate ρ with which they discount their own. Hence, the family chooses individual consumption and labor policies $\{c_{kt}, l_{pkt}, l_{xkt}, x_{\text{En}kt}, x_{\text{Eq}kt}\}$ and portfolio policies $\{\alpha_t\}$ (defined below) for all times t and all individuals k (born and unborn) to maximize the total expected family utility

$$U_t \equiv \max_{\{c_{k,t+\tau}, l_{k,t+\tau}, x_{k,t+\tau}\}} \left\{ \int_0^\infty e^{-\rho\tau} u_{t+\tau} d\tau \right\}, \tag{C.97}$$

where, again, family utility is defined by (C.96).

A consequence of this setting is that, for any $\varpi_t \in (0, L_t]$, each family will represent a fully diversified mass of individuals in which the proportions of each labor and entrepreneurial type equal the overall population proportions ($\bar{s}$ and ω). All families will therefore be symmetric and the equilibrium decisions at the individual level will be the same for all ϖ_t. Without

loss of generality, then, we can set $\varpi_t = L_t$ and solve the problem of a representative family composed of all individuals in the economy.

C.3.2 Optimal household policies

Even in the presence of yet-nonexistent cohorts in the household's maximization, the derivation of the household's HJB equation is standard. Under $\varpi_t = L_t$, the utility maximand (C.97) can be written recursively as

$$U_t = \int_t^T e^{-\rho(s-t)} u_s ds + e^{-\rho(T-t)} U_T, \tag{C.98}$$

for any $T \geq t$ (I suppress the maximization for convenience). Now, multiplying this equation by $e^{-\rho t}$ and adding $\int_{t_0}^t e^{-\rho s} u_s ds$ with $t_0 < t$ to both sides, we get

$$\int_{t_0}^t e^{-\rho s} u_s ds + e^{-\rho t} U_t = \int_{t_0}^T e^{-\rho s} u_s ds + e^{-\rho T} U_T. \tag{C.99}$$

This equation means the expression on the left-hand side is a martingale; therefore, it is also a local martingale, satisfying the differential equation

$$\frac{d}{dt} \left(\int_{-\infty}^t e^{-\rho s} u_s ds + e^{-\rho t} U_t \right) = 0. \tag{C.100}$$

Evaluating this expression and bringing back the maximization implies the HJB equation

$$\rho U_t = \sup_{\alpha_t, \{c_{kt}, l_{kt}, x_{kt}\}} \left\{ u_t + \dot{U}_t \right\}. \tag{C.101}$$

To solve this, we must express U_t as a function of the state variables of the household.

The first state variable of the household is its total financial wealth W_t. The household accumulates wealth through three endogenous sources of income,[16]

$$w_{\mathrm{p}t} L_{\mathrm{p}t} + w_{\mathrm{x}t} L_{\mathrm{x}t} = \int_0^{L_t} (w_{\mathrm{p}t} l_{\mathrm{p}kt} + w_{\mathrm{x}t}(l_{\mathrm{x}kt} - x_{\mathrm{En}kt} - x_{\mathrm{Eq}kt})) dk. \tag{C.102}$$

---

[16]Note that $l_{\mathrm{x}t}$ represents hours supplied to both incumbent firms and private entrepreneurial efforts, so we must subtract out entrepreneurial labor hours $x_{\mathrm{En}kt}$ and $x_{\mathrm{Eq}kt}$.

The second source is new-firm creation from entrepreneurs. Let $dJ_{\mathrm{n}kt} \in \{0,1\}$ denote a Poisson jump that equals one if individual k succeeds in starting a new firm via new-variety creation. The assumption that entrepreneurs have the same R&D technologies as single-product firms means

$$\mathbb{E}_t dJ_{\mathrm{n}kt} = \Phi_{\mathrm{n}}(x_{\mathrm{En}kt})dt = (1-\alpha)\varphi x_{\mathrm{En}kt}^{\varepsilon}dt. \tag{C.103}$$

Define $dJ_{\mathrm{q}kt}$ analogously for a new firm formed via quality innovation. Let $v_{\mathrm{En}kt}$ and $v_{\mathrm{Eq}kt}$ denote the market value of a newly created firm. Then the total flow of new-firm value equals

$$V_{\mathrm{E}t} \equiv \frac{1}{dt}\int_0^{L_t}(v_{\mathrm{En}kt}dJ_{\mathrm{n}kt} + v_{\mathrm{Eq}kt}dJ_{\mathrm{q}kt})dk \tag{C.104}$$

The fact that there is a continuum of entrepreneurs means two things. First, policies will be symmetric across technologies and types (high- versus low-productivity a_i), so there are only four unique policies $\{x_{\mathrm{En}ht}, x_{\mathrm{Eq}ht}, x_{\mathrm{En}\ell t}, x_{\mathrm{Eq}\ell t}\}$. Second, we can apply a law of large numbers to express this value as a deterministic function of policies:

$$V_{\mathrm{E}t} = \bar{s}L_t \sum_{i\in\{h,\ell\}} \omega_i(\Phi_{\mathrm{n}}(x_{\mathrm{En}it})\bar{v}_{\mathrm{En}it} + \Phi_{\mathrm{q}}(x_{\mathrm{Eq}it})\bar{v}_{\mathrm{Eq}it}), \tag{C.105}$$

where $\bar{v}_{\mathrm{En}it}$ and $\bar{v}_{\mathrm{Eq}it}$ are the expected values of a new firm. By the results in the previous section, we will have $\bar{v}_{\mathrm{En}it} = \bar{v}_{\mathrm{n}it}y_t$ and $\bar{v}_{\mathrm{Eq}it} = \bar{v}_{\mathrm{q}it}y_t$. Finally, the household receives capital income from trading in the stocks of incumbent firms and in the riskfree bond with return r_{ft}. Let α_t denote the share of wealth invested in the stock market[17] and

$$r_{Vt} \equiv \frac{D_t}{V_t} + \frac{\dot{V}_t - V_{\mathrm{E}t}}{V_t} \tag{C.106}$$

the return on all *incumbent* firms in the economy.[18] Combining these three sources with consumption expenditures $C_t \equiv \int_0^{L_t} c_{kt}dk$, wealth grows according to

$$\dot{W}_t = W_t(r_{ft} + \alpha_t(r_{Vt} - r_{ft})) + V_{\mathrm{E}t} + w_{\mathrm{p}t}L_{\mathrm{p}t} + w_{\mathrm{x}t}L_{\mathrm{x}t} - C_t. \tag{C.107}$$

---

[17]A more precise setup would be to let the household choose portfolio weight α_{it} in every firm $i \in [0, M_t]$ in the market, then clear the market with $\alpha_{it} = V_{it}/V_t$. This would ultimately yield the same solution.

[18]$\dot{V}_t$ denotes growth inclusive of entry and exit, so subtracting entry $V_{\mathrm{E}t}$ gives incumbent value growth.

As a reality check, note that imposing market clearing for wealth $W_t = V_t$ (and $\alpha_t = 1$) and consumption $C_t = Y_t$ implies

$$\dot{W}_t = (D_t + \dot{V}_t - V_{Et}) + V_{Et} + w_{pt}L_{pt} + w_{xt}L_{xt} - C_t = \dot{V}_t + Y_t - C_t = \dot{V}_t, \qquad \text{(C.108)}$$

consistent with the fact that all tradable wealth in the economy is held in claims to (existing) firms' cash flows (i.e., the stock market).

The second state variable of the household captures expected growth in off-balance-sheet wealth, such as future labor income. We will prove that aggregate output Y_t and time t are sufficient information to summarize this state. Consider, as an example, the present value of aggregate income from production labor:

$$H_{pt} \equiv \int_0^\infty e^{-\int_0^\tau r_{f,t+\tau'}d\tau'} w_{p,t+\tau}L_{p,t+\tau}d\tau. \qquad \text{(C.109)}$$

We showed in Section C.1.3 that total income $w_{pt}L_{pt} = \Lambda_t Y_t$ for labor share Λ_t a deterministic function of time. Thus, the present value of production-labor income can be stated as

$$H_{pt} = Y_t \int_0^\infty e^{-\int_0^\tau (r_{f,t+\tau'}-g_{Y,t+\tau'})d\tau'} \Lambda_{t+\tau}d\tau = Y_t h_p(t), \qquad \text{(C.110)}$$

the product of aggregate output and a deterministic function of time. Similar arguments can be made for human wealth from skilled labor and new-firm creation. Unlike wealth W_t, which has an endogenous accumulation equation, the household takes aggregate output and its growth rate $\dot{Y}_t/Y_t = g_{Yt}$ as given.

To solve the HJB equation (C.101), I first conjecture that U_t is a function of wealth, aggregate output, and time, implying that

$$\rho U_t = \sup_{\alpha_t, \{c_{kt}, l_{kt}, x_{kt}\}} \left\{ u_t + \frac{\partial U_t}{\partial W_t}\dot{W}_t + \frac{\partial U_t}{\partial Y_t}\dot{Y}_t + \frac{\partial U_t}{\partial t} \right\}. \qquad \text{(C.111)}$$

Consistent with the absence of systematic risks, the first-order condition for the allocation α_t implies the return restriction

$$r_{Vt} = r_{ft}, \qquad \text{(C.112)}$$

as otherwise there would be an arbitrage opportunity. The first-order conditions for consumption are

$$\frac{c_{kt}^{-1}}{L_t} = \frac{\partial U_t}{\partial W_t}, \tag{C.113}$$

which implies that all individuals consume the same amount $c_{kt} = c_t = C_t/L_t$. The first-order conditions for production laborers are

$$\frac{\chi l_{pkt}^{1/\zeta}}{L_t} = \frac{\partial U_t}{\partial W_t} w_{\mathrm{p}t}, \tag{C.114}$$

which, similarly, implies that $l_{pkt} = l_{\mathrm{p}t}$. Likewise, for skilled laborers, we have

$$\frac{\chi l_{xkt}^{1/\zeta}}{L_t} = \frac{\partial U_t}{\partial W_t} w_{\mathrm{x}t}, \tag{C.115}$$

so $l_{xkt} = l_{\mathrm{x}t}$. Finally, the first-order conditions for type-a_i entrepreneurial effort into each technology is

$$w_{\mathrm{x}t} = \varepsilon\varphi x_{\mathrm{En}it}^{\varepsilon-1}\bar{v}_{\mathrm{n}it}y_t \quad \text{and} \quad w_{\mathrm{x}t} = \varepsilon\varphi x_{\mathrm{Eq}it}^{\varepsilon-1}\bar{v}_{\mathrm{q}it}y_t. \tag{C.116}$$

Entrepreneurs weigh the expected market value of their investment against the foregone skilled wage they would have received if they had instead worked for an incumbent. Dividing these by y_t implies that the entrepreneurs have exactly the same first-order conditions as incumbents, and therefore the exact same R&D policies conditional on type.

Next, conjecture that the solution to the value function is

$$U_t = k(t) + \rho^{-1}\log\left(\frac{W_t + Y_t h(t)}{L_t}\right) \tag{C.117}$$

for some functions $k(t)$ and $h(t)$ of time, for which we will solve. In line with the intuition above, the function $h(t)$ will be such that $Y_t h(t)$ represents the present value of non-financial wealth. Foreseeing this, let us define total wealth as

$$\overline{W}_t \equiv W_t + Y_t h(t). \tag{C.118}$$

It will prove to be the case that $\overline{W}_t/L_t$ equals the present value of per-capita consump-

tion. Now substitute the conjecture (C.117) into the first-order conditions (C.113) through (C.115). We then have per capita consumption

$$c_t = \rho \frac{\overline{W}_t}{L_t},$$ (C.119)

the standard result for a unit EIS. Combining this with the market-clearing condition $c_t = Y_t/L_t$, the labor supply first-order conditions (C.114) and (C.115) imply

$$l_{\mathrm{p}t} = \left(\chi^{-1} \frac{w_{\mathrm{p}t}}{Y_t/L_t} \right)^{\zeta} \quad \text{and} \quad l_{\mathrm{x}t} = \left(\chi^{-1} \frac{w_{\mathrm{x}t}}{Y_t/L_t} \right)^{\zeta}.$$ (C.120)

These labor supply curves can be combined with the labor demand curves from firms to solve for labor hours and wages. For production labor, we have the aggregate demand curve (C.35) with $L_{\mathrm{p}t} = (1 - \bar{s})l_{\mathrm{p}t}L_t$, implying the solution

$$l_{\mathrm{p}t} = \left(\frac{\Lambda_t}{\chi(1 - \bar{s})l_{\mathrm{p}t}} \right)^{\zeta} = \left(\frac{\Lambda_t}{\chi(1 - \bar{s})} \right)^{\frac{\zeta}{1+\zeta}}.$$ (C.121)

For skilled labor, the policies (C.70) through (C.72) imply total demand from incumbents

$$\int_0^{M_t} X_{it} di = \sum_{i \in \{h,\ell\}} f_{it}(x_{\mathrm{n}it} + x_{\mathrm{q}it} + x_{\mathrm{m}it})N_t = \hat{w}_{\mathrm{x}t}^{-1/(1-\varepsilon)} \kappa_t^{\mathrm{inc}} N_t,$$ (C.122)

where, from (C.70) to (C.72), the object

$$\kappa_t^{\mathrm{inc}} \equiv \sum_{i \in \{h,\ell\}} f_{it}\hat{w}_{\mathrm{x}t}^{1/(1-\varepsilon)}(x_{\mathrm{n}it} + x_{\mathrm{q}it} + x_{\mathrm{m}it})$$ (C.123)

is independent of the wage $\hat{w}_{\mathrm{x}t}$. Recall that labor hours $l_{\mathrm{x}t}$ include hours supplied to both incumbents and private entrepreneurial efforts. Thus, the total skilled labor supply that must meet this demand is total hours $L_{\mathrm{x}t} = \bar{s}l_{\mathrm{x}t}L_t$ minus entrepreneurial hours

$$\int_0^{L_t} (x_{\mathrm{En}kt} + x_{\mathrm{Eq}kt})dk = \sum_{i \in \{h,\ell\}} \omega_i(x_{\mathrm{n}it} + x_{\mathrm{q}it})\bar{s}L_t = \hat{w}_{\mathrm{x}t}^{-1/(1-\varepsilon)} \kappa_t^{\mathrm{ent}} \bar{s}L_t,$$ (C.124)

where, analogous to κ_t^{inc},

$$\kappa_t^{\text{ent}} \equiv \sum_{i\in\{h,\ell\}} \omega_i \hat{w}_{\text{x}t}^{1/(1-\varepsilon)}(x_{\text{n}it} + x_{\text{q}it}). \tag{C.125}$$

Now equate aggregate labor supply with demand,

$$(\chi^{-1}\hat{w}_{\text{x}t})^\zeta \bar{s}L_t - \hat{w}_{\text{x}t}^{-1/(1-\varepsilon)}\kappa_t^{\text{ent}}\bar{s}L_t = \hat{w}_{\text{x}t}^{-1/(1-\varepsilon)}\kappa_t^{\text{inc}}N_t, \tag{C.126}$$

and solve for the wage:

$$\hat{w}_{\text{x}t} = \left(\chi^\zeta\left(\kappa_t^{\text{ent}} + \kappa_t^{\text{inc}}\frac{N_t}{\bar{s}L_t}\right)\right)^{\frac{1-\varepsilon}{1+\zeta(1-\varepsilon)}}. \tag{C.127}$$

Substituting this back into the supply curve gives

$$l_{\text{x}t} = (\chi^{-1}\hat{w}_{\text{x}t})^\zeta. \tag{C.128}$$

To verify the conjecture, substitute the optimal policies above and the conjecture (C.117) into the HJB (C.111). First, note that the utility flow becomes

$$u_t = \log\rho + \log\frac{\overline{W}_t}{L_t} - (1-\bar{s})\chi\frac{l_{\text{p}t}^{1+1/\zeta}}{1+1/\zeta} - \bar{s}\chi\frac{l_{\text{x}t}^{1+1/\zeta}}{1+1/\zeta}, \tag{C.129}$$

where the labor supply terms are functions of time only. Second, note that $W_t = V_t$. Thus, the HJB becomes

$$\rho k(t) = \log\rho - (1-\bar{s})\chi\frac{l_{\text{p}t}^{1+1/\zeta}}{1+1/\zeta} - \bar{s}\chi\frac{l_{\text{x}t}^{1+1/\zeta}}{1+1/\zeta}$$
$$+ \rho^{-1}\frac{V_t}{\overline{W}_t}\frac{\dot{V}_t}{V_t} + \rho^{-1}\frac{Y_t h(t)}{\overline{W}_t}\frac{\dot{Y}_t}{Y_t} + k'(t) + \rho^{-1}\frac{Y_t}{\overline{W}_t}h'(t) \quad \text{(C.130)}$$

Because the ratio V_t/Y_t is a function of time only (proven formally in the next section), so too are the ratios

$$\frac{V_t}{\overline{W}_t} = \frac{V_t/Y_t}{V_t/Y_t + h(t)} \quad \text{and} \quad \frac{Y_t h(t)}{\overline{W}_t} = \frac{h(t)}{V_t/Y_t + h(t)}. \tag{C.131}$$

A.51

This confirms the form of the conjecture (C.117). However, the ordinary differential equation (C.130) only characterizes $k(t)$ in terms of the function $h(t)$. The next section computes $h(t)$ analytically. In all parts of the paper, including the transition path, the specific functional form of $k(t)$ turns out to be irrelevant, because policies and resource allocations do not explicitly depend on $k(t)$. All that matters is that it exists, which (C.130) confirms.

C.3.3 State-price density and asset prices

The state-price density of the household equals

$$\xi_t = e^{-\rho t}\frac{\partial u_t}{\partial c_t} = e^{-\rho t}c_t^{-1}, \tag{C.132}$$

where, again, per-capita consumption c_t equals per-capita output y_t. This implies the law of motion

$$\frac{\dot{\xi}_t}{\xi_t} = -\rho - g_{yt}. \tag{C.133}$$

By the absence of arbitrage, the riskfree rate is equal to

$$r_{ft} = -\frac{\dot{\xi}_t}{\xi_t} = \rho + g_{yt}, \tag{C.134}$$

and all risk premia are equal to zero. Hence, the growth-adjusted discount rate applied to each firm, described in Appendix C.2.1, is

$$\rho_{it} = r_{ft} - g_{yt} = \rho. \tag{C.135}$$

We can use this discount rate to compute $h(t)$ in (C.117). Note that the present value of per-capita consumption equals

$$\int_0^\infty e^{-\int_0^\tau r_{f,t+\tau'}d\tau'}c_{t+\tau}d\tau = c_t\int_0^\infty e^{-\rho\tau}d\tau = \frac{c_t}{\rho} = \frac{\overline{W}_t}{L_t}, \tag{C.136}$$

where the last equation follows from the household's optimal consumption policy. Thus, per-capita total wealth $\overline{W}_t/L_t$ is the present value of consumption. Using the definition of

total wealth and solving for $h(t)$, we get

$$h(t) = \frac{1}{\rho} - \frac{V_t}{Y_t}, \tag{C.137}$$

where

$$\frac{V_t}{Y_t} = \frac{1}{L_t} \int_0^{M_t} \hat{V}_{it} di = \frac{N_t}{L_t} \int_{-\infty}^{\infty} \sum_{i_1 \in \{h,\ell\}} \sum_{i_2 \in \{h,\ell,\varnothing\}} v_{i_1 i_2 \hat{q} t} f_{i_1 i_2 \hat{q} t} d\hat{q}, \tag{C.138}$$

which is indeed a function of time only. Substituting (C.137) and (C.138) back into (C.130) characterizes the solution for $k(t)$, with terminal condition $\lim_{t \to \infty} k(t)$ equal to the constant solution in the balanced-growth equilibrium.

C.4 Distribution

This section determines the law of motion for the distribution $f_{\mathcal{Z}t}$ by deriving the Kolmogorov Forward Equations (KFEs) in terms of the R&D and M&A policies determined in the previous section.

C.4.1 Blueprint distribution

The final step in characterizing the equilibrium is to calculate the evolution of the joint distribution $f_{i_1 i_2 \hat{q} t}$ of leader-follower pairs $\{i_1, i_2\}$ and relative qualities $\hat{q} \equiv \log(q/Q)$. To characterize how the distribution evolves, we derive and solve the Kolmogorov Forward Equations (KFEs). Henceforth, we will use the shorthand notation $\{\Phi_{\mathrm{n}it}, \Phi_{\mathrm{q}it}, \Psi_{it}\}$ to denote the policy-dependent intensities $\{\Phi_{\mathrm{n}}(x_{\mathrm{n}it}), \Phi_{\mathrm{q}}(x_{\mathrm{q}it}), \Psi(x_{\mathrm{m}it})\}$ and will use the notation

$$\nu_{it} \equiv \omega_i \frac{\bar{s} L_t}{N_t} \tag{C.139}$$

to denote the ratio of type-a_i entrepreneurs relative to the blueprint stock N_t.

Note that relative quality $\hat{q}$ depreciates continuously over time. Because $\hat{q}$ is a continuous variable, it is most intuitive to think about how this affects the distribution using a discretized time grid, so relative quality declines in increments $\hat{q}_t - \hat{q}_{t-\Delta t} = -g_{Qt}\Delta t$, and then take the limit as the increments Δt go to zero. Note that the change in blueprint types from solely

A.53

this depreciation is equal to

$$\Delta N_{i_1 i_2 \hat{q} t} = N_{i_1 i_2, \hat{q} + g_{Qt} \Delta t, t - \Delta t} - N_{i_1 i_2 \hat{q}, t - \Delta t}. \tag{C.140}$$

Dividing by $N_{t - \Delta t}$ implies

$$f_{i_1 i_2 \hat{q} t}(1 + g_{Nt} \Delta t) - f_{i_1 i_2 \hat{q}, t - \Delta t} = f_{i_1 i_2, \hat{q} + g_{Qt} \Delta t, t - \Delta t} - f_{i_1 i_2 \hat{q}, t - \Delta t}. \tag{C.141}$$

Dividing by Δt and taking the limit as the time interval becomes small implies

$$\dot{f}_{i_1 i_2 \hat{q} t} + g_{Nt} f_{i_1 i_2 \hat{q} t} = \lim_{\Delta t \to 0} \frac{f_{i_1 i_2, \hat{q} + g_{Qt} \Delta t, t - \Delta t} - f_{i_1 i_2 \hat{q}, t - \Delta t}}{g_{Qt} \Delta t} \frac{g_{Qt} \Delta t}{\Delta t} = \frac{\partial f_{i_1 i_2 \hat{q} t}}{\partial \hat{q}} g_{Qt}. \tag{C.142}$$

This is the standard KFE for a process with only negative drift.

It is straightforward to combine this drift with the jumps across characteristics $\{i_1, i_2\}$. First, consider monopoly products $i_2 = \varnothing$ with leaders $i_1 = i$. The change in the number of blueprints equals

$$\dot{N}_{i \varnothing \hat{q} t} = \frac{\partial f_{i \varnothing \hat{q} t}}{\partial \hat{q}} g_{Qt} N_t + \Phi_{\mathrm{n}it}(N_{it} + \omega_i \bar{s} L_t) \Gamma_{\hat{q}} + \mathbb{I}_{i=h} \Psi_{ht} N_{ht} f_{\ell \varnothing \hat{q} t}$$
$$- \delta_t N_{i \varnothing \hat{q} t} - \mathbb{I}_{i=\ell} \Psi_{ht} f_{ht} N_{i \varnothing \hat{q} t}. \tag{C.143}$$

Dividing by N_t and using the fact that $\dot{f}_{Zt} = \dot{N}_{Zt}/N_t - (\dot{N}_t/N_t)f_{Zt}$ implies the KFE

$$\dot{f}_{i \varnothing \hat{q} t} + g_{Nt} f_{i \varnothing \hat{q} t} = \frac{\partial f_{i \varnothing \hat{q} t}}{\partial \hat{q}} g_{Qt} + \Phi_{\mathrm{n}it}(f_{it} + \nu_{it}) \Gamma_{\hat{q}} + (\mathbb{I}_{i=h} - \mathbb{I}_{i=\ell}) \Psi_{ht} f_{ht} f_{\ell \varnothing \hat{q} t} - \delta_t f_{i \varnothing \hat{q} t}. \tag{C.144}$$

Similarly, for a same-type leader-follower pair $i_1 = i_2 = i$, and letting $-i \equiv \{h, \ell\} \setminus \{i\}$, we have blueprint flows

$$\dot{N}_{ii \hat{q} t} = \frac{\partial f_{ii \hat{q} t}}{\partial \hat{q}} g_{Qt} N_t + \Phi_{\mathrm{q}it}(N_{it} + \omega_i \bar{s} L_t)(f_{ii, \hat{q} - \log \lambda, t} + f_{i, -i, \hat{q} - \log \lambda, t} + f_{i \varnothing, \hat{q} - \log \lambda, t})$$
$$+ \mathbb{I}_{i=h} \Psi_{ht} N_{ht} f_{\ell h \hat{q} t} - \delta_t N_{ii \hat{q} t} - \mathbb{I}_{i=\ell} \Psi_{ht} f_{ht} N_{\ell \ell \hat{q} t}, \tag{C.145}$$

A.54

implying the KFEs

$$\dot{f}_{ii\hat{q}t} + g_{Nt}f_{ii\hat{q}t} = \frac{\partial f_{ii\hat{q}t}}{\partial \hat{q}}g_{Qt} + \Phi_{\mathrm{q}it}(f_{it}+\nu_{it})(f_{ii,\hat{q}-\log\lambda,t} + f_{i,-i,\hat{q}-\log\lambda,t} + f_{i\varnothing,\hat{q}-\log\lambda,t})$$
$$+ \mathbb{I}_{i=h}\Psi_{ht}f_{ht}f_{\ell h\hat{q}t} - \delta_t f_{ii\hat{q}t} - \mathbb{I}_{i=\ell}\Psi_{ht}f_{ht}f_{\ell\ell\hat{q}t}, \quad \text{(C.146)}$$

Finally, for an opposite-type pair $i_1 = i$ and $i_2 = -i$, we have

$$\dot{N}_{i,-i,\hat{q}t} = \frac{\partial f_{i,-i,\hat{q}t}}{\partial \hat{q}}g_{Qt}N_t + \Phi_{\mathrm{q}it}(N_{it}+\omega_i\bar{s}L_t)(f_{-ii,\hat{q}-\log\lambda,t} + f_{-i,-i,\hat{q}-\log\lambda,t} + f_{-i\varnothing,\hat{q}-\log\lambda,t})$$
$$+ \mathbb{I}_{i=h}\Psi_{ht}N_{ht}f_{\ell\ell\hat{q}t} - \delta_t N_{i,-i,\hat{q}t} - \mathbb{I}_{i=\ell}\Psi_{ht}f_{ht}N_{\ell h\hat{q}t}, \quad \text{(C.147)}$$

and thus

$$\dot{f}_{i,-i,\hat{q}t} + g_{Nt}f_{i,-i,\hat{q}t} = \frac{\partial f_{i,-i,\hat{q}t}}{\partial \hat{q}}g_{Qt} + \Phi_{\mathrm{q}it}(f_{it}+\nu_{it})(f_{-ii,\hat{q}-\log\lambda,t} + f_{-i,-i,\hat{q}-\log\lambda,t} + f_{-i\varnothing,\hat{q}-\log\lambda,t})$$
$$+ \mathbb{I}_{i=h}\Psi_{ht}f_{ht}f_{\ell\ell\hat{q}t} - \delta_t f_{i,-i,\hat{q}t} - \mathbb{I}_{i=\ell}\Psi_{ht}f_{ht}f_{\ell h\hat{q}t}. \quad \text{(C.148)}$$

Given R&D and M&A policies, the KFEs (C.144), (C.146), and (C.148) fully characterize the time path of the blueprint distribution.

We can aggregate these KFEs to characterize the evolution of the total blueprint share across firm types f_{it}. Specifically, integrating the KFEs for $i_1 = h$ across qualities and summing over follower types implies[19]

$$\dot{f}_{ht} + g_{Nt}f_{ht} = \Phi_{\mathrm{n}ht}(f_{ht}+\nu_{ht}) + \Phi_{\mathrm{q}ht}(f_{ht}+\nu_{ht}) + \Psi_{ht}f_{ht}f_{\ell t} - \delta_t f_{ht}. \quad \text{(C.149)}$$

Substituting in g_{Nt} and δ_t implies

$$\dot{f}_{ht} = (\Phi_{\mathrm{n}ht} + \Phi_{\mathrm{q}ht})(f_{ht}+\nu_{ht})f_{\ell t} - (\Phi_{\mathrm{n}\ell t} + \Phi_{\mathrm{q}\ell t})(f_{\ell t}+\nu_{\ell t})f_{ht} + \Psi_{ht}f_{ht}f_{\ell t}. \quad \text{(C.150)}$$

This expression gives some insight into why equilibria are stable and eventually converge to a balanced-growth path in which both types maintain positive market share. If low-

---

[19]As a reality check, note that doing the same for $i_1 = \ell$ and then adding it to this equation yields the expression $g_{Nt} = \sum_i \Phi_{\mathrm{n}it}(f_{it}+\nu_{it})$, which is true by definition.

productivity firms have all the market share ($f_{\ell t} = 1$), then

$$\dot{f}_{ht} = (\Phi_{nht} + \Phi_{qht})\nu_{ht} > 0. \tag{C.151}$$

The flow of new high-productivity entrants increases the market share. Conversely, if high-productivity firms produce all goods ($f_{ht} = 1$), then

$$\dot{f}_{ht} = -(\Phi_{n\ell t} + \Phi_{q\ell t})\nu_{\ell t} < 0. \tag{C.152}$$

In both of these extreme cases, there are no flows from M&A because, in the first case ($f_{\ell t} = 1$), there are no buyers looking for targets; and in the second ($f_{ht} = 1$), there are no targets to identify.

C.5 Solution to the balanced-growth equilibrium

In a balanced-growth equilibrium, we have a stationary distribution $\{f_{i_1 i_2 \hat{q}}\}$ and a constant ratio L_t/N_t. It is straightforward to see that, if we substitute these constants into the above equilibrium conditions, then all time dependence disappears. We are then left with the equilibrium conditions laid out in the main text.

C.6 Proof of Proposition 1

Proof of Proposition 1. For monopoly products ($i_2 = \varnothing$), high- and low-productivity leaders have the same markup: $\mu_{h\varnothing} = \mu_{\ell\varnothing} = \eta/(\eta-1)$. For $i_2 \in \{h, \ell\}$,

$$\frac{\mu_{hi_2}}{\mu_{\ell i_2}} = \frac{\min\{\lambda a_h/a_{i_2}, \eta/(\eta-1)\}}{\min\{\lambda a_\ell/a_{i_2}, \eta/(\eta-1)\}} \geq 1. \tag{C.153}$$

Specifically, $\mu_{hi_2} > \mu_{\ell i_2}$ if $\lambda a_\ell/a_{i_2} < \eta/(\eta-1)$, and $\mu_{hi_2} = \mu_{\ell i_2}$ otherwise. This verifies the claim that $\mu_{hi_2} \geq \mu_{\ell i_2}$. Now consider revenues:

$$\frac{p_{hi_2\hat{q}} y_{hi_2\hat{q}}}{p_{\ell i_2\hat{q}} y_{\ell i_2\hat{q}}} = \left(\frac{a_h/\mu_{hi_2}}{a_\ell/\mu_{\ell i_2}}\right)^{\eta-1}. \tag{C.154}$$

A.56

If $\mu_{hi_2} = \mu_{\ell i_2}$ (which we just showed occurs when $i_2 = \varnothing$ or when $\lambda a_\ell / a_{i_2} \geq \eta/(\eta - 1)$), then this ratio is clearly greater than one, so $p_{hi_2\hat{q}} y_{hi_2\hat{q}} > p_{\ell i_2\hat{q}} y_{\ell i_2\hat{q}}$. If instead $\mu_{hi_2} > \mu_{\ell i_2}$, then it must be that either $\mu_{hi_2}/\mu_{\ell i_2} = a_h/a_\ell$, so $p_{hi_2\hat{q}} y_{hi_2\hat{q}} = p_{\ell i_2\hat{q}} y_{\ell i_2\hat{q}}$; or $\mu_{hi_2}/\mu_{\ell i_2} = (\eta/(\eta - 1))/(\lambda a_\ell/a_{i_2}) \in (1, a_h/a_\ell)$, so $p_{hi_2\hat{q}} y_{hi_2\hat{q}} > p_{\ell i_2\hat{q}} y_{\ell i_2\hat{q}}$. This verifies the claim that $p_{hi_2\hat{q}} y_{hi_2\hat{q}} \geq p_{\ell i_2\hat{q}} y_{\ell i_2\hat{q}}$. Finally, to see that $\pi_{hi_2\hat{q}} > \pi_{\ell i_2\hat{q}}$, note that in all of the above cases, whenever $\mu_{hi_2} = \mu_{\ell i_2}$ we have $p_{hi_2\hat{q}} y_{hi_2\hat{q}} > p_{\ell i_2\hat{q}} y_{\ell i_2\hat{q}}$, and whenever $p_{hi_2\hat{q}} y_{hi_2\hat{q}} = p_{\ell i_2\hat{q}} y_{\ell i_2\hat{q}}$ we have $\mu_{hi_2} > \mu_{\ell i_2}$. Thus,

$$\frac{\pi_{hi_2\hat{q}}}{\pi_{\ell i_2\hat{q}}} = \frac{(1 - \mu_{hi_2}^{-1})}{(1 - \mu_{\ell i_2}^{-1})} \frac{p_{hi_2\hat{q}} y_{hi_2\hat{q}}}{p_{\ell i_2\hat{q}} y_{\ell i_2\hat{q}}} > 1, \tag{C.155}$$

confirming that high-productivity firms earn strictly greater operating profits on a blueprint.

Next, let us prove the claim that $v_{hi_2\hat{q}}^{\mathrm{P}} > v_{\ell i_2\hat{q}}^{\mathrm{P}}$ and $v_h^{\mathrm{G}} > v_\ell^{\mathrm{G}}$. The proof is clearest if we first consider the case without M&A ($\psi = 0$). In this case, the ratio of product values is

$$\frac{v_{hi_2\hat{q}}^{\mathrm{P}}}{v_{\ell i_2\hat{q}}^{\mathrm{P}}} = \frac{\pi_{hi_2\hat{q}}}{\pi_{\ell i_2\hat{q}}} > 1. \tag{C.156}$$

This implies that the ratio of growth values is

$$\frac{v_h^{\mathrm{G}}}{v_\ell^{\mathrm{G}}} = \frac{\Phi_{\mathrm{n}h} \bar{v}_{\mathrm{n}h}^{\mathrm{P}} + \Phi_{\mathrm{q}h} \bar{v}_{\mathrm{q}h}^{\mathrm{P}}}{\Phi_{\mathrm{n}\ell} \bar{v}_{\mathrm{n}\ell}^{\mathrm{P}} + \Phi_{\mathrm{q}\ell} \bar{v}_{\mathrm{q}\ell}^{\mathrm{P}}} > 1, \tag{C.157}$$

because $\bar{v}_{\mathrm{k}h}^{\mathrm{P}} > \bar{v}_{\mathrm{k}\ell}^{\mathrm{P}}$ means $\Phi_{\mathrm{k}h} > \Phi_{\mathrm{k}\ell}$. The intuition when we add M&A ($\psi > 0$) is similar, but we now have to rule out the possibility that low-productivity firms will want to acquire high-productivity blueprints. Note that we can only have $v_{\ell i_2\hat{q}}^{\mathrm{G}} > v_{hi_2\hat{q}}^{\mathrm{G}}$ for some blueprint type if there exists at least one product for which $v_{\ell i_2\hat{q}}^{\mathrm{P}} > v_{hi_2\hat{q}}^{\mathrm{P}}$. Suppose this is the case. The ratio of product values for such a product will equal

$$\frac{v_{hi_2\hat{q}}^{\mathrm{P}}}{v_{\ell i_2\hat{q}}^{\mathrm{P}}} = (1 - \theta_\ell^{\mathrm{P}}) \frac{\pi_{hi_2\hat{q}}}{\pi_{\ell i_2\hat{q}}} + \theta_\ell^{\mathrm{P}}, \tag{C.158}$$

where

$$\theta_\ell^{\mathrm{P}} \equiv \frac{\Psi_\ell f_\ell (1 - \varrho)}{\rho + \delta + (\eta - 1) g_Q + \Psi_\ell f_\ell (1 - \varrho)} \in [0, 1). \tag{C.159}$$

But then $\pi_{hi_2\hat{q}} > \pi_{\ell i_2\hat{q}}$ implies that $v_{hi_2\hat{q}}^{\mathrm{P}} > v_{\ell i_2\hat{q}}^{\mathrm{P}}$, because (C.158) is a weighted average of one and a number greater than one. This contradicts the assumption, meaning we cannot

have $v^{\mathrm{G}}_{\ell i_2 \hat q} > v^{\mathrm{G}}_{h i_2 \hat q}$ for any blueprint type. Hence, it is always the case that $v^{\mathrm{G}}_{h i_2 \hat q} > v^{\mathrm{G}}_{\ell i_2 \hat q}$ for all blueprint types, which implies that

$$\frac{v^{\mathrm{P}}_{\ell i_2 \hat q}}{v^{\mathrm{P}}_{h i_2 \hat q}} = (1 - \theta^{\mathrm{P}}_h)\frac{\pi_{h i_2 \hat q}}{\pi_{\ell i_2 \hat q}} + \theta^{\mathrm{P}}_h, \tag{C.160}$$

where

$$\theta^{\mathrm{P}}_h \equiv \frac{\Psi_h f_h (1 - \varrho)}{\rho + \delta + (\eta - 1)g_Q + \Psi_h f_h (1 - \varrho)} \in [0, 1). \tag{C.161}$$

This implies that $v^{\mathrm{P}}_{h i_2 \hat q} > v^{\mathrm{P}}_{\ell i_2 \hat q}$ for all blueprint types, confirming the claim.

The second claim of the proposition follows from the fact that $x_{\mathrm{k}h} > x_{\mathrm{k}\ell}$ if and only if $\bar v_{\mathrm{k}h} > \bar v_{\mathrm{k}\ell}$, which is implied by $v^{\mathrm{P}}_{h i_2 \hat q} > v^{\mathrm{P}}_{\ell i_2 \hat q}$ and $v^{\mathrm{G}}_h > v^{\mathrm{G}}_\ell$. $\qquad\square$

D   Derivations for distribution of firms

The characteristics of a firm can be summarized by its productivity type $a_i \in \{a_h, a_\ell\}$, the number blueprints it operates $n \in \mathbb{N} \setminus \{0\}$, and the characteristics of those blueprints $\{i_{2jt}, \hat q_{jt}\}_{j=1}^n \in \{h, \ell, \varnothing\}^n \times \mathbb{R}^n$. Thus, the entire firm distribution is summarized over all possible combinations of these characteristics. This distribution is too complicated to express analytically; however, we can derive closed-form solutions for the simpler distribution of types a_i, blueprint counts n, and ages $\tau \in \mathbb{R}_+$. As explained in the main text and below, this lower-dimensional distribution yields several useful insights about the equilibrium.

D.1   Notation

Let $M_t(a_i, n, \tau)$ denote the measure of firms with characteristics $\{a_i, n, \tau\}$ and $m_{it}(n, \tau) \equiv M_t(a_i, n, \tau)/M_t$ the share of such firms. Marginal distributions omit arguments: for example, $m_{it}(\tau) = \sum_{n=1}^\infty m_{it}(n, \tau)$. Note that the mass M_t only includes firms that are currently operating ($n \geq 1$), not firms that have exited in the past. Another useful distribution will be the density of *firm-products* $p_{it}(n, \tau) \equiv nM_t(a_i, n, \tau)/N_t$, which represents the share of *blueprints* held by firms with the corresponding characteristics. The firm distribution and

firm-product distribution are linked by the equation

$$m_{it}(n, \tau) = \frac{N_t}{M_t} \frac{p_{it}(n, \tau)}{n}, \tag{D.1}$$

which implies that the average number of blueprints per firm is $N_t/M_t = \sum_{n=1}^{\infty} n m_{it}(n)$.

It is convenient to define the firm-specific blueprint-gain and blueprint-loss intensities:

$$\bar{\Phi}_{it} \equiv \Phi_{\mathrm{n}it} + \Phi_{\mathrm{q}it} + \mathbb{I}_{i=h} \Psi_{ht} f_{\ell t}, \tag{D.2}$$

$$\bar{\delta}_{it} \equiv \delta_t + \mathbb{I}_{i=\ell} \Psi_{ht} f_{ht}. \tag{D.3}$$

The former is the intensity, per current blueprint, with which a firm is expected to generate new blueprints via R&D or M&A; the latter is the intensity with which a firm loses a blueprint via creative destruction or M&A exit.

D.2 Derivation of KFEs

The KFEs for the firm distribution are derived similarly to those for the blueprint distribution. First, to see how the distribution evolves with respect to age τ, suppose there is no R&D or M&A ($\bar{\Phi}_{it} = \bar{\delta}_{it} = 0$) and that time evolves in discrete increments of length Δt. The change in the mass of firms $M_{it}(n, \tau)$ from time $t - \Delta t$ to time t equals

$$\Delta M_{it}(n, \tau) = M_{i,t-\Delta t}(n, \tau - \Delta t) - M_{i,t-\Delta t}(n, \tau). \tag{D.4}$$

Dividing by $M_{i,t-\Delta t} \Delta t$ and taking limits implies

$$\dot{m}_{it}(n, \tau) + g_{Mt} m_{it}(n, \tau) = \lim_{\Delta t \to 0} -\frac{m_{i,t-\Delta t}(n, \tau) - m_{i,t-\Delta t}(n, \tau - \Delta t)}{\Delta t} = -\frac{\partial m_{it}(n, \tau)}{\partial \tau}. \tag{D.5}$$

where $g_{Mt} \equiv \dot{M}_t/M_t$ is the net growth rate of the number of firms.

Now consider the case in which firms do R&D and M&A. The mass of firms with $n \geq 2$ blueprints increases when firms with $n - 1$ products successfully gain a blueprint, increases when firms with $n + 1$ products lose a blueprint, and decreases when firms with n products

A.59

gain or lose blueprints. Hence, we have the KFEs

$$\dot{m}_{it}(n,\tau) + g_{Mt}m_{it}(n,\tau) = -\frac{\partial m_{it}(n,\tau)}{\partial \tau} + \bar{\Phi}_{it}[(n-1)m_{it}(n-1,\tau) - nm_{it}(n,\tau)]$$

$$+ \bar{\delta}_{it}[(n+1)m_{it}(n+1,\tau) - nm_{it}(n,\tau)]. \quad \text{(D.6)}$$

The KFEs for firms with $n=1$ blueprint is similar, but there are now no inflows from firms with $n-1$ products and, at $\tau=0$, there is also entry by entrepreneurs:

$$\dot{m}_{it}(1,\tau) + g_{Mt}m_{it}(1,\tau) = -\frac{\partial m_i(1,\tau)}{\partial \tau} + \mathbb{I}_{\tau=0}(\Phi_{\mathrm{n}it} + \Phi_{\mathrm{q}it})\nu_{it}$$

$$- \bar{\Phi}_{it}m_{it}(1,\tau) + \bar{\delta}_{it}[2m_{it}(2,\tau) - m_{it}(1,\tau)]. \quad \text{(D.7)}$$

In all that follows, we assume a balanced-growth equilibrium, which implies constant intensities $\bar{\Phi}_i$ and $\bar{\delta}_i$, a stationary firm distribution $m_i(n,\tau)$, and a constant growth rate of firms g_M equal to the growth rate of blueprints g_N.

D.3 Distribution of firm size and age

There are three steps in deriving the full distribution $m_i(n,\tau)$: first, solve for the distribution of size given age and type $m(n|a_i,\tau)$; then, solve for the marginal distribution of age and type $m_i(\tau)$; and finally, use Bayes' rule to get the full distribution $m_i(n,\tau) = m(n|a_i,\tau)m_i(\tau)$.

D.3.1 Distribution of size given age and type

Every firm begins its life with one product, so the initial condition is $m(n|a_i,\tau=0) = \mathbb{I}_{n=1}$. From this initial condition, the conditional distribution evolves with firm age as

$$\frac{\partial m(n|a_i,\tau)}{\partial \tau} = \bar{\Phi}_i[(n-1)m(n-1|a_i,\tau) - nm(n|a_i,\tau)]$$

$$+ \bar{\delta}_i[(n+1)m(n+1|a_i,\tau) - nm(n|a_i,\tau)]. \quad \text{(D.8)}$$

The solution to this differential-difference equation is the function

$$m(n|\tau, a_i) = \frac{(\bar{\Phi}_i - \bar{\delta}_i)e^{-(\bar{\Phi}_i - \bar{\delta}_i)\tau}}{\bar{\Phi}_i - \bar{\delta}_i e^{-(\bar{\Phi}_i - \bar{\delta}_i)\tau}} \left(\frac{\bar{\Phi}_i(1 - e^{-(\bar{\Phi}_i - \bar{\delta}_i)\tau})}{\bar{\Phi}_i - \bar{\delta}_i e^{-(\bar{\Phi}_i - \bar{\delta}_i)\tau}} \right)^{n-1}. \tag{D.9}$$

This solution can be derived via the probability-generating function of Kendall (1948), as laid out in Luttmer (2011) and Klette and Kortum (2004). Alternatively, one could verify this solution by simply substituting (D.9) into the conditional KFEs (D.8).

D.3.2 Distribution of age and type

Aggregating the KFEs for $m_i(n, \tau)$ over n for any $\tau > 0$, we have

$$g_N m_i(\tau) = -\frac{\partial m_i(\tau)}{\partial \tau} - \bar{\delta}_i m_i(1, \tau). \tag{D.10}$$

Dividing by $m_i(\tau)$ and integrating from over τ implies that the solution to this is

$$m_i(\tau) = \kappa_i \exp\left\{ -\left(g_N \tau + \bar{\delta}_i \int_0^\tau m(1|a_i, \tau')d\tau' \right) \right\}, \tag{D.11}$$

for some constants of integration κ_i. Using the conditional size distribution (D.9) at $n = 1$,

$$m(1|\tau, a_i) = \frac{(\bar{\Phi}_i - \bar{\delta}_i)e^{-(\bar{\Phi}_i - \bar{\delta}_i)\tau}}{\bar{\Phi}_i - \bar{\delta}_i e^{-(\bar{\Phi}_i - \bar{\delta}_i)\tau}}, \tag{D.12}$$

the integral in (D.11) evaluates to

$$\bar{\delta}_i \int_0^\tau m(1|a_i, \tau')d\tau' = \log(|\bar{\Phi}_i - \bar{\delta}_i e^{-(\bar{\Phi}_i - \bar{\delta}_i)\tau'}|)\Big|_0^\tau = \log\left(\frac{\bar{\Phi}_i - \bar{\delta}_i e^{-(\bar{\Phi}_i - \bar{\delta}_i)\tau}}{\bar{\Phi}_i - \bar{\delta}_i} \right). \tag{D.13}$$

If $\bar{\Phi}_i > \bar{\delta}_i$ (type-a_i firms grow on average), then this grows asymptotically to a constant $\log(\bar{\Phi}_i/(\bar{\Phi}_i - \bar{\delta}_i))$. If $\bar{\Phi}_i < \bar{\delta}_i$, then this can be written as

$$\bar{\delta}_i \int_0^\tau m(1|a_i, \tau')d\tau' = \log\left(\frac{\bar{\delta}_i e^{(\bar{\delta}_i - \bar{\Phi}_i)\tau} - \bar{\Phi}_i}{\bar{\delta}_i - \bar{\Phi}_i} \right) = (\bar{\delta}_i - \bar{\Phi}_i)\tau + \log\left(\frac{\bar{\delta}_i - \bar{\Phi}_i e^{-(\bar{\delta}_i - \bar{\Phi}_i)\tau}}{\bar{\delta}_i - \bar{\Phi}_i} \right), \tag{D.14}$$

which asymptotically grows linearly with τ.

The last step in solving for $m_i(\tau)$ is finding the constants κ_i in (D.11). These expressions need to integrate to the total firm-type shares $m_i = \int_0^\infty m_i(\tau)d\tau$, so it must be that

$$\kappa_i = \frac{m_i}{\bar{m}_i}, \quad \text{where} \quad \bar{m}_i \equiv \int_0^\infty \exp\left\{-\left(g_N\tau + \bar{\delta}_i \int_0^\tau m(1|a_i, \tau')d\tau'\right)\right\} d\tau. \tag{D.15}$$

To solve for m_i, and thus for $m_i(\tau)$, note first that, at $\tau = 0$, we have

$$m_i(\tau = 0) = \kappa_i = \frac{m_i}{\bar{m}_i}, \tag{D.16}$$

and hence the conditional share of type-a_i firms at $\tau = 0$ is

$$m(a_i|\tau = 0) = \frac{m_i(\tau = 0)}{m_h(\tau = 0) + m_\ell(\tau = 0)} = \frac{m_i/\bar{m}_i}{m_h/\bar{m}_h + m_\ell/\bar{m}_\ell}. \tag{D.17}$$

Second, note that the conditional share at $\tau = 0$ is also equal to the relative rates of entry:

$$m(a_i|\tau = 0) = \frac{\omega_i(\Phi_{\mathrm{n}i} + \Phi_{\mathrm{q}i})}{\omega_h(\Phi_{\mathrm{n}h} + \Phi_{\mathrm{q}h}) + \omega_\ell(\Phi_{\mathrm{n}\ell} + \Phi_{\mathrm{q}\ell})}. \tag{D.18}$$

Equating these two expressions and solving for m_i implies

$$m_i = \frac{\omega_i(\Phi_{\mathrm{n}i} + \Phi_{\mathrm{q}i})\bar{m}_i}{\omega_h(\Phi_{\mathrm{n}h} + \Phi_{\mathrm{q}h})\bar{m}_h + \omega_\ell(\Phi_{\mathrm{n}\ell} + \Phi_{\mathrm{q}\ell})\bar{m}_\ell}. \tag{D.19}$$

Substituting this back into (D.16), we have

$$\kappa_i = \frac{m_i}{\bar{m}_i} = \frac{\omega_i(\Phi_{\mathrm{n}i} + \Phi_{\mathrm{q}i})\bar{m}_i}{\omega_h(\Phi_{\mathrm{n}h} + \Phi_{\mathrm{q}h})\bar{m}_h + \omega_\ell(\Phi_{\mathrm{n}\ell} + \Phi_{\mathrm{q}\ell})\bar{m}_\ell}, \tag{D.20}$$

and hence a closed-form solution for $m_i(\tau)$.

Finally by Bayes' rule, we can multiply the conditional distribution (D.9) with the marginal distribution (D.11) to get the solution $m_i(n, \tau)$.

D.4 Derivation of firm entry and exit rates

The firm entry rate $\mathcal{E}_{\text{entry}}$ is defined as the flow of new firms being created relative to the number of incumbent firms:

$$\mathcal{E}_{\text{entry},t} = \frac{\sum_{i\in\{h,\ell\}}(\Phi_{\text{n}it} + \Phi_{\text{q}it})\omega_i\bar{s}L_t}{M_t} = \frac{N_t}{M_t}\sum_{i\in\{h,\ell\}}(\Phi_{\text{n}it} + \Phi_{\text{q}it})\nu_{it}. \tag{D.21}$$

The ratio N_t/M_t can be computed from the marginal distribution $m_t(n) = \sum_i \int_0^\infty m_{it}(n,\tau)d\tau$:

$$\frac{N_t}{M_t} = \sum_{n=1}^\infty nm_t(n). \tag{D.22}$$

The exit rate is likewise defined as the flow of firms existing relative to the number of incumbent firms:

$$\mathcal{E}_{\text{exit},t} = \frac{\sum_{i\in\{h,\ell\}}\bar{\delta}_{it}M_{it}(n=1)}{M_t} = \sum_{i\in\{h,\ell\}}\bar{\delta}_{it}m_{it}(n=1). \tag{D.23}$$

In a balanced-growth equilibrium, the entry and exit rates are constant and are linked to each other by the relation

$$g_L = \mathcal{E}_{\text{entry}} - \mathcal{E}_{\text{exit}}, \tag{D.24}$$

because the growth rate in the number of firms $g_M = \dot{M}_t/M_t$ is equal to the growth rate of the population g_L.

D.5 Derivation of Pareto tail coefficients

As Luttmer (2011) shows, if $\bar{\Phi}_i > \bar{\delta}_i$, then the distribution of blueprints $m_i(n)$ exhibits a Pareto tail with tail index $\theta_i > 0$. To represent the fact that the probability density function $m_i(n)$ behaves like $n^{-(1+\theta_i)}$, I use the notation $m_i(n) \overset{\lim}{\sim} n^{-(1+\theta_i)}$.[20] To show this, it is easiest to first show that the firm-product distribution $p_i(n) \propto nm_i(n) \overset{\lim}{\sim} n^{-\theta_i}$. The tail index is

---

[20]Specifically, I use the notation $f(n) \overset{\lim}{\sim} g(n)$ to represent the fact that there exists a finite constant $k > 0$ such that $\lim_{n\to\infty} \frac{f(n)}{g(n)} = k$.

defined as the limit of the sequence

$$\theta_i \equiv \lim_{n \to \infty} -n \frac{p_i(n+1) - p_i(n)}{p_i(n)}. \tag{D.25}$$

Intuitively, this tail index behaves like the continuous elasticity $\lim_{n \to \infty} -\partial \log p_i(n) / \partial \log n$ and is finite if $\log p_i(n)$ falls at the same rate as (or more slowly than) $\log n$ rises. If the distribution exhibits a tail that is thinner than a Pareto distribution (which will occur for $\bar{\bar{\Phi}}_i \leq \bar{\delta}_i$), then the tail index will be infinite.

Consider the case $\bar{\bar{\Phi}}_i > \bar{\delta}_i$. To compute the Pareto tail coefficient, note that the KFE for $m_i(n)$ is

$$g_N m_i(n) = \bar{\bar{\Phi}}_i[(n-1)m_i(n-1) - nm_i(n)] + \bar{\delta}_i[(n+1)m_i(n+1) - nm_i(n)]. \tag{D.26}$$

Dividing both sides by N_t, we can rewrite this in terms of the firm-product distribution

$$g_N \frac{1}{n} p_i(n) = \bar{\bar{\Phi}}_i[p_i(n-1) - p_i(n)] + \bar{\delta}_i[p_i(n+1) - p_i(n)]. \tag{D.27}$$

Multiplying both sides by $n/p_i(n)$ and taking limits as $n \to \infty$, we then get the tail coefficient

$$\theta_i = \frac{g_N}{\bar{\bar{\Phi}}_i - \bar{\delta}_i}. \tag{D.28}$$

If instead $\bar{\bar{\Phi}}_i < \bar{\delta}_i$, then the tail is thinner than Pareto. We instead have $p_i(n) \overset{\lim}{\sim} (\bar{\bar{\Phi}}_i / \bar{\delta}_i)^n$. One can verify this by dividing (D.27) by $p_i(n)$, substituting this in, and taking $n \to \infty$. As shown in Luttmer (2011), the tail index for the entire distribution of firms (i.e., including both high and low types) is

$$\theta = \min\{\theta_h, \theta_\ell\} = \min\left\{ \frac{g_N}{[\bar{\bar{\Phi}}_h - \bar{\delta}_h]^+}, \frac{g_N}{[\bar{\bar{\Phi}}_\ell - \bar{\delta}_\ell]^+} \right\} = \theta_h. \tag{D.29}$$

High-productivity firms grow faster than low-productivity firms on average, so they dominate the tail.

As a final step, let us prove that the tail index θ of the blueprint distribution also applies

A.64

to the distribution of employment and the distributions of rescaled firm revenues, profits, and value. Firm i's total employment per blueprint equals

$$L_{it} = l_{\mathrm{p}}(1 - \bar{s})\frac{L_t}{N_t}\left(\frac{a_i}{\bar{a}}\right)^{\eta-1} n_{it}v_{\eta it}, \tag{D.30}$$

where

$$v_{\eta it} \equiv \frac{1}{n_{it}}\sum_{j=1}^{n_{it}}\left(\frac{q_{ijt}}{Q_t}\right)^{\eta-1}\left(\frac{\mu_{ijt}}{\mathcal{M}_{\eta-1}}\right)^{-\eta} \tag{D.31}$$

is an average of qualities and markups across products. Because the mix of product gains and losses is invariant to a firm's size, this average $v_{\eta it}$ converges in probability to a constant as $n \to \infty$: there exists a finite constant $\bar{v}_\eta(a_i) > 0$ (which may depend on the type a_i) such that

$$\lim_{n\to\infty}\mathbb{P}(v_{\eta it} = \bar{v}_\eta(a_i)) = 1. \tag{D.32}$$

I henceforth use the notation $v_{\eta it} \xrightarrow{\mathrm{P}} \bar{v}_\eta(a_i)$ to denote this probability limit. This means that

$$\frac{L_{it}}{n_{it}} \xrightarrow{p} l_{\mathrm{p}}(1 - \bar{s})\frac{L_t}{N_t}\left(\frac{a_i}{\bar{a}}\right)^{\eta-1}\bar{v}_\eta(a_i), \tag{D.33}$$

and thus that $L \overset{\mathrm{lim}}{\sim} n$ (in a probabilistic sense). Among very large firms, product portfolios are well-diversified, so employment is linear in n. It follows that the Pareto tail of the employment distribution looks the same as that of the blueprint distribution. The exact same logic applies to the rescaled sales, profit, and value distributions:

$$\frac{p_{it}y_{it}}{Y_t/L_t} \overset{\mathrm{lim}}{\sim} n_{it}, \quad \hat{\Pi}_{it} \overset{\mathrm{lim}}{\sim} n_{it}, \quad \text{and} \quad \hat{V}_{it} \overset{\mathrm{lim}}{\sim} n_{it}. \tag{D.34}$$

E   Numerical solution and simulation methods

E.1   Balanced-growth path: Equilibrium solution method

This section outlines the numerical method used to solve the system of equilibrium conditions in Section 3.2. The basic idea is two loops: an outer loop over the blueprint distribution $\{f_{i_1 i_2 \hat{q}}\}$ and an inner loop over the ratio L_t/N_t, which is constant in a BGE.

E.1.1 State space

The states needed to calculate blueprint values $\{v^{\mathrm{P}}_{i_1 i_2}, v^{\mathrm{G}}_{i_1}\}$ are $\{i_1, i_2\}$ (recall that we solved out $\hat{q}$ analytically, so we do not need to consider this state), a six-dimensional grid. To solve for the blueprint distribution, we need the states $\{i_1, i_2, \hat{q}\}$, which is infinite-dimensional. Thus, we discretize the grid for $\hat{q}$ on an evenly spaced grid with increments $\log \lambda / m$ for some positive integer m. The grid has $2K + 1$ points evenly centered around zero:

$$\hat{q} \in \left\{ -K \frac{\log \lambda}{m}, -(K-1) \frac{\log \lambda}{m}, \dots, -\frac{\log \lambda}{m}, 0, \frac{\log \lambda}{m}, \dots, (K-1) \frac{\log \lambda}{m}, K \frac{\log \lambda}{m} \right\}.$$

Thus, the full numerical state space has $6 \times (2K + 1)$ points.

E.1.2 Overview of solution algorithm

Ultimately, we are interested are the values $\{v^{\mathrm{P}}_{i_1 i_2}, v^{\mathrm{G}}_i\}$ and the distribution $\{f_{i_1 i_2 \hat{q}}\}$ that solve the HJBs and KFEs. These unknowns depend on, and jointly imply, macroeconomic aggregates, firm markups and profits, and R&D and M&A policies $\{x_{\mathrm{n}i}, x_{\mathrm{q}i}, x_{\mathrm{m}i}\}$. Starting from an initial guess for values and distributions (iteration 0), the algorithm at iteration k is as follows:

1. Given the distribution $\{f^{k-1}_{i_1 i_2 \hat{q}}\}$ and policies $\{x^{k-1}_{\mathrm{n}i}, x^{k-1}_{\mathrm{q}i}, x^{k-1}_{\mathrm{m}i}\}$, compute the profit rates $\{\pi^k_{i_1 i_2}\}$ and values $\{v^{\mathrm{P},k}_{i_1 i_2}\}$ and $\{v^{\mathrm{G},k}_i\}$ from the HJBs.

2. Under the new values $\{v^{\mathrm{P},k}_{i_1 i_2}\}$ and $\{v^{\mathrm{G},k}_i\}$ and the old distribution $\{f^{k-1}_{i_1 i_2 \hat{q}}\}$, use an inner loop to solve for policies $\{x^k_{\mathrm{n}i}, x^k_{\mathrm{q}i}, x^k_{\mathrm{m}i}\}$ and the aggregates $\{LN^k, \hat{w}^k_{\mathrm{x}}, f^k_h\}$ that are consistent with balanced growth.

3. Using the new values and policies, solve for the distribution $\{f^k_{i_1 i_2 \hat{q}}\}$; use this distribution to re-compute macroeconomic aggregates.

4. Check if $\{f^k_{i_1 i_2 \hat{q}}\}$ has converged; if not, return to step 1.

This iterative process exploits the stability of a stationary distribution: starting from an arbitrary distribution, the model will converge to the balanced-growth path.

E.1.3 Initialization

The initialization requires a guess for the distribution $\{f^0_{i_1 i_2 \hat{q}}\}$, policies $\{x^0_{ni}, x^0_{qi}, x^0_{mi}\}$, and ratio LN^0. Given that the speed of convergence is insensitive to this initialization, I set the distribution to be normal across $\hat{q}$ (with standard deviation small relative to the grid) and uniform across pairs $\{i_1, i_2\}$. I set unit policies $x^0_{ni} = x^0_{qi} = 1$ and $x^0_{mh} = 1$. I set $LN^0 = 10$.

Before starting the algorithm, we also need to compute some aggregate variables using these initial guesses. These include the aggregates $\{\bar{a}^0, \mathcal{M}^0_{\eta-1}\}$, which we need to calculate profits; and the growth rates $\{\delta^0, g^0_N\}$ and rescaled skilled wage $\hat{w}^0_x$, which are needed to compute values and distributions. These computations are explained in further detail below.

E.1.4 Solving HJBs

The markups of a product $\mu_{i_1 i_2}$ are independent of product values and the distribution, so these can be computed with initialized values. The profit rate at iteration k then equals

$$\pi^k_{i_1 i_2} = \left(1 - \frac{1}{\mu_{i_1 i_2}}\right) \left(\frac{a_{i_1}/\mu_{i_1 i_2}}{\bar{a}^{k-1}/\mathcal{M}^{k-1}_{\eta-1}}\right)^{\eta-1} LN^{k-1}, \tag{E.1}$$

where $\bar{a}^{k-1}$ and $\mathcal{M}^{k-1}_{\eta-1}$ are computed using $\{f^{k-1}_{i_1 i_2 \hat{q}}\}$.

Using the distribution $\{f^{k-1}_{i_1 i_2 \hat{q}}\}$, the policies $\{x^{k-1}_{ni}, x^{k-1}_{qi}, x^{k-1}_{mi}\}$, and the ratio LN^{k-1}, we can compute product values as

$$v^{P,k}_{i_1 i_2} = \frac{\pi^k_{i_1 i_2} + \mathbb{I}_{i=\ell} \Psi^{k-1}_h f^{k-1}_h (1 - \varrho) v^{P,k}_{h i_2}}{\rho + \delta^{k-1} + (\eta - 1) g^{k-1}_Q + \mathbb{I}_{i=\ell} \Psi^{k-1}_h f^{k-1}_h (1 - \varrho)}. \tag{E.2}$$

Using these product values, we can compute expected product values, $\{\bar{v}^k_{ni}, \bar{v}^k_{qi}, \bar{v}^k_{mi}\}$, with which we can compute the value of growth opportunities. Because of M&A, growth values are interdependent in the HJB, so we need to invert a two-by-two system: for $v^G = [v^G_h \quad v^G_\ell]^\top$,

$$C^{G,k} v^{G,k} = \tilde{v}^{G,k}, \tag{E.3}$$

where (suppressing k superscripts for the moment)

$$C^{\mathrm{G}} = \begin{bmatrix} \rho + \delta - (1-\varepsilon)(\Phi_{\mathrm{n}h} + \Phi_{\mathrm{q}h} + \Psi_h f_\ell \varrho) & (1-\varepsilon)\Psi_h f_\ell \varrho \\ -\Psi_h f_h(1-\varrho) & \rho + \delta - (1-\varepsilon)(\Phi_{\mathrm{n}\ell} + \Phi_{\mathrm{q}\ell}) + \Psi_h f_h(1-\varrho) \end{bmatrix}. \tag{E.4}$$

and

$$\tilde{v}^{\mathrm{G}} = \begin{bmatrix} (1-\varepsilon)(\Phi_{\mathrm{n}h}\bar{v}^{\mathrm{P}}_{\mathrm{n}h} + \Phi_{\mathrm{q}h}\bar{v}^{\mathrm{P}}_{\mathrm{q}h} + \Psi_h f_\ell \bar{v}^{\mathrm{P}}_{\mathrm{m}h}) \\ (1-\varepsilon)(\Phi_{\mathrm{n}\ell}\bar{v}^{\mathrm{P}}_{\mathrm{n}\ell} + \Phi_{\mathrm{q}\ell}\bar{v}^{\mathrm{P}}_{\mathrm{q}\ell}) \end{bmatrix}. \tag{E.5}$$

Inverting this system gives us $\{v_{i_1}^{\mathrm{G},k}\}$.

E.1.5 Inner loop: Solving for policies and growth rates

A challenge presented by the balanced-growth path is that we need $g_N = g_L$. This means that we need to simultaneously find the ratio LN, the policies $\{x_{\mathrm{n}i}^k, x_{\mathrm{q}i}^k, x_{\mathrm{m}i}^k\}$, and the total market share f_h that are consistent with this condition and with each other. The relations among these objects are highly nonlinear, so I implement an iterative procedure.

Within iteration k, we re-initialize the procedure at iteration $l = 0$, setting $\{LN^0, \hat{w}_{\mathrm{x}}^0, f_h^0\}$ equal to their values $\{LN^{k-1}, \hat{w}_{\mathrm{x}}^{k-1}, f_h^{k-1}\}$ from the previous outer-loop iteration. We can then calculate the new-variety R&D policies

$$x_{\mathrm{n}i}^l = \left(\frac{\varepsilon(1-\alpha)\varphi}{\hat{w}_{\mathrm{x}}^{l-1}} (\bar{v}_{\mathrm{n}i}^{\mathrm{P},k} + v_i^{\mathrm{G},k}) \right)^{1/(1-\varepsilon)}, \tag{E.6}$$

and do the same for $x_{\mathrm{q}i}^k$ and $x_{\mathrm{m}i}^k$. These imply $\{\Phi_{\mathrm{n}i}^l, \Phi_{\mathrm{q}i}^l, \Psi_i^l\}$. Next, compute the LN^l that solves $g_N = g_L$, or

$$LN^l = \frac{g_L - \sum_{i\in\{\ell,h\}} f_i^l \Phi_{\mathrm{n}i}^l}{\bar{s} \sum_{i\in\{\ell,h\}} \omega_i \Phi_{\mathrm{n}i}^l} \tag{E.7}$$

Then find f_h^l that solves

$$(g_L + \Phi_{\mathrm{q}h}^l(\nu_h^l + f_h) + \Phi_{\mathrm{q}\ell}^l(\nu_\ell^l + f_\ell))f_h^l = (\Phi_{\mathrm{n}h}^l + \Phi_{\mathrm{q}h}^l)(\nu_h^l + f_h^l) + \Psi_h^l f_h^l(1 - f_h^l). \tag{E.8}$$

Substituting $f_\ell^l = 1 - f_h^l$, this is a quadratic equation,

$$\underbrace{-(\Phi_{nh} + \Phi_{qh})\nu_h}_{\phi_0^l} + \underbrace{(g_L + \Phi_{qh}(\nu_h - 1) + \Phi_{q\ell}(\nu_\ell + 1) - \Phi_{nh} - \Psi_h)}_{\phi_1^l} f_h^l$$

$$+ \underbrace{(\Phi_{qh} + \Psi_h - \Phi_{q\ell})}_{\phi_2^l}(f_h^l)^2 = 0. \quad \text{(E.9)}$$

the solution to which is the positive root

$$f_h^l = -\frac{\phi_1^l}{2\phi_2^l} + \sqrt{\left(\frac{\phi_1^l}{2\phi_2^l}\right)^2 - \frac{\phi_0^l}{\phi_2^l}}. \quad \text{(E.10)}$$

Finally, we can re-solve for the skilled wage

$$\hat{w}_x^l = \left((\hat{w}_x^{l-1})^{\frac{1}{1-\varepsilon}} \sum_{i \in \{\ell,h\}} \left(\omega_i(x_{ni}^l + x_{qi}^l) + \frac{1}{\bar{s}LN^l} f_i^l(x_{ni}^l + x_{qi}^l + x_{mi}^l)\right)\right)^{\frac{1-\varepsilon}{1+\zeta(1-\varepsilon)}} \quad \text{(E.11)}$$

The criterion for the convergence of this inner loop is that LN is constant between iterations: $|LN^l - LN^{l-1}| < \epsilon_{\text{inner}}^{\text{tol}}$, where I use the threshold $\epsilon_{\text{inner}}^{\text{tol}} = 10^{-8}$. At this final iteration l, we take these values for $\{x_{ni}^k, x_{qi}^k, x_{mi}^k\}$, $\{\Phi_{ni}^k, \Phi_{qi}^k, \Psi_i^k\}$, LN^k, $\hat{w}_x^k$, and f_h^k.

E.1.6 Solving KFEs

Given f_h, we can solve the KFEs as a $(6 \times (2K+1))$-dimensional linear system. For notational simplicity, suppress the iteration index k in what follows. Note that the derivative over relative quality $\hat{q}$ can be approximately by the discretization

$$\frac{\partial f_{i_1 i_2 \hat{q}}}{\partial \hat{q}} = \frac{f_{i_1 i_2, \hat{q} + \log \lambda/m} - f_{i_1 i_2 \hat{q}}}{\log \lambda/m} \quad \text{(E.12)}$$

We take a forward difference because $\hat{q}$ is falling over time.

We will use the following notation. Let $C_{ss'}$ denote a $(2K+1) \times (2K+1)$ matrix of the

A.69

form

$$
C_{ss'} = \begin{bmatrix} c_{ss'}^{(0)} & c_{ss'}^{(1)} & c_{ss'}^{(2)} & c_{ss'}^{(3)} & \cdots \\ c_{ss'}^{(-1)} & c_{ss'}^{(0)} & c_{ss'}^{(1)} & c_{ss'}^{(2)} & \cdots \\ c_{ss'}^{(-2)} & c_{ss'}^{(-1)} & c_{ss'}^{(0)} & c_{ss'}^{(1)} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{E.13}
$$

for constants $c_{ss'}^{(j)}$ to be defined. The indices s correspond to the states $\{i_1, i_2\}$ ordered from 1 to 6 as $\{\{h, \varnothing\}, \{\ell, \varnothing\}, \{h, h\}, \{\ell, \ell\}, \{h, \ell\}, \{\ell, h\}\}$; the same is true of s'. The interpretation of element $c_{ss'}^{(j)}$ is thus the transition rate into the current state s and quality $\hat{q}$ (the row) from the state s' with quality j increments higher than $\hat{q}$.

Now, let us concatenate these block matrices into the 6×6 block matrix

$$
C = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{16} \\ C_{21} & C_{22} & \cdots & C_{26} \\ \vdots & \vdots & \ddots & \vdots \\ C_{61} & C_{62} & \cdots & C_{66} \end{bmatrix}, \tag{E.14}
$$

which ultimately is a $6(2K+1)) \times 6(2K+1))$ matrix.

Letting $\mathbf{f}$ denote the $(6(2K+1))$-dimensional vector of the distribution $f_{i_1 i_2 \hat{q}}$ with element order corresponding to the order in C, we can write the KFEs as the linear system

$$
C\mathbf{f} = \mathbf{v}, \tag{E.15}
$$

where the right-hand side vector $\mathbf{v}$ is a concatenation of the new-variety quality distribution and a vector of zeros:

$$
\mathbf{v} \equiv [\Phi_{nh}(\nu_h + f_h)\Gamma^\top, \Phi_{n\ell}(\nu_\ell + f_\ell)\Gamma^\top, 0_{4(2K+1)}^\top]^\top. \tag{E.16}
$$

(Recall that we have assumed $\Gamma_{\hat{q}}$ is a point mass at $\hat{q} = \log \lambda$.) From this equation, it follows

that the diagonal block matrices C_{ss} have elements

$$c_{ss}^{(0)} = g_N + \delta + \mathbb{I}_{s \in \{2,4,6\}} \Psi_h f_h + \frac{g_Q}{\log \lambda/m}, \tag{E.17}$$

$$c_{ss}^{(1)} = -\frac{g_Q}{\log \lambda/m}, \tag{E.18}$$

where $\mathbb{I}_{s \in \{2,4,6\}} = 1$ if $s \in \{2,4,6\}$. All other elements $c_{ss}^{(j)}$ equal zero. The diagonals of C_{12}, C_{36}, and C_{54} are

$$c_{12}^{(0)} = c_{36}^{(0)} = c_{54}^{(0)} = -\Psi_h f_h. \tag{E.19}$$

For $\{h, h\}$ we have non-zero elements of C_{31}, C_{33}, and C_{35} equal to

$$c_{31}^{(-m)} = c_{33}^{(-m)} = c_{35}^{(-m)} = -\Phi_{\mathrm{q}h}(\nu_h + f_h). \tag{E.20}$$

Similarly, for $\{\ell, \ell\}$,

$$c_{42}^{(-m)} = c_{44}^{(-m)} = c_{46}^{(-m)} = -\Phi_{\mathrm{q}\ell}(\nu_\ell + f_\ell). \tag{E.21}$$

And finally, for $\{h, \ell\}$,

$$c_{52}^{(-m)} = c_{54}^{(-m)} = c_{56}^{(-m)} = -\Phi_{\mathrm{q}h}(\nu_h + f_h), \tag{E.22}$$

and

$$c_{61}^{(-m)} = c_{63}^{(-m)} = c_{65}^{(-m)} = -\Phi_{\mathrm{q}\ell}(\nu_\ell + f_\ell). \tag{E.23}$$

All other elements of C are set equal to zero.[21]

Because we have discretized the distribution and therefore truncated the relative quality distribution, we need to adjust coefficients at the boundary to account for the (very small) number of blueprints that would exit the economy at the boundaries of the state space. These exits would occur in one of two cases: first, when a product with the minimum quality $\hat{q}_{\min}$ depreciates; and second, when a product with one of the m highest possible qualities $\hat{q} \in \{\hat{q}_{\max} - (m-1)\log\lambda/m, \ldots, \hat{q}_{\max} - \log\lambda/m, \hat{q}_{\max}\}$ gets hit with a quality innovation. The former is resolved by assuming there is no depreciation at this lower bound, meaning

---

[21]The vast majority of elements in C are zero, so it is stored numerically as a sparse matrix for efficiency.

$g_Q/(\log \lambda/m)$ must be subtracted from the top left element of every C_{ss} matrix. The latter is resolved by assuming that the quality of maximal-quality blueprints cannot be further improved, which means δ must be subtracted from the m bottom-right diagonal element of every C_{ss} matrix.

E.1.7 Convergence criteria and final computations

The convergence criterion is that the entire distribution $f_{i_1 i_2 \hat{q}}$ has converged between iterations. I implement this by checking whether

$$||\mathbf{f}^k - \mathbf{f}^{k-1}|| < \epsilon_{\text{outer}}^{\text{tol}}, \tag{E.24}$$

where $|| \cdot ||$ is the Euclidean norm and $\epsilon_{\text{outer}}^{\text{tol}}$ is a very small tolerance threshold (I choose $\epsilon_{\text{outer}}^{\text{tol}} = 10^{-10}$). If this condition is satisfied, then the iteration is complete. Otherwise, use the distribution $\mathbf{f}^k$ to recompute the aggregates $\{\bar{a}^k, \mathcal{M}_{\eta-1}^k\}$ and begin iteration $k+1$.

E.1.8 Solution without population and variety growth

While this entire paper assumes positive population growth ($g_L > 0$) and new-variety creation ($\alpha < 1$), it nests the special case without population or variety growth ($g_L = 1 - \alpha = 0$), which is the assumption in a large class of endogenous-growth models. The above solution method will fail in this case for two reasons, requiring two numerical changes:

1. The ratio L_t/N_t will be indeterminate, so we need to set it to some exogenous initial value and omit the step of solving for LN in the inner loop. Some papers assume $L_t/N_t = 1$, but this need not be the case. This will imply $g_N = g_L = 0$ and $g_Q > 0$.

2. The matrix inversion that allows us to solve for $\mathbf{f}$ will no longer work numerically, because the right-hand-side vector will equal a vector of zeros and the matrix C will be singular. The solution to this is proposed in Achdou et al. (2022): set an arbitrary element (say, element j) of the zero vector equal to 0.01, then set the jth diagonal of matrix C equal to one and every other element in the jth row equal to zero (thus fixing $f(j) = 0.01$). Then, after inverting this system, re-normalize the solution to this

A.72

system so that it sums to one.

E.2 Balanced-growth path: Simulation of firm distribution

The distribution of firm sales, profits, valuations, and other variables is computed via simulation. Specifically, using the fact that we know the distribution of type and age $m_i(\tau)$, moments can be calculated in two steps. First, simulate a single cohort of firms over their lifecycle from $\tau = 0$ onward, computing conditional moments at each τ. Then, use $m_i(\tau)$ to compute unconditional moments after the simulation is complete.

E.2.1 Dimension reduction and state space

While Appendix D shows that we can analytically solve for the firm distribution over the states $\{a_i, n, \tau\}$, this distribution does not contain enough information to infer the distributions of sales, profits, or valuations. For that, we also need to know the distribution over possible combinations of blueprint characteristics $\{i_{2j}, \hat{q}_{jt}\}_{j=1}^n$ for every n and a_i. The dimensionality of this space is far too large to solve numerically: for a grid of n with $n_{\max}$ points and a grid of $\hat{q}$ with $2K+1$ points, there are $\sum_{n=1}^{n_{\max}} 2(3(2K+1))^n$ possibilities. To resolve this problem, I first show that we can collapse a firm's quality into just three sufficient statistics, then approximate the distribution of these statistics over the firm lifecycle via simulation.

Each firm has $n = \sum_{i_2 \in \{h, \ell, \varnothing\}} n_{i_2}$ blueprints. Define the average quality of blueprints within the firm with followers i_2 as

$$\tilde{q}_{i_2} \equiv \frac{1}{n_{i_2}} \sum_{j=1}^{n_{i_2}} e^{(\eta-1)\hat{q}_j}. \tag{E.25}$$

We will show that $\{n_{i_2}, \tilde{q}_{i_2}\}_{i_2 \in \{h, \ell, \varnothing\}}\}$ is sufficient information for characterizing the full distribution of sales, profits, R&D and M&A expenditures, cash flows, and market values. The rescaled revenues of firm i are

$$\frac{p_{it} y_{it}}{Y_t / L_t} = \frac{L_t}{N_t} \left(\frac{a_{i_1}}{\bar{a}}\right)^{\eta-1} \sum_{i_2 \in \{h, \ell, \varnothing\}} n_{i_2} \tilde{q}_{i_2} \left(\frac{\mu_{i_1 i_2}}{\mathcal{M}_{\eta-1}}\right)^{-(\eta-1)}. \tag{E.26}$$

A.73

The firm-level markup is

$$\mu_{it} = \left(\frac{p_{ijt}y_{ijt}}{p_{it}y_{it}} \mu_{ijt}^{-1} \right)^{-1} = \frac{\sum_{i_2 \in \{h,\ell,\varnothing\}} n_{i_2} \tilde{q}_{i_2} \mu_{i_1 i_2}^{-(\eta-1)}}{\sum_{i_2 \in \{h,\ell,\varnothing\}} n_{i_2} \tilde{q}_{i_2} \mu_{i_1 i_2}^{-\eta}}. \tag{E.27}$$

and hence rescaled profits are

$$\hat{\Pi}_{it} = \left(1 - \frac{1}{\mu_{it}} \right) \frac{p_{it}y_{it}}{Y_t/L_t}. \tag{E.28}$$

From this it is straightforward to see that cashflows equal $\hat{D}_{it} = \hat{\Pi}_{it} - \hat{w}_x x_i$, and that rescaled firm value equals

$$\hat{V}_{it} = \sum_{i_2 \in \{h,\ell,\varnothing\}} n_{i_2} (v_{i_1 i_2}^{\mathrm{P}} \tilde{q}_{i_2} + v_{i_1}^{\mathrm{G}}). \tag{E.29}$$

These expressions confirm that the states $\{a_i, \{n_{i_2}, \tilde{q}_{i_2}\}_{i_2 \in \{h,\ell,\varnothing\}}\}$ are sufficient information for the variables in which we are interested.

The simulation exploits the fact that we know the distribution of ages and types $m_i(\tau)$ in closed-form. This means that, for each a_i, we can simulate a single cohort of firms from inception ($\tau = 0$) to some terminal age ($\tau = \tau_{\max}$), keeping track of moments conditional on $\{a_i, \tau\}$, and then use $m_i(\tau)$ to compute unconditional moments after simulating. I discretize the age grid with equal time increments Δt:

$$\tau \in \{0, \Delta t, \dots, \tau_{\max} - \Delta t, \tau_{\max}\}, \tag{E.30}$$

where the maximum age $\tau_{\max}$ is chosen such that the truncated mass of firms above $\tau_{\max}$, $1 - \int_{\tau_{\max}}^{\infty} m(\tau)d\tau$, is close to zero. For convenience, I re-index age τ to time t and let time start at $t = 0$. The simulation begins with M_{sim} firms for each a_i. At each time step, I store the firms' characteristics $\{n_{i_2 t}, \tilde{q}_{i_2 t}\}_{i_2 \in \{h,\ell,\varnothing\}}$.

E.2.2 Initialization

At the beginning of its lifecycle ($\tau = 0$), each firm begins with a single product $n_{i_2} = 1$ for some i_2 that has its own initial relative quality $\hat{q}_{i_2 0}$. A share $\Phi_{ni}/(\Phi_{ni} + \Phi_{qi})$ of these entrants will begin with a blueprint of type $i_2 = \varnothing$, for which the relative quality is $\hat{q}_{\varnothing 0} = \log \lambda$. A

share $\Phi_{\text{n}i} f_h / (\Phi_{\text{n}i} + \Phi_{\text{q}i})$ will possess type $i_2 = h$ with relative quality $\hat{q}_{h0} \sim f_{hi_2, \hat{q} - \log \lambda} / f_h$. And the remaining entrants will have type $i_2 = \ell$ with analogous quality. Numerically, to be able to draw $\hat{q}_{i_2 0}$ from the given quality distributions, I create piecewise CDFs over the discretized support for $\hat{q}$ and invert simulated qualities from a set of uniform random variables. Each $\hat{q}_{i_2 0}$ implies a corresponding starting point $\{\tilde{q}_{i_2 0}\}$, with $\tilde{q}_{i_2 0} = 0$ for those i_2 with $n_{i_2} = 0$. For each firm, we store the values $\{\tilde{q}_{i_2 0}, n_{i_2 0}\}_{i_2}$.

E.2.3 Simulation of firm characteristics

At each time step $t > 0$, we simulate the evolution of blueprint counts and quality statistics. At time $t + \Delta t$, the simulated number of monopoly products for a firm with $n_{\varnothing t}$ current blueprints is

$$n_{\varnothing, t + \Delta t} = \max\{n_{\varnothing t} + J^+_{\varnothing, t + \Delta t} - J^-_{\varnothing, t + \Delta t}, 0\}, \tag{E.31}$$

where

$$J^+_{\varnothing, t + \Delta t} \sim \text{Poisson}((\Phi_{\text{n}i} + \mathbb{I}_{i_1 = h} \Psi_h f_{\ell\varnothing}) n_t \Delta t), \tag{E.32}$$

is the number of type-$\varnothing$ blueprints added and

$$J^-_{\varnothing, t + \Delta t} \sim \text{Poisson}((\delta + \mathbb{I}_{i_1 = \ell} \Psi_h f_h) n_{\varnothing t} \Delta t), \tag{E.33}$$

is the number of type-$\varnothing$ blueprints lost. Changes in n_h and n_ℓ are simulated accordingly. The average quality of the portfolio of monopoly products then evolves approximately according to the discretized law of motion

$$\tilde{q}_{\varnothing, t + \Delta t} = \frac{1}{n_{\varnothing, t + \Delta t}} \left(n_{\varnothing t} \tilde{q}_{\varnothing t} e^{-(\eta - 1) g_Q \Delta t} + \tilde{q}^+_{\varnothing t} J^+_{\varnothing, t + \Delta t} - \tilde{q}^-_{\varnothing t} J^-_{\varnothing, t + \Delta t} \right). \tag{E.34}$$

The first term on the right-hand side is the current average quality, after depreciation. The second term is the change in average quality from new blueprints:

$$\tilde{q}^+_{\varnothing t} J^+_{\varnothing, t + \Delta t} \equiv \sum_{j=1}^{J^+_{\varnothing, t + \Delta t}} e^{(\eta - 1) \hat{q}^{\text{new}}_{\varnothing j t}}, \tag{E.35}$$

where the new blueprints' qualities are drawn independently from the distribution

$$
\hat{q}_{\varnothing jt}^{\text{new}} \sim \begin{cases} \mathbb{I}_{\hat{q}=\log \lambda} & \text{with probability } \Phi_{\text{n}i}/(\Phi_{\text{n}i} + \mathbb{I}_{i=h}\Psi_h f_{\ell\varnothing}), \\ f_{\ell\varnothing\hat{q}|\ell\varnothing} & \text{with probability } \mathbb{I}_{i=h}\Psi_h f_{\ell\varnothing}/(\Phi_{\text{n}i} + \mathbb{I}_{i=h}\Psi_h f_{\ell\varnothing}). \end{cases} \tag{E.36}
$$

The third and final term in (E.34) captures the change from lost blueprints, with $\tilde{q}_{\varnothing t}^{-}$ representing the average quality of those blueprints. The exact distribution from which $\tilde{q}_{\varnothing t}^{-}$ is drawn requires knowledge of the entire current quality distribution $\{\hat{q}_{\varnothing jt}\}_{j=1}^{n_{\varnothing t}}$. Because it is infeasible to keep track of all of these qualities (as explained above), I approximate the average quality of a lost blueprint with the average quality of all blueprints: $\tilde{q}_{\varnothing t}^{-} \approx \tilde{q}_{\varnothing t}$. This approximation has two appealing properties. First, it means that the change in total quality from product loss will on average be correct for all firms. Second, the approximation will be nearly exact for firms with large n (and these large firms get the most weight in macroeconomic moments). I confirm that this approximation does not introduce significant error by verifying that simulation-implied moments that depend on the quality distribution do not differ significantly from the exact equilibrium moments computed before the simulation.

The procedure for simulating the blueprint counts and average qualities for the other two types $i_2 \in \{h, \ell\}$ is analogous to this one, with the appropriate changes to the Poisson intensities and the quality distributions.

E.2.4 Computation of moments

As shown above, the set $\{a_i, \{n_{i_2}, \tilde{q}_{i_2}\}_{i_2}\}$ is sufficient information to compute a firm's (rescaled) sales, profits, and value. Any moment (mean, covariance, etc.) of any function these variables at time step t (age τ) can be expressed in terms of expected values (means). For example, to compute the cross-sectional variance of firm markups, we can use the fact that

$$
\text{var}(\mu_{it}|a_i, \tau) = \mathbb{E}[\mu_{it}^2|a_i, \tau] - \mathbb{E}[\mu_{it}|a_i, \tau]^2, \tag{E.37}
$$

computing and storing each of these two expectations at each iteration of the simulation. Then, after the simulation, we can use the distribution $m_i(\tau)$ to compute

$$\mathbb{E}[\mu_{it}] = \sum_{i \in \{h,\ell\}} \int_0^\infty m_i(\tau) \mathbb{E}[\mu_{it}|a_i, \tau] d\tau. \tag{E.38}$$

and

$$\text{var}(\mu_{it}) = \sum_{i \in \{h,\ell\}} \int_0^\infty m_i(\tau) \left(\text{var}(\mu_{it}|a_i, \tau) + \mathbb{E}[\mu_{it} - \mathbb{E}[\mu_{it}]|a_i, \tau]^2 \right) d\tau, \tag{E.39}$$

The same logic applies to any moment of the firm distribution in which we are interested.

E.3 Transition path: Equilibrium solution method

E.3.1 State space and initial and terminal conditions

Starting at some initial time t_0, a transition-path equilibrium begins with a blueprint distribution $\{f_{i_1 i_2 \hat{q} t_0}\}$ and ratio L_{t_0}/N_{t_0}. After t_0, it may feature any arbitrary time path of model parameters, provided that those parameters eventually stabilize at some constant values (and are consistent with a well-defined solution). This appendix shows the numerical solution for declining innovation productivity φ_t: the economy transitions from an initially high innovation productivity $\varphi_{t_0} = \varphi^+$ to a steady state with low productivity $\lim_{t \to \infty} \varphi_t = \varphi^-$. Extending to other parameters just requires adding t subscripts to those parameters. Because every transition-path equilibrium converges to a balanced-growth equilibrium, the terminal condition for the solution method is the terminal balanced-growth path.

The relevant state space for a transition path includes both the balanced-growth states $\{i_1, i_2, \hat{q}\}$ and the new time dimension $t \geq t_0$. Numerically, we can discretize this time grid into increments Δt as

$$t \in \{t_{\min}, t_{\min} + \Delta t, \ldots, t_{\max} - \Delta t, t_{\max}\}, \tag{E.40}$$

where $t_{\min} = t_0$ and $t_{\max}$ is sufficiently large that the economy has converged near the new steady state. The solution method then takes as its terminal condition ($t = t_{\max}$) the balanced-growth equilibrium, solved as described in Appendix E.1 above.

E.3.2 Overview of solution algorithm

The algorithm is a backward-forward iteration scheme:

1. Starting from the final steady-state value functions $\{v^{\mathrm{P}}_{i_1 i_2 \hat{q} t_{\max}}\}$ and $\{v^{\mathrm{G}}_{i_1 t_{\max}}\}$, solve the HJBs backward to $t = t_{\min}$, re-computing R&D and M&A policies at each time step.

2. Starting from the initial steady-state distribution $f_{i_1 i_2 \hat{q} t_{\min}}$, solve the KFEs forward to $t = t_{\max}$, re-solving for profits, aggregates (including L_t/N_t), and R&D and M&A policies at each time step.

3. Check if $\{f_{i_1 i_2 \hat{q} t}\}$ has converged at all points in the the state space $\{i_1, i_2, \hat{q}, t\}$; if not, return to step 1.

E.3.3 Initialization

To initiate this algorithm, we need an initial guess for the path of the distribution over time t. The solution is not particularly sensitive to this choice. For simplicity, I start with the stationary distributions under $\varphi_{t_{\min}}$ and $\varphi_{t_{\max}}$ ($f^0_{i_1 i_2 \hat{q} t_{\max}}$ and $f^0_{i_1 i_2 \hat{q} t_{\max}}$, respectively), then interpolate them along the time grid as if they evolve in proportion to the parameter φ_t:

$$f^0_{i_1 i_2 \hat{q} t} = f^0_{i_1 i_2 \hat{q} t_{\min}} + \frac{\varphi_{t_{\min}} - \varphi_t}{\varphi_{t_{\min}} - \varphi_{t_{\max}}}(f^0_{i_1 i_2 \hat{q} t_{\max}} - f^0_{i_1 i_2 \hat{q} t_{\min}}) \tag{E.41}$$

To solve the initial HJB equations, we need a sequence $\{\Phi_{\mathrm{n} i_1 t}, \Phi_{\mathrm{q} i_1 t}, \Psi_{i_1 t}, LN_t\}$ from $t = t_{\min}$ to $t = t_{\max} - \Delta t$. I therefore similarly interpolate $\{v^{\mathrm{P}}_{i_1 i_2 \hat{q} t}\}$ and $\{v^{\mathrm{G}}_{i_1 t}\}$, then use these interpolated series to compute an initial sequence $\{\Phi^0_{\mathrm{n} i_1 t}, \Phi^0_{\mathrm{q} i_1 t}, \Psi^0_{i_1 t}, LN^0_t\}$.

E.3.4 Solving HJBs backwards

HJBs are solved backward, starting from the known terminal values $\{v^{\mathrm{P}}_{i_1 i_2 \hat{q} t_{\max}}\}$ and $\{v^{\mathrm{G}}_{i_1 t_{\max}}\}$. The HJB equation for product values is

$$(\rho + \delta_t + (\eta - 1)g_{Qt})v^{\mathrm{P}}_{i_1 i_2 \hat{q} t} = \pi_{i_1 i_2 \hat{q} t} + \mathbb{I}_{i_1 = \ell} \Psi_{ht} f_{ht}(1 - \varrho)(v^{\mathrm{P}}_{h i_2 \hat{q} t} - v^{\mathrm{P}}_{\ell i_2 \hat{q} t}) + \frac{\partial}{\partial t} v^{\mathrm{P}}_{i_1 i_2 \hat{q} t}. \tag{E.42}$$

Approximating the time derivative with a forward difference implies, at iteration k,

$$v_{i_1 i_2 \hat{q} t}^{P,k} = \frac{v_{i_1 i_2 \hat{q}, t+\Delta t}^{P,k}/\Delta t + \pi_{i_1 i_2 \hat{q} t}^{k-1} + \mathbb{I}_{i_1=\ell} \Psi_{ht}^{k-1} f_{ht}^{k-1}(1-\varrho) v_{h i_2 \hat{q} t}^{P,k}}{1/\Delta t + \rho + \delta_t^{k-1} + (\eta-1) g_{Qt}^{k-1} + \mathbb{I}_{i_1=\ell} \Psi_{ht}^{k-1} f_{ht}^{k-1}(1-\varrho)}. \tag{E.43}$$

Like in Appendix E.1, the value of growth opportunities $v_t^G = [v_{ht}^G \quad v_{\ell t}^G]^\top$ solves the two-by-two linear system

$$C_t^{G,k} v_t^{G,k} = \tilde{v}_t^{G,k}, \tag{E.44}$$

where (suppressing k superscripts for the moment)

$$\tilde{v}_t^G = \begin{bmatrix} v_{h,t+\Delta t}^G/\Delta t + (1-\varepsilon)(\Phi_{nht} \bar{v}_{nht}^P + \Phi_{qht} \bar{v}_{qht}^P + \Psi_{ht} f_{\ell t} \bar{v}_{mht}^P) \\ v_{\ell,t+\Delta t}^G/\Delta t + (1-\varepsilon)(\Phi_{n\ell t} \bar{v}_{n\ell t}^P + \Phi_{q\ell t} \bar{v}_{q\ell t}^P) \end{bmatrix} \tag{E.45}$$

and where

$$C_t^G = \begin{bmatrix} c_{hht} & c_{h\ell t} \\ c_{\ell ht} & c_{\ell\ell t} \end{bmatrix}, \tag{E.46}$$

with entries

$$c_{hht} = 1/\Delta t + \rho + \delta_t - (1-\varepsilon)(\Phi_{nht} + \Phi_{qht} + \Psi_{ht} f_{\ell t} \varrho), \tag{E.47}$$

$$c_{h\ell t} = (1-\varepsilon)\Psi_{ht} f_{\ell t} \varrho, \tag{E.48}$$

$$c_{\ell ht} = -\Psi_{ht} f_{ht}(1-\varrho), \tag{E.49}$$

$$c_{\ell\ell t} = 1/\Delta t + \rho + \delta_t - (1-\varepsilon)(\Phi_{n\ell t} + \Phi_{q\ell t}) + \Psi_{ht} f_{ht}(1-\varrho). \tag{E.50}$$

Inverting this system gives us $\{v_{i_1 t}^{G,k}\}$.

E.3.5   Solving KFEs forward

After solving HJBs backward, we solve forward from $t_{\min}$ (i.e., from $LN_{t_{\min}}$ and $f_{i_1 i_2 \hat{q} t_{\min}}$), taking the path of blueprint values as given. At each time step, compute the wage $\hat{w}_{xt}^k$ and policies $\{x_{ni_1 t}^k, x_{qi_1 t}^k, x_{mi_1 t}^k\}$. These imply intensities $\{\Phi_{ni_1 t}^k, \Phi_{qi_1 t}^k, \Psi_{i_1 t}^k\}$ and δ_t^k and growth

rates g_{Nt}^k and g_{Qt}^k. Using these growth rates, we have

$$LN_{t+\Delta t}^k = LN_t^k \exp\{(g_L - g_{Nt}^k)\Delta t\} \tag{E.51}$$

Next, we solve forward for the total blueprint share $f_{h,t+\Delta t}^k$ using the aggregated KFE

$$\frac{\partial}{\partial t} f_{ht} + g_{Nt} f_{ht} = (\Phi_{nht} + \Phi_{qht})(\nu_{ht} + f_{ht}) + \Psi_{ht} f_{ht} f_{\ell t} - \delta_t f_{ht}. \tag{E.52}$$

Discretizing this expression over time with a forward difference implies that the distribution at time $t + \Delta t$ in iteration k equals

$$f_{h,t+\Delta t}^k = f_{ht}^k + \left((\Phi_{nht}^k + \Phi_{qht}^k)(\nu_h^k + f_{ht}^k) + \Psi_{ht}^k f_{ht}^k (1 - f_{ht}^k) - (\delta_t^k + g_{Nt}^k) f_{ht}^k\right) \Delta t. \tag{E.53}$$

Finally, by the same reasoning as for the total blueprint shares, we can discretize the full system of KFEs over the grids for t and $\hat{q}$ and solve forward in similar fashion. As in the balanced-growth case, we must make two numerical adjustments to ensure products are not lost at the boundaries of the quality grid. First, we must eliminate the loss from quality depreciation at $\hat{q}_{\min}$; second, we must eliminate the loss from quality improvements in the uppermost set of qualities. See Appendix E.1 for specifics.

For the distribution of monopoly products ($\{i_1 = i, i_2 = \varnothing\}$), we have

$$f_{i\varnothing\hat{q},t+\Delta t}^k = f_{i\varnothing\hat{q}t}^k + \left((f_{i\varnothing,\hat{q}+\log\lambda/m,t}^k - \mathbb{I}_{\hat{q}>\hat{q}_{\min}} f_{i\varnothing\hat{q}t}^k)\frac{g_{Qt}}{\log\lambda/m} + \Phi_{nit}^k(\nu_{it}^k + f_{it}^k)\mathbb{I}_{\hat{q}=\log\lambda} \right.$$
$$\left. + \mathbb{I}_{i=h}\Psi_{ht}^k f_{ht}^k f_{l\varnothing\hat{q}t}^k - ((1 - \mathbb{I}_{\hat{q}>\hat{q}_{\max}-\log\lambda})\delta_t^k + \mathbb{I}_{i=\ell}\Psi_{ht}^k f_{ht}^k + g_{Nt}^k) f_{i\varnothing\hat{q}t}^k \right) \Delta t. \tag{E.54}$$

For the distribution of same-type competitors ($i_1 = i_2 = i$), we have

$$f_{ii\hat{q},t+\Delta t}^k = f_{ii\hat{q}t}^k + \left((f_{ii,\hat{q}+\log\lambda/m,t}^k - \mathbb{I}_{\hat{q}>\hat{q}_{\min}} f_{ii\hat{q}t}^k)\frac{g_{Qt}}{\log\lambda/m}\right.$$
$$+ \Phi_{qit}^k(\nu_{it}^k + f_{it}^k)[f_{ih,\hat{q}-\log\lambda,t}^k + f_{i\ell,\hat{q}-\log\lambda,t}^k + f_{i\varnothing,\hat{q}-\log\lambda,t}^k]$$
$$\left. + \mathbb{I}_{i=h}\Psi_{ht}^k f_{ht}^k f_{\ell h\hat{q}t}^k - ((1 - \mathbb{I}_{\hat{q}>\hat{q}_{\max}-\log\lambda})\delta_t^k + \mathbb{I}_{i=\ell}\Psi_{ht}^k f_{ht}^k + g_{Nt}^k) f_{ii\hat{q}t}^k \right) \Delta t. \tag{E.55}$$

A.80

For $i_1 = h$ and $i_2 = \ell$, we have

$$f^k_{h\ell\hat{q},t+\Delta t} = f^k_{h\ell\hat{q}t} + \left((f^k_{h\ell,\hat{q}+\log\lambda/m,t} - \mathbb{I}_{\hat{q}>\hat{q}_{\min}} f^k_{h\ell\hat{q}t}) \frac{g_{Qt}}{\log\lambda/m} \right.$$

$$+ \Phi^k_{\mathrm{q}ht}(\nu^k_{ht} + f^k_{ht})[f^k_{\ell\ell,\hat{q}-\log\lambda,t} + f^k_{\ell h,\hat{q}-\log\lambda,t} + f^k_{\ell\varnothing,\hat{q}-\log\lambda,t}]$$

$$\left. + \Psi^k_{ht} f^k_{ht} f^k_{\ell\ell\hat{q}t} - ((1 - \mathbb{I}_{\hat{q}>\hat{q}_{\max}-\log\lambda})\delta^k_t + g^k_{Nt}) f^k_{h\ell\hat{q}t} \right) \Delta t. \quad \text{(E.56)}$$

And for $i_1 = \ell$ and $i_2 = h$, we have

$$f^k_{\ell h\hat{q},t+\Delta t} = f^k_{\ell h\hat{q}t} + \left((f^k_{\ell h,\hat{q}+\log\lambda/m,t} - \mathbb{I}_{\hat{q}>\hat{q}_{\min}} f^k_{\ell h\hat{q}t}) \frac{g_{Qt}}{\log\lambda/m} \right.$$

$$+ \Phi^k_{\mathrm{q}\ell t}(\nu^k_{\ell t} + f^k_{\ell t})[f^k_{hh,\hat{q}-\log\lambda,t} + f^k_{h\ell,\hat{q}-\log\lambda,t} + f^k_{h\varnothing,\hat{q}-\log\lambda,t}]$$

$$\left. - ((1 - \mathbb{I}_{\hat{q}>\hat{q}_{\max}-\log\lambda})\delta^k_t + \Psi^k_{ht} f^k_{ht} + g^k_{Nt}) f^k_{\ell h\hat{q}t} \right) \Delta t. \quad \text{(E.57)}$$

Finally, at each time step in this iteration, we can re-compute the macroeconomic aggregates that will be needed for the next iteration of HJBs: the aggregates $\{\bar{a}^k_t, \mathcal{M}^k_{\eta-1,t}\}$, from which we get profits $\pi^k_{i_1 i_2 t}$.

E.3.6 Convergence criteria and final computations

The convergence criterion is that the entire distribution $f_{i_1 i_2 \hat{q}t}$ has converged over the entire time path between iterations. I implement this by checking whether the average error at any given point in time

$$\frac{||\mathbf{f}^k - \mathbf{f}^{k-1}||}{\dim(t)} < \epsilon^{\mathrm{tol}}_{\mathrm{outer}}, \quad \text{(E.58)}$$

where $||\cdot||$ is the Euclidean norm, $\dim(t) \equiv 1 + (t_{\max} - t_{\min})/\Delta t$ is the number of points in the time grid, and $\epsilon^{\mathrm{tol}}_{\mathrm{outer}}$ is the same (very small) tolerance threshold used in the balanced-growth solution algorithm. If this condition is satisfied, then the iteration is complete. Otherwise, we reinitiate at iteration $k+1$.

E.4 Transition path: Simulation of firm distribution

The simulation of firms along the transition path consists of two steps: first, simulate a panel of firms along the balanced-growth path; and then, simulate that panel of firms forward under the transition-path equilibrium, adding a new cohort of firms at each iteration.

In the first step, we can first simulate a cohort of M_{sim} firms using the simulation method in Appendix E.2 above. Then, for each type a_i and age $\tau \in \{0, \ldots, \tau_{\max} - \Delta\tau, \tau_{\max}\}$, we keep $M_{\text{sim}}(a_i, \tau) = \lfloor M_{\text{sim}} \frac{m_i(\tau)}{m(\tau=0)} \rfloor$ of them, giving each a unique firm identifier, so that the distribution of firms in the panel reflects the true theoretical distribution $m_i(\tau)$.

In the second step, we simulate this panel forward in time $t \in \{t_0, \ldots, t_{\text{end}} - \Delta t, t_{\text{end}}\}$. The method for simulating a firm forward and for drawing a cohort of new firms is, again, the same as in Appendix E.2 above, the only difference being that at each point in time we now use the updated equilibrium path of the distribution, aggregates, and firm policies and values. To determine the number of new firms in each cohort, we need to compute the evolution of the entry rate (D.21). At the beginning of the transition ($t = t_0$), we have a constant ratio N_t/M_t and constant entry and exit rates $\mathcal{E}_{\text{entry}}$ and $\mathcal{E}_{\text{exit}}$ from the balanced-growth path. Over the transition, the exit rate $\mathcal{E}_{\text{exit},t+\Delta t}$ equals the number of exiting firms divided by the total number of firms (per unit of time Δt). The number of blueprints per firm evolves as

$$\frac{N_{t+\Delta t}}{M_{t+\Delta t}} = \frac{N_t}{M_t} \exp\{(g_{Nt} - (\mathcal{E}_{\text{entry},t} - \mathcal{E}_{\text{exit},t}))\Delta t\}. \tag{E.59}$$

We can then use this ratio to compute the entry rate $\mathcal{E}_{\text{entry},t+\Delta}$, which determines the number of new firms that must be simulated.

F   Details of model estimation

F.1 Estimation procedure

I estimate the parameters by simulated method of moments (SMM). In the first step, I search for the parameters that minimize the distance between the model-implied moments and data moments. Recall that Θ denotes the vector of estimated parameters. Let $\hat{m}$ denote the vector of empirical moments and $m(\Theta)$ the corresponding moments in the model. Let Σ

denote the sample covariance matrix of moments $\hat{m}$. Then the SMM estimator is defined as

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}} \left\{ (m(\Theta) - \hat{m})^\top \Sigma (m(\Theta) - \hat{m}) \right\}. \tag{F.1}$$

I find this minimum-distance estimate using the numerical approach recently proposed by Catherine et al. (2023), which has three steps.[22] First, I solve and simulate the model for 3000 pseudorandom combinations of parameters generated from a Halton sequence. Second, using the resulting dataset, I fit a cubic polynomial (with interactions) between parameters and moments. To increase precision around the correct estimate, I apply OLS weights proportional to the inverse squared error of each point. Third, using this fitted model, I apply the Nelder-Mead simplex algorithm to locally search for the minimum-distance estimate $\hat{\Theta}$.

I compute standard errors of moments using the following asymptotic results. Let J denote the Jacobian matrix of moments with respect to parameters: for moment $m_j \in m(\Theta)$ and parameter $\theta_k \in \Theta$, element (j, k) of the Jacobian equals

$$J_{jk} = \frac{\partial m_j}{\partial \theta_k}. \tag{F.2}$$

Standard asymptotic results imply that the (estimated) covariance matrix of parameters is given by

$$\hat{\Omega} = (J^\top \Sigma^{-1} J)^{-1}. \tag{F.3}$$

Standard errors are then given by the square root of the diagonal elements of $\hat{\Omega}$:

$$\operatorname{se}(\hat{\Theta}) = \sqrt{\operatorname{diag}(\hat{\Omega})}. \tag{F.4}$$

Numerically, I approximate the derivatives in J using a double-sided finite difference with increments equal to 1% of each estimated parameter value (Judd, 1998). I discuss the computation of the matrix Σ next.

In the estimation of the balanced-growth path, Θ is the vector of parameters (57) and $\hat{m}$ and $m(\Theta)$ the corresponding data moments. In the transition estimation, Θ is the vector

---

[22]The benefit of this approach is that the computationally intensive step (step one) does not depend on the empirical moments, so robustness checks with respect to moments can be performed at low cost.

of parameters $\{\kappa_\varphi, t_{\min}, \varphi_\infty\}$ governing (66) and the moments are the 45 values of aggregate R&D-to-value $\{w_{xt}X_{rt}/V_t\}$ from 1975–2020.

F.2   Computation of empirical moments and standard errors

As explained in Section 4.2, I estimate seven parameters on the balanced-growth path from seven moments in micro and macro data. The moments themselves are computed as described in the main text. The five moments depending on value, sales, profits, R&D, and M&A are computed using CRSP/Compustat merged with SDC Platinum. Profits in profit-to-sales ratios are defined as sales minus cost of goods sold. Value-to-sales and profit-to-sales ratios are computed using data from 1965–1975. R&D-to-value and M&A-to-value must be computed after this because, as explained in Appendix A, R&D reporting begins in 1975 and M&A transaction data begin in 1978.

Output growth is the annualized percentage growth rate of real per capita quarterly GDP, as reported by the BEA. Specifically, from year t to $t+0.25$ (i.e., across quarters), annualized growth equals $g_{y,t+0.25} = (y_{t+0.25}/y_t)^4 - 1$. The firm entry rate in year t is computed from the Business Dynamics Statistics (BDS) as the ratio of the number of entrants from year $t-1$ to year t (computed as the net change in the number of firms from $t-1$ to t plus the number of firm deaths from $t-1$ to t) divided by the number of firms in year $t-1$.

Finally, for the sales-weighted standard deviations of value-to-sales and profit-to-sales, I rescale each variable by the corresponding aggregate value to remove level effects. For example, for value-to-sales, the moment (in the model notation) is equal to

$$\frac{\sqrt{\mathrm{var}_{py}(V_{it}/(p_{it}y_{it}))}}{V_t/Y_t} = \sqrt{\int_0^{M_t} \frac{p_{it}y_{it}}{Y_t} \left(\frac{V_{it}/(p_{it}y_{it})}{V_t/Y_t} - 1\right)^2 di}. \tag{F.5}$$

This rescaling removes the effect of the level of the aggregate ratios, which are mainly affected by other parameters. To ensure that the model captures the same amount of valuation dispersion as in the empirical decompositions, I do the same winsorizing of the cross-sectional data in each year (10th and 90th percentiles) as in the decompositions. Winsorizing also has the benefit of removing extreme dispersion in valuations that is caused by transitory shocks outside of the model (e.g., discount-rate shocks), which would lead to overestimates of the

productivity gap a_h/a_ℓ and therefore overstate the effect of reallocation on the transition.

Because some of the moments must use non-overlapping samples or time periods, I compute the empirical covariance matrix of moments Σ in diagonal blocks, setting the unavailable off-diagonal elements to zero.[23] For aggregate value-to-sales and the two cross-sectional standard deviations, I estimate the 3×3 covariance matrix by bootstrap. Specifically, I sample firms from Compustat in each year from 1965–1975, 1000 times, recomputing the moments each time, then compute the covariance of these sampled moments. Likewise, for aggregate R&D-to-value and M&A-to-value, I estimate a 2×2 matrix from 1975–1985. Finally, the diagonal elements corresponding to output growth and the firm entry rate (which are aggregated data) equal the respective squared standard errors.

F.3   Additional details on identification of parameters

F.3.1   Identification of balanced-growth path

As further evidence of identification, Figure F.9 plots the effect of changing model parameters on the corresponding identifying moments in the estimation of the balanced-growth equilibrium. In each case, the parameter on the x-axis is changed and all other parameters are held fixed at their estimated values. The direction of these local changes confirm the intuition for identification given in the main text.

F.3.2   Identification of transition path

Section 4.3 presents the equilibrium R&D-to-value ratio under the special case of inelastic labor supply $(\zeta = 0)$. Here I derive this in the more general case of elastic labor supply $(\zeta \geq 0)$. The total amount of R&D spending equals

$$w_{\mathrm{xt}} X_{\mathrm{rt}} = w_{\mathrm{xt}} N_t \sum_{i \in \{h,\ell\}} f_{it}(x_{\mathrm{n}it} + x_{\mathrm{q}it}) \tag{F.6}$$

$$= w_{\mathrm{xt}}^{-\frac{\varepsilon}{1-\varepsilon}} (\varepsilon\varphi_t)^{\frac{1}{1-\varepsilon}} N_t \left(\sum_{i \in \{h,\ell\}} f_{it} \left(((1-\alpha)\bar{v}_{\mathrm{n}it})^{\frac{1}{1-\varepsilon}} + (\alpha\bar{v}_{\mathrm{q}it})^{\frac{1}{1-\varepsilon}} \right) \right). \tag{F.7}$$

---

[23]An alternative approach could be to look at "worst-case" entries for the unknown elements of the covariance matrix, as recently proposed by Cocci and Plagborg-Møller (in press). This would have a minor effect on the main results, which turn out to be very robust to moments.

**Figure F.9: Relationship between parameters and identifying moments**
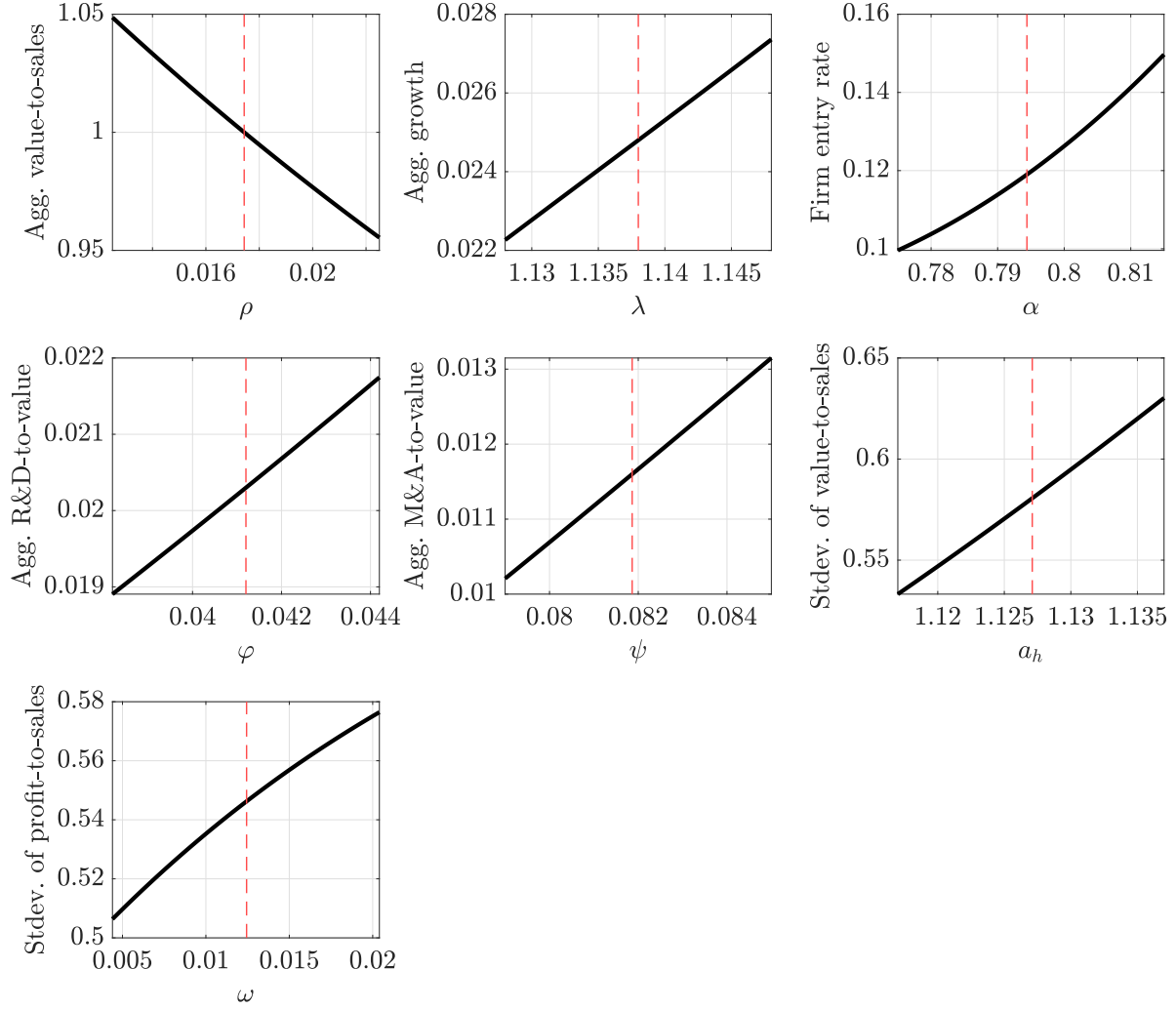
The figure plots the effect of changing each parameter in the model on the corresponding identifying moment. The vertical red lines denote the estimated value. Each panel keeps all other parameters fixed at their estimated values and changes only the parameter on the x-axis.

Substituting in the equilibrium wage and dividing by value then implies

$$\frac{w_{xt}X_{rt}}{V_t} = \underbrace{\varphi^{1+\frac{\zeta\varepsilon}{1+\zeta(1-\varepsilon)}}}_{\text{prod.}} \times \underbrace{\frac{\bar{v}_r^{1+\frac{\zeta\varepsilon}{1+\zeta(1-\varepsilon)}}}{V_t/N_t}}_{\text{values}} \times \underbrace{\varepsilon^{1+\frac{\zeta\varepsilon}{1+\zeta(1-\varepsilon)}}\left(1 + \frac{N_t}{\bar{s}L_t}\left(1 + \left(\frac{\psi}{\varphi}\frac{\bar{v}_m}{\bar{v}_r}\right)^{\frac{1}{1-\varepsilon}}\right)\right)^{-\frac{\varepsilon}{1+\zeta(1-\varepsilon)}}}_{\text{substitution between factors and technologies}},$$

(F.8)

which reduces to the equation in the main text if $\zeta = 0$. Thus, if we again apply the intuition that $\Delta\log\bar{v}_r \approx \Delta\log(V_t/N_t)$ and that the substitution term is insensitive to parameter

A.86

Figure F.10: Sensitivity of R&D-to-value to other parameter shifts
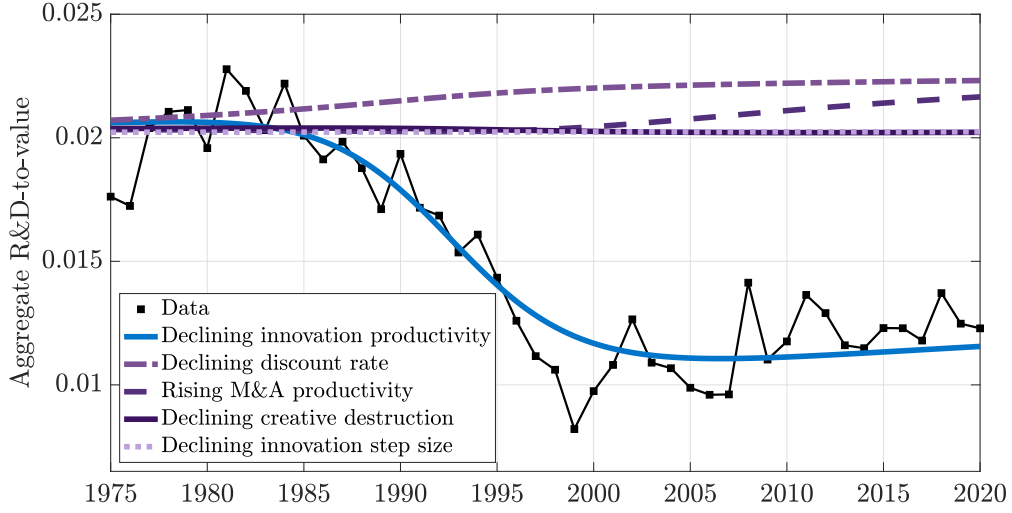
The figure plots the model-implied change in the aggregate R&D-to-value ratio in response to an exogenous change in various model parameters. In each line, only the one parameter changes; all others remain fixed at their pre-transition values. See the text for more details

changes, we have that

$$\Delta \log \frac{w_{\mathrm{x}t} X_{\mathrm{r}t}}{V_t} \approx \left(\frac{1 + \zeta}{1 + \zeta(1 - \varepsilon)} \right) \Delta \log \varphi_t + \left(\frac{\zeta \varepsilon}{1 + \zeta(1 - \varepsilon)} \right) \Delta \log \bar{v}_{\mathrm{r}t}, \tag{F.9}$$

$$= 1.2 \times \Delta \log \varphi_t + 0.2 \times \Delta \log \bar{v}_{\mathrm{r}t}, \tag{F.10}$$

where the second equality substitutes in the calibrated values $\zeta = \varepsilon = 1/2$. There is now a labor-supply effect: if research productivity or valuations rise, then labor supply endogenously increases, so R&D spending goes up slightly more than value. For a decline in φ_t, net labor supply effects turn out to be close to zero, because φ_t falls but $\bar{v}_{\mathrm{r}t}$ rises.

Figure F.10 illustrates this intuition. The figure plots the response of aggregate R&D-to-value to various alternative parameter shifts in the model. These shifts include a decline in discount rates ρ, an increase in M&A productivity ψ, a decline in the creative destruction rate α, and a decline in the innovation step size λ. Some of these have been proposed as explanations for other moments in the data—for example, a decline in discount rates could cause a rise in aggregate valuations, and a decline in innovation size (ideas becoming more incremental) could explain a decline in growth. However, none of these shifts can explain the decline in R&D-to-value. The intuition is as just laid out above. In fact, some of these

A.87

shifts, like the decline in discount rates, even cause R&D-to-value to *increase*, because of the positive labor-supply effects.

F.4 Robustness check: Alternative identification of ω

Here I consider an alternative identification strategy for the firm-heterogeneity parameters $\{a_h/a_\ell, \omega\}$ that instead matches the amount of concentration in the firm distribution. This is useful both as a robustness check and as an illustration of how this model can quantitatively explain concentration. In the baseline estimation, these parameters were jointly identified by the sales-weighted cross-sectional variances of value-to-sales and profit-to-sales, under the argument that these moments were differentially sensitivity to each parameter. An alternative strategy is to identify ω from the Pareto tail of the employment distribution. The right tail of the firm distribution is populated by high-productivity firms; the lower is ω, the fewer high-productivity firms there are, and the greater the number of blueprints in any given high-productivity firm. In other words, a greater share of blueprints becomes concentrated in a smaller number of high-productivity firms.

The Pareto tail coefficient $\theta \in [1, \infty)$ is defined and derived in Appendix D.5. For high levels of employment L_i, the distribution follows $\mathbb{P}(L_i > L) \propto L^{-\theta}$, which corresponds to a density $m(L) \propto L^{-(\theta+1)}$. A lower value of θ represents a thicker Pareto tail (more concentration). I estimate θ in the data from the share of employment in large firms. Consider two (large) values of employment, a lower bound $\underline{L}$ and an upper bound $\bar{L} > \underline{L}$. The share of employment held by firms with more than $\bar{L}$ employees, as a fraction of the total employment of all firms with more than $\underline{L}$ employees, is given by

$$\frac{s_L(\bar{L})}{s_L(\underline{L})} \equiv \frac{\int_{\bar{L}}^\infty L m(L) dL}{\int_{\underline{L}}^\infty L m(L) dL} = \frac{\int_{\bar{L}}^\infty L^{-\theta} dL}{\int_{\underline{L}}^\infty L^{-\theta} dL} = \left(\frac{\bar{L}}{\underline{L}}\right)^{-(\theta-1)}. \tag{F.11}$$

Thus, we can infer the coefficient θ as

$$\theta = 1 - \frac{\log(s_L(\bar{L})/s_L(\underline{L}))}{\log(\bar{L}/\underline{L})}. \tag{F.12}$$

I compute this in the Business Dynamics Statistics, which reports the employment shares

Table F.4: Alternative estimates of parameters for initial balanced-growth path

| | Parameters | | Identifying moments | | |
	Notation	Estimate	Moment	Data	Model
Preferences					
Discount factor	ρ	0.0085 (0.0169)	Agg. value-to-sales	1.0006 (0.0230)	1.0006
Innovation (R&D)					
Innovation productivity	φ_0	0.0415 (0.0041)	Agg. R&D-to-value	0.0204 (0.0008)	0.0204
Quality innovation size	λ	1.1296 (0.0281)	Agg. output growth	0.0248 (0.0040)	0.0248
Relative quality productivity	α	0.7992 (0.0096)	Firm entry rate	0.1191 (0.0048)	0.1191
Acquisition (M&A)					
M&A search productivity	ψ	0.0795 (0.0095)	Agg. M&A-to-value	0.0116 (0.0013)	0.0116
Firm heterogeneity					
Productivity gap	a_h/a_ℓ	1.1283 (0.0509)	St. dev. of value-to-sales	0.5806 (0.0303)	0.5806
High-prod. entrep. share	ω	0.0011 (0.0008)	Employment Pareto tail	1.2786 (0.0027)	1.2793

The table reports the estimates of parameters in the balanced-growth equilibrium before the transition. The estimation is the same as the baseline (see Table 2), except the high-productivity entrepreneur share ω is now instead identified by the Pareto tail of the employment distribution. See the text for details

for employment bins in each year; I use $\underline{L} = 5000$ and $\bar{L} = 10000$. The empirical moment used in the estimation is then the average of annual estimates $\hat{\theta}_t$ from 1978–1985.

Table F.4 reports the results of this alternative estimation.[24] Most of the parameters take very similar values to the baseline estimation. The exception is ω, which takes a much lower value. The reason is that there is substantial concentration in the data. Indeed, as the Pareto tail approaches the lower bound $\theta = 1$ (Zipf's law), which represents maximal concentration, the estimated value of ω approaches zero.[25]

This lower value of ω does not have a substantial impact on the main results. The estimated decline in innovation productivity is virtually identical, around 47% in total. The

---

[24]In the SMM algorithm, I actually use the inverse of the Pareto coefficient $1/\theta$ as the moment, because $\theta = g_L/(\bar{\Phi}_h - \bar{\delta}_h)$ is discontinuous (i.e., changes from negative infinity to positive infinity) in any part of the parameter space over which we cross from $\bar{\Phi}_h > \bar{\delta}_h$ to $\bar{\Phi}_h < \bar{\delta}_h$.

[25]As discussed in Section 7, this is no longer true if we allow for type switching (i.e., high-productivity firms becoming low-productivity). In this case, we could still have a thick-tailed distribution with a higher value of ω and a higher productivity gap a_h/a_ℓ (so most growth occurs at the beginning of a firm's life).

decline in per-capita output growth is also similar. The rise in aggregate value to-sales is larger than in the baseline—about 96% overall, compared to 83% in the baseline. This is because there is even more reallocation when we start from a low value of ω: the total high-productivity share of sales goes from about 10% to about 30%, whereas it goes from about 15% to about 30% in the baseline. Overall, the predictions are the same.

G Model extensions: Derivations and results

G.1 Systematic risk and risk aversion

Consider the assumptions (72)–(74). Letting $Z_t = e^{z_t}$, we have

$$\frac{dZ_t}{Z_{t-}} = \left(\kappa_z(\bar{z} - z_t) + \frac{1}{2}\sigma_z^2 \right) dt + \sigma_z dB_t + (e^{-\zeta_z} - 1)dJ_t. \tag{G.1}$$

For all values of Z_t, the static production allocations across firms will be the same as in the case without systematic risk, and aggregate output will equal

$$Y_t = N_t^{1/(\eta-1)} Q_t \bar{a}_t Z_t \Omega_t L_{\mathrm{pt}}. \tag{G.2}$$

Hence, if the distribution is stationary, then expected per-capita output growth equals

$$g_{yt} \equiv \mathbb{E}_t \left[\frac{dy_t}{y_{t-}} \right] \frac{1}{dt} = \bar{g}_y - \kappa_z z_t, \tag{G.3}$$

where

$$\bar{g}_y \equiv \frac{1}{\eta-1} g_N + g_Q + \kappa_z \bar{z} + \frac{1}{2}\sigma_z^2 + p_Z(e^{-\zeta_z} - 1). \tag{G.4}$$

Expected growth is affine in z, so we should expect the same of asset prices.

Asset prices can be fully characterized by trading in a riskfree bond and the (per-capita) consumption claim. Let V_{yt} denote the value of this claim. The return on the consumption claim is, by definition, its yield plus its capital gains:

$$dR_t = \frac{y_t}{V_{yt}}dt + \frac{dV_{yt}}{V_{yt-}} = (r_{ft} + \nu_t)dt + \sigma_{Vt}dB_t + (e^{-\zeta_{Vt}} - 1)dJ_t, \tag{G.5}$$

for some risk premium ν and risk coefficient σ_V and ζ_V, for which we will solve. To simplify the proof, first conjecture that risk premia are constant ($\nu_t = \nu$) and the stock's risk coefficients are constants ($\sigma_{Vt} = \sigma_V$ and $\zeta_{Vt} = \zeta_V$), which follow from i.i.d. risks.

Letting α_t denote the share of the household's wealth W_t invested in the consumption claim, the household's wealth evolves according to[26]

$$\frac{dW_t}{W_{t-}} = \left(r_{ft} + \alpha_t \nu - \frac{c_t}{W_t}\right) dt + \alpha_t \sigma_V dB_t + \alpha_t(e^{-\zeta_V} - 1). \tag{G.6}$$

The household's HJB equation (suppressing time subscripts) is then

$$0 = \sup_{\{c,l,\alpha\}} \left\{ U_W(W(r_f + \alpha\nu) - c) + \frac{1}{2}U_{WW}W^2\alpha^2\sigma_V^2 - U_z\kappa_z z + \frac{1}{2}U_{zz}\sigma_z^2 + U_{Wz}W\alpha\sigma_V\sigma_z \right.$$
$$\left. + p_z[U(W(1 + \alpha(e^{-\zeta_V} - 1)), z - \zeta_z) - U(W, z)] + u(c, l, U) \right\}. \tag{G.7}$$

The first-order condition for consumption is

$$U_W = u_c(c, l, U) = \rho(1 - \gamma)\frac{U}{c}, \tag{G.8}$$

and that for the portfolio share is

$$0 = U_W W\nu + U_{WW}W^2\alpha\sigma_V^2 + U_{Wz}W\sigma_V\sigma_z + p_z U_W(W(1 + \alpha(e^{-\zeta_V} - 1)), z - \zeta_z)W(e^{-\zeta_V} - 1). \tag{G.9}$$

Conjecture that the solution to this system is

$$U(W, z) = \frac{W^{1-\gamma}}{1 - \gamma}\exp\{(1 - \gamma)(a_z + b_z z)\}, \tag{G.10}$$

for constants a_z and b_z. Under this conjecture, we have

$$c = \rho W \tag{G.11}$$

---

[26]Technically, in the notation of the main model, this is the evolution of total wealth per person $\overline{W}_t/L_t$, not market wealth W_t. I use W_t to simplify notation.

and

$$0 = \nu - \gamma\alpha\sigma_V^2 + (1-\gamma)b_z\sigma_V\sigma_z + p_z[(1+\alpha(e^{-\zeta_V}-1))^{-\gamma}(e^{-\zeta_V}-1)e^{-(1-\gamma)b_z\zeta_z}]. \qquad \text{(G.12)}$$

Households consume a constant fraction of their wealth, the standard unit-EIS result. Their portfolios are governed by the classic Merton (1969) risk-return tradeoff and a hedging demand against shocks to aggregate productivity z. Market clearing requires that the bond be in zero net supply, so we can set $\alpha = 1$ and $\sigma_V = \sigma_z$ and $\zeta_V = \zeta_z$ and solve for the risk premium

$$\nu = (\gamma + (\gamma-1)b_z)\sigma_z^2 + p_z[e^{(\gamma+(\gamma-1)b_z)\zeta_z}(1-e^{-\zeta_z})], \qquad \text{(G.13)}$$

which is indeed a constant, as conjectured. Standard asset-pricing results (see, e.g., Paron (2022)) and market clearing $c = y$ imply that the riskfree interest rate equals

$$r_{ft} = \rho + g_{yt} - \gamma\sigma_z^2 - p_z[e^{\gamma\zeta_z}(1-e^{-\zeta_z})], \qquad \text{(G.14)}$$

where g_{yt} is defined above. Interest rates are increasing in impatience ρ and growth g_{yt} and decreasing in precautionary savings demand from risk and risk aversion. Putting these together, the total expected return on the stock is

$$r_{ft} + \nu = \rho + g_{yt} + (\gamma-1)b_z\sigma_z^2 + p_z[e^{\gamma\zeta_z}(e^{(\gamma-1)b_z\zeta_z}-1)(1-e^{-\zeta_z})]. \qquad \text{(G.15)}$$

If $b_z \in (-1,0)$ (which I show below), then the risk premium is strictly increasing in risk aversion but the total expected return is decreasing in risk aversion, due to the dominance of the interest-rate decline.

Let us now verify the conjecture for the value function and, in particular, solve for b_z. Intuitively, we should expect $b_z < 0$, since high aggregate productivity z means low future consumption growth due to mean reversion. Substituting $U(W,z)$, $c = \rho W$, and market

A.92

clearing into the HJB equation, then dividing by $(1-\gamma)U$, implies

$$0 = r_f + \nu - \rho - \frac{1}{2}\gamma\sigma_z^2 - b_z\kappa_z z + \frac{1}{2}(1-\gamma)b_z^2\sigma_z^2 + (1-\gamma)b_z\sigma_z^2$$
$$+ \frac{p_z}{1-\gamma}[e^{(\gamma-1)(1+b_z)\zeta_z} - 1] + \rho(\log\rho - a_z - b_z z). \quad (G.16)$$

Substituting in r_f, g_y, and ν, then collecting coefficients on z, we get

$$b_z = -\frac{\kappa_z}{\kappa_z + \rho} \in (-1, 0). \quad (G.17)$$

Collecting constants, we get

$$a_z = \log\rho + \frac{1}{\rho}\left(\bar{g}_y - \frac{1}{2}(\gamma + (\gamma-1)b_z^2)\sigma_z^2\right.$$
$$\left. + p_z\left[e^{\gamma\zeta_z}(e^{(\gamma-1)b_z\zeta_z} - 1)(1 - e^{-\zeta_z}) + \frac{e^{(\gamma-1)(1+b_z)\zeta_z} - 1}{1-\gamma}\right]\right). \quad (G.18)$$

Substituting b_z back into the risk premium expression implies

$$\nu = \left(1 + \frac{\rho}{\kappa_z + \rho}(\gamma-1)\right)\sigma_z^2 + p_z\left[e^{\left(1 + \frac{\rho}{\kappa_z+\rho}(\gamma-1)\right)\zeta_z}(1 - e^{-\zeta_z})\right]. \quad (G.19)$$

This verifies the conjectured value function and completes the household's problem.

To price firms, it is useful to first derive an expression for the law of motion of the state-price density ξ of the representative household. Using the results of Paron (2022), we have

$$\frac{d\xi_t}{\xi_t} = -r_{ft}dt - (\gamma + (\gamma-1)b_z)\sigma_z dB_t + (e^{(\gamma+(\gamma-1)b_z)\zeta_z} - 1)(dJ_t - p_z dt). \quad (G.20)$$

By the absence of arbitrage, it must be that

$$r_{ft}dt = -\mathbb{E}_t\frac{d\xi_t}{\xi_t}, \quad (G.21)$$

and that

$$\nu dt = -\mathbb{E}_t\frac{d[\xi, y]_t}{\xi_t y_t}, \quad (G.22)$$

both of which are consistent with the results above.

Finally, let us assess the consequences of risk and risk aversion for firms' valuations and decisions. Recall that the value of the firm equals

$$V_{it} = \mathbb{E}_t \left[\int_0^\infty \frac{\xi_{t+\tau}}{\xi_t} D_{i,t+\tau} d\tau \right], \tag{G.23}$$

which we can scale by y_t to get

$$\hat{V}_{it} = \mathbb{E}_t \left[\int_0^\infty \frac{\xi_{t+\tau}}{\xi_t} \frac{y_{t+\tau}}{y_t} \hat{D}_{i,t+\tau} d\tau \right], \tag{G.24}$$

where the scaled dividend equals

$$\hat{D}_{it} = \sum_{j=1}^n \pi_{ijt} - \hat{w}_{\mathrm{xt}} x_{it} n_{it}, \tag{G.25}$$

and is thus independent of y and Z. After some algebra, this expression can be rewritten

$$\int_{-\infty}^0 (\xi_{t+\tau} y_{t+\tau}) \hat{D}_{i,t+\tau} d\tau + (\xi_t y_t) \hat{V}_{it} = \mathbb{E}_t \left[\int_{-\infty}^{\bar{\tau}} (\xi_{t+\tau} y_{t+\tau}) \hat{D}_{i,t+\tau} d\tau \right] + \mathbb{E}_t \left[(\xi_{t+\bar{\tau}} y_{t+\bar{\tau}}) \hat{V}_{i,t+\bar{\tau}} \right], \tag{G.26}$$

and so the expression on the left-hand side is a martingale. It is hence a local martingale, and has zero expected drift:

$$0 = \xi_t y_t \hat{D}_{it} dt + \mathbb{E}_t [d(\xi_t y_t \hat{V}_{it})]. \tag{G.27}$$

Because $\hat{V}$ is independent of ξ and y, this expression can be rewritten

$$0 = \hat{D}_{it} dt + \mathbb{E}_t \left[\frac{d(\xi_t y_t)}{\xi_t y_t} \right] \hat{V}_{it} + \mathbb{E}_t [d\hat{V}_{it}]. \tag{G.28}$$

A.94

Using the household's state-price density along with no-arbitrage conditions, we know

$$\mathbb{E}_t \left[\frac{d(\xi_t y_t)}{\xi_t y_t} \right] = \mathbb{E}_t \frac{d\xi_t}{\xi_t} + \mathbb{E}_t \frac{d[\xi, y]_t}{\xi_t y_t} + \mathbb{E}_t \frac{dy_t}{y_t} \tag{G.29}$$

$$= -(r_{ft} + \nu - g_{yt})dt \tag{G.30}$$

$$\equiv -\rho^* dt, \tag{G.31}$$

which is a constant independent of the state z_t (and of the expected growth rate g_{yt}). Substituting this back in, we have the scaled, risk-adjusted HJB equation

$$\rho^* \hat{V}_{it} = \hat{D}_{it} + \mathbb{E}_t[d\hat{V}_{it}], \tag{G.32}$$

where the firm's expected growth rate is of the exact same form as it was with log utility and no aggregate risk. This proves that $\hat{V}$ is independent of aggregate state variables.

The only difference between this model and the riskless model is that, where there was once discounting by the constant ρ, there is is now discounting by the risk-adjusted constant

$$\rho^* = \rho - (\gamma - 1)\frac{\kappa_z}{\kappa_z + \rho}\sigma_z^2 - p_z[e^{\gamma \zeta_z}(1 - e^{-(\gamma-1)\frac{\kappa_z}{\kappa_z+\rho}\zeta_z})(1 - e^{-\zeta_z})]. \tag{G.33}$$

If $\gamma = 1$, as before, then $\rho^* = \rho$, so the model truly is identical with and without systematic risk. If $\gamma > 1$, then the risk-adjusted discount factor is lower than the parameter ρ, meaning that, in order to match the same set of moments in the model with risk, we will need a higher value of ρ to get the same ρ^*. Most importantly, this means that, from the firm's perspective, whether discount rates are driven by impatience or by risk is irrelevant to R&D decisions; all that matters is the total effect of these channels.

G.2   Physical capital

Suppose instead the production function of firms takes as an input both labor and capital:

$$Y_{ijt} = a_i L_{ijt}^{1-\beta} K_{ijt}^{\beta}. \tag{G.34}$$

Like labor, physical capital K can be used in the production of any good. It follows that a firm owning capital stock K_i will use what it needs of its capital stock and, if it needs more (less), rent from (to) another firm. Consequently, it is equivalent to assume that the total capital stock is owned by a single, competitive "capital sector" which produces capital and rents it out to goods-producing firms at a rental rate r_{Kt}.

Defining the capital-labor ratio as $k = K/L$, an incumbent firm's profits equal

$$\Pi_{ijt} = \max_{p_{jt}, k_{ijt}} \left(p_{jt} - \frac{w_{\mathrm{p}t} + r_{Kt} k_{ijt}}{q_{ijt} a_i k_{ijt}^{\beta}} \right) p_{jt}^{-\eta} \left(Y_t - \sum_{-i \neq i} q_{-ijt} Y_{-ijt} \right) \tag{G.35}$$

The first-order condition for the capital-labor ratio is

$$\beta w_{\mathrm{p}t} k_{ijt}^{-1-\beta} = (1-\beta) r_{Kt} k_{ijt}^{-\beta}, \tag{G.36}$$

and therefore equals

$$k_{ijt} = \frac{\beta}{1-\beta} \frac{w_{\mathrm{p}t}}{r_{Kt}}. \tag{G.37}$$

This also means that

$$\frac{w_{\mathrm{p}t} + r_{Kt} k_{ijt}}{k_{ijt}^{\beta}} = \frac{r_{Kt}}{\beta} k_{ijt}^{1-\beta} = \left(\frac{w_{\mathrm{p}t}}{1-\beta} \right)^{1-\beta} \left(\frac{r_{Kt}}{\beta} \right)^{\alpha}, \tag{G.38}$$

so the quality-adjusted marginal cost of a firm is

$$mc_{ijt} = \frac{1}{q_{ijt} a_i} \left(\frac{w_{\mathrm{p}t}}{1-\beta} \right)^{1-\beta} \left(\frac{r_{Kt}}{\beta} \right)^{\beta}. \tag{G.39}$$

Hence, the optimal markup for the leader i_1 is the same as in the case without capital:

$$\mu_{i_1 jt} = \min \left\{ \frac{a_{i_1} q_{i_1 jt}}{a_{i_2} q_{i_2 jt}}, \frac{\eta}{\eta - 1} \right\}. \tag{G.40}$$

All of this implies that the labor and capital shares of revenue for product j are

$$\frac{w_{\mathrm{p}t} L_{ijt}}{p_{ijt} Y_{ijt}} = \frac{1-\beta}{\mu_{ijt}} \quad \text{and} \quad \frac{r_{Kt} K_{ijt}}{p_{ijt} Y_{ijt}} = \frac{\beta}{\mu_{ijt}}, \tag{G.41}$$

A.96

respectively. Consequently, the aggregate expenditure shares of output will equal

$$\Lambda_{Lt} = \frac{w_{\text{pt}} L_{\text{pt}}}{Y_t} = (1 - \beta)\Lambda_t \quad \text{and} \quad \Lambda_{Kt} = \frac{r_{Kt} K_t}{Y_t} = \beta \Lambda_t \qquad \text{(G.42)}$$

with Λ_t equal to an inverse of the aggregate markup, as in the labor-only benchmark model.

Capital is accumulated according to

$$\dot{K}_t = I_t - \delta K_t. \qquad \text{(G.43)}$$

The objective of the capital sector is to choose a capital investment policy I_t to maximize the present value V_{Kt} of its rental income less investment cost, $r_{Kt} K_t - I_t$, where the rental rate r_{Kt} is taken as given since the capital sector is perfectly competitive. The total value of the sector thus satisfies the HJB equation

$$r_{ft} V_{Kt} = r_{Kt} K_t - I_t + \frac{\partial V_{Kt}}{\partial K_t}(I_t - \delta K_t). \qquad \text{(G.44)}$$

The investment first-order condition is

$$\frac{\partial V_{Kt}}{\partial K_t} = 1, \qquad \text{(G.45)}$$

which implies

$$V_{Kt} = \frac{r_{Kt} - \delta}{r_{ft}} K_t = K_t, \qquad \text{(G.46)}$$

and hence that

$$r_{Kt} = r_{ft} + \delta. \qquad \text{(G.47)}$$

Using the fact that the aggregate capital expenditure share equals,

$$\Lambda_{Kt} = \frac{r_K K_t}{Y_t} = \beta \Lambda, \qquad \text{(G.48)}$$

it must be that the growth rate of capital equals

$$g_{Kt} = g_{Yt} + g_{\Lambda t}, \qquad \text{(G.49)}$$

A.97

where $g_{\Lambda t} = \dot{\Lambda}_t/\Lambda_t$ is zero on a balanced-growth path.[27] Thus, the equilibrium investment-capital ratio equals

$$\frac{I_t}{K_t} = \delta + g_{Yt} + g_{\Lambda t}. \tag{G.50}$$

Higher total output growth, either through technological innovation or population growth, necessitates more capital, increasing the intensity of capital investment. If Λ_t rise (markups fall), then firms increase investment to increase the capital share. Along the transition path, the investment rate will therefore fall both because the rate of economic growth falls and because the aggregate markup rises.

The only remaining change in the model is the fact that resources must now be allocated to the creation of new capital. Hence, total consumption will equal

$$C_t = Y_t - I_t = \left(1 - \frac{I_t}{Y_t}\right) Y_t. \tag{G.51}$$

Because I_t/Y_t is a constant in a stationary equilibrium, consumption growth and output growth are identical, preserving this characteristic of the baseline model.

This decline in investment intensity amid booming market valuations aligns with the literature studying the divergence of average Q (the value of the stock market relative to capital) and marginal q of investment (Crouzet & Eberly, 2023). In this model, average Q rises, for the reasons laid out in the paper, while marginal q always equals $\partial V_K/\partial K = 1$. The reason is that, in this model, most of the value of the market comes from intangible capital—blueprints, generated by R&D investment—and it is the present value of profits from intangible capital that drives average Q up. Physical capital merely serves to scale up production once new ideas are in place, and therefore stagnates as output growth stagnates.

---

[27]If g_K is too high, then there will be too much capital and r_K will fall below $r_f + \delta$, defying the investment first-order condition.

G.3 Cournot competition

Suppose firms choose quantities Y_{ijt} instead of prices p_{ijt} when they operate within a product market. The conditions for household demand remain the same, so

$$p_{ijt} = q_{ijt} \left(\frac{Y_{jt}}{Y_t} \right)^{-1/\eta} = q_{ijt} p_{jt}. \tag{G.52}$$

The profit maximization problem of a firm becomes

$$\Pi_{ijt} = \max_{Y_{ijt}} \left(p_{ijt} - \frac{w_{\mathrm{p}t}}{a_i} \right) Y_{ijt} \tag{G.53}$$

$$= \max_{Y_{ijt}} \left(\left(\frac{\sum_{i'} q_{i'jt} Y_{i'jt}}{Y_t} \right)^{-1/\eta} - \frac{w_{\mathrm{p}t}}{a_i q_{ijt}} \right) q_{ijt} Y_{ijt}. \tag{G.54}$$

The first-order condition for an interior solution is

$$p_{jt} - \frac{w_{\mathrm{p}t}}{a_i q_{ijt}} = \frac{1}{\eta} p_{jt}^{\eta+1} \frac{q_{ijt} Y_{ijt}}{Y_t}. \tag{G.55}$$

Summing over the m_{jt} producing firms i then implies

$$m_{jt} \left(p_{jt} - \frac{w_{\mathrm{p}t}}{\overline{aq}_{jt}} \right) = \frac{1}{\eta} p_{jt}^{\eta+1} \frac{Y_{jt}}{Y_t}, \tag{G.56}$$

where

$$\overline{aq}_{jt} \equiv \left(\frac{1}{m_{jt}} \sum_{i'=1}^{m_{jt}} (a_i q_{ijt})^{-1} \right)^{-1} \tag{G.57}$$

is a harmonic average of quality-adjusted productivity in the product market. Substituting in the household demand curve then implies that the product-level price equals

$$p_{jt} = \underbrace{\frac{\eta m_{jt}}{\eta m_{jt} - 1}}_{\equiv \mu_{jt}} \frac{w_{\mathrm{p}t}}{\overline{aq}_{jt}}, \tag{G.58}$$

where we have defined μ_{jt} as the markup over product-level marginal cost. Because all producers must set $p_{ijt} = q_{ijt}p_{jt}$, firm i's markup over marginal cost equals

$$\mu_{ijt} \equiv \frac{p_{ijt}}{w_{\mathrm{p}t}/a_i} = \mu_{jt}\frac{a_iq_{ijt}}{\overline{aq}_{jt}} = \frac{\eta m_{jt}}{\eta m_{jt}-1}\frac{a_iq_{ijt}}{\overline{aq}_{jt}}. \tag{G.59}$$

Among the m_{jt} producing firms, those with higher quality-adjusted productivity relative to the average will earn higher markups.

The question that remains is which firms will choose to produce and which will decide not to operate. That is, what determines m_{jt}? A firm will choose to produce ($Y_{ijt} > 0$) if and only if it will earn positive profits, or, equivalently, a markup greater than one:

$$\frac{\eta m_{jt}}{\eta m_{jt}-1}\frac{a_iq_{ijt}}{\overline{aq}_{jt}} \geq 1. \tag{G.60}$$

Under our parameterization $\lambda > a_h/a_\ell$, the markups of producers are exactly ranked by their quality rankings. Consequently, the equilibrium number of producers m is the number such that (i) the mth highest-quality firm earns a positive markup and (ii) the $(m+1)$th highest-quality firm would have a negative markup if it decided to produce.

This intuition can be used to find all possible equilibrium market structures. First, note that, given any set of potential producers, a unique equilibrium exists with a finite number of producing firms. Formally, there exists a finite $\bar{m} \in \mathbb{R}$ such that, for any combination of high- and low-productivity blueprint holders, no product market will ever have $m_{jt} > \bar{m}$ producing firms. This maximum occurs when the lowest-quality firm is high-productivity and all higher-quality blueprint holders are low-productivity (since, given m, this maximizes $a_iq_{ijt}/\overline{aq}_{jt}$ for the worst firm). In this case, the necessary condition for participation is

$$\frac{\eta\bar{m}}{\eta\bar{m}-1}\frac{1}{\bar{m}}\left(1 + \frac{a_\ell}{a_h}\sum_{k=1}^{\bar{m}-1}\lambda^{-k}\right) \geq 1, \tag{G.61}$$

or, evaluating this sum,

$$\frac{a_\ell}{a_h}\frac{\lambda^{-1} - \lambda^{-\bar{m}}}{1 - \lambda^{-1}} \geq \bar{m} - \frac{\eta+1}{\eta}. \tag{G.62}$$

The maximum number of possible producers is then the value of $\bar{m}$ that makes this an

equality.[28] Second, note that, under the parameter estimates in the main model, we will have $\bar{m} > 2$. To see this, substitute $m = 2$ into (G.60), to see that the condition for the follower $(a_i = a_{i_2})$ to produce is

$$\frac{\eta}{\eta - 1} > \lambda \frac{a_{i_1}}{a_{i_2}}, \tag{G.63}$$

which always holds in the estimated Bertrand model.

From the fact of finite $m < \bar{m}$, we could, in principle, characterize all possible market structures and their evolution in response to successful R&D and M&A. Still, one can see how the dimensionality in this Cournot world is potentially much larger than in the Bertrand world. For every new market structure we add into the model, we must add states to the HJBs and KFEs. This is, in principle, feasible, but does not add much to the model. Moreover, it adds an additional step to estimating the model: for each set of parameters, one must re-solve for the possible set of markets, which will change as $\{\eta, \lambda, a_h/a_\ell\}$ change.

G.4 Life-cycle productivity dynamics

Suppose that, with Poisson intensity $p_{h\ell}$, high-productivity firms become low-productivity firms indefinitely. In this case, blueprint characteristics are still summarized by $\{i_1, i_2, \hat{q}\}$, and we make the following alterations to the equilibrium conditions in Appendix C. Of course, blueprints with high-type leaders $(i_1 = h)$ transition to low-type leaders $(i_1 = \ell)$ with intensity $p_{h\ell}$. Likewise, for any blueprint with a high-type follower $(i_2 = h)$, the follower may become low-productivity $(i_2 = \ell)$ with intensity $p_{h\ell}$. These possibilities mean that the HJB equations (C.85) and (C.86) become

$$(\rho + \delta_t + (\eta - 1)g_{Qt})v_{i_1 i_2 t}^{\mathrm{P}} = \pi_{i_1 i_2 t} + \mathbb{I}_{i_1 = \ell}\Psi(x_{\mathrm{m}ht})f_{ht}(1 - \varrho)(v_{h i_2 t}^{\mathrm{P}} - v_{\ell i_2 t}^{\mathrm{P}}) + \frac{\partial v_{i_1 i_2 t}^{\mathrm{P}}}{\partial t}$$
$$+ p_{h\ell}(\mathbb{I}_{i_2 = h}(v_{i_1 h t}^{\mathrm{P}} - v_{i_1 \ell t}^{\mathrm{P}}) - \mathbb{I}_{i_1 = h}(v_{h i_2 t}^{\mathrm{P}} - v_{\ell i_2 t}^{\mathrm{P}})), \tag{G.64}$$

---

[28]Intuitively, as $\bar{m}$ increases, the upper bound on the left-hand side approaches a constant and the lower bound on the right-hand side rises to infinity, so $\bar{m}$ must be finite.

$$(\rho + \delta_t)v_{i_1 t}^{\mathrm{G}} = (1 - \varepsilon)(\Phi_{\mathrm{n}}(x_{\mathrm{n}it})\bar{v}_{\mathrm{n}it} + \Phi_{\mathrm{q}}(x_{\mathrm{q}it})\bar{v}_{\mathrm{q}it} + \Psi(x_{\mathrm{m}it})\bar{v}_{\mathrm{m}it})$$

$$+ \mathbb{I}_{i=\ell}\Psi(x_{\mathrm{m}ht})f_{ht}(1 - \varrho)(v_{ht}^{\mathrm{G}} - v_{\ell t}^{\mathrm{G}}) + \frac{\partial v_{i_1 t}^{\mathrm{G}}}{\partial t} - p_{h\ell}\mathbb{I}_{i_1=h}(v_{ht}^{\mathrm{G}} - v_{\ell t}^{\mathrm{G}}). \quad \text{(G.65)}$$

The second and only other alteration we must make is to the KFEs (C.144), (C.146), and (C.148), which become

$$\dot{f}_{i\varnothing\hat{q}t} + g_{Nt}f_{i\varnothing\hat{q}t} = \frac{\partial f_{i\varnothing\hat{q}t}}{\partial\hat{q}}g_{Qt} + \Phi_{\mathrm{n}it}(f_{it} + \nu_{it})\Gamma_{\hat{q}}$$

$$+ (\mathbb{I}_{i=h} - \mathbb{I}_{i=\ell})\Psi_{ht}f_{ht}f_{\ell\varnothing\hat{q}t} - (\delta_t + p_{h\ell}\mathbb{I}_{i=h})f_{i\varnothing\hat{q}t} + p_{h\ell}\mathbb{I}_{i=l}f_{h\varnothing\hat{q}t}, \quad \text{(G.66)}$$

$$\dot{f}_{ii\hat{q}t} + g_{Nt}f_{ii\hat{q}t} = \frac{\partial f_{ii\hat{q}t}}{\partial\hat{q}}g_{Qt} + \Phi_{\mathrm{q}it}(f_{it} + \nu_{it})(f_{ii,\hat{q}-\log\lambda,t} + f_{i,-i,\hat{q}-\log\lambda,t} + f_{i\varnothing,\hat{q}-\log\lambda,t})$$

$$+ \mathbb{I}_{i=h}\Psi_{ht}f_{ht}f_{\ell h\hat{q}t} - (\delta_t + 2p_{h\ell}\mathbb{I}_{i=h})f_{ii\hat{q}t} - \mathbb{I}_{i=\ell}\Psi_{ht}f_{ht}f_{\ell\ell\hat{q}t} + p_{h\ell}\mathbb{I}_{i=l}(f_{\ell h\hat{q}t} + f_{h\ell\hat{q}t}), \quad \text{(G.67)}$$

$$\dot{f}_{i,-i,\hat{q}t} + g_{Nt}f_{i,-i,\hat{q}t} = \frac{\partial f_{i,-i,\hat{q}t}}{\partial\hat{q}}g_{Qt} + \Phi_{\mathrm{q}it}(f_{it} + \nu_{it})(f_{-ii,\hat{q}-\log\lambda,t} + f_{-i,-i,\hat{q}-\log\lambda,t} + f_{-i\varnothing,\hat{q}-\log\lambda,t})$$

$$+ \mathbb{I}_{i=h}\Psi_{ht}f_{ht}f_{\ell\ell\hat{q}t} - (\delta_t + p_{h\ell})f_{i,-i,\hat{q}t} - \mathbb{I}_{i=\ell}\Psi_{ht}f_{ht}f_{\ell h\hat{q}t} + p_{h\ell}f_{hh\hat{q}t}. \quad \text{(G.68)}$$

All other equilibrium conditions remain the same.

One can see by looking at these KFEs how $p_{h\ell} > 0$ will affect the estimates. Type switching will, all else equal, reduce the total blueprint share f_h of high-productivity firms. In order to match the other moments, then, the model will require a higher productivity gap a_h/a_ℓ and a higher share of high-productivity entrepreneurs ω. In other words, a larger share of entrants will be high-productivity firms, which tend to grow exceedingly quickly over the early part of their life cycles before ultimately slowing down. This would allow the model to match the empirical fact that large firms tend to have had higher R&D intensities and growth rates earlier in life (Luttmer, 2011; Acemoglu et al., 2018; Akcigit & Kerr, 2018).