R&D Product Dispersion and Growth*

Nils H. Lehr†
Boston University

March 2022

Abstract

This paper documents that average R&D products (ARDPs), defined as the value created from innovation divided by the associated cost, are highly dispersed among US listed, innovation-intensive firms and that this dispersion has increased over time. This finding is surprising ex-ante as workhorse endogenous growth models predict ARDP equalization across firms. After ruling out benign drivers, I argue that ARDP dispersion can be interpreted as misallocation. I develop a growth rate decomposition for a general class of endogenous growth models that allows me to directly link ARDP dispersion to economic growth. Combing data and model I estimate US growth could be up to 30% larger without ARDP dispersion and that its rise can account for up to 50% of the declining long-run growth rate. I investigate potential mechanisms and present reduced form and quantitative evidence in favor of adjustment cost and labor market power.

*I am especially grateful to my advisors Stephen Terry, Pascual Restrepo, David Lagakos, and Tarek Hassan for their guidance, feedback, and encouragement on this project. I benefited from the advice of many including Adam Guren, Diego Restuccia, Maarten De Ridder, Sara Moreira, and the participants in the BU Macro Lunch Seminar.

†Email: nilslehr@bu.edu Mailing Address: Boston University, Dept. of Economics, 270 Bay State Rd., Room 515, Boston, MA 02215.

1 Introduction

This paper documents that the average R&D product (ARDP), defined as the value created by R&D divided by the associated cost, is highly dispersed across and within US listed, innovation-intensive firms and that this dispersion has increased over time. This finding is ex-ante surprising as workhorse models of endogenous growth predict ARDP equalization across firms.[1]

In a series of tests I confirm that this dispersion is not driven by measurement error, misspecification, or other benign factors. Instead, I argue that this finding can be rationalized by interpreting the dispersion as frictions or indicators of misallocation. Following this interpretation I show that dispersion in ARDPs is associated with lower economy growth in a wide class of models according to a simple formula decomposing growth into a frictionless benchmark and an adjustment factor linked to frictions. Combining data and model, I find that the dispersion in ARDPs is associated with up to 30% lower growth, or 0.47 percentage points p.a. against a baseline of 1.5% annual growth. Furthermore, I find that rising ARDP dispersion can explain up to 50% of the long-run slowdown of economic growth.

Naturally, the question arises as to the source of misallocation. I focus on two potential sources, adjustment cost and labor market power, and provide suggestive reduced form evidence. For adjustment cost, I show that the growth rate of R&D expenditure is auto-correlated and that R&D expenditure is auto-correlated conditional on current R&D productivity. Both findings are a sign of adjustment cost in standard models. For labor market power, I show that ARDPs are increasing in firm-level labor productivity, which has been linked to labor market power in Seegmiller (2021), and inventor employment concentration.

To estimate the empirical contribution of both mechanisms to ARDP dispersion, I propose a quantitative Schumpeterian growth model and estimate it via SMM. The estimated parameters suggest a significant role for both frictions and qualitatively reproduce a set of facts that a frictionless model cannot explain. The estimated model suggests that both frictions jointly can explain 14% and 26% of the overall and within-firm dispersion in measured ARDPs respectively.

This paper contributes to three strands of the literature. Firstly, the paper is closely related to the growing literature on the slowdown in productivity growth (Gordon, 2016;

[1] This prediction is shared among expanding variety models as surveyed in Gancia and Zilibotti (2005) and Schumpeterian growth models as surveyed in Aghion et al. (2014). It is also applicable to more recent contributions including Acemoglu and Cao (2015); Akcigit and Kerr (2018); de Ridder (2021) and Aghion et al. (2022).

Syverson, 2017).[2] The existing literature has focused on two explanations. Firstly, Bloom et al. (2020) argue that R&D productivity has been declining exogenously in the long run, which can explain a decline in productivity growth if innovation expenditure is not rising fast enough. Secondly, de Ridder (2021); Aghion et al. (2022); Liu et al. (2022) argue for an increasing wedge between public and private incentives that reduces the productivity improvement per unit of R&D and thus economic growth. I complement these perspectives by highlighting the importance rising of frictions in the private market as a source of lower productivity growth and lower R&D productivity.

Secondly, the paper adds to the literature on misallocation in endogenous growth models, which has mostly focused on the importance of incentive misalignment across planner and private firms.[3] Starting with Romer (1990) the literature has recognized that the private market under-provides innovation due to dynamic externalities and has investigated a range of mechanisms and policy instruments. Important exceptions to this focus include Terry (2017) and Terry et al. (2021), who consider misalignment of incentives between managers and owners as a source of misallocation. I complement this literature by highlighting the importance of private frictions, providing a simple growth rate decomposition linking them to economic growth and estimating their contribution.

Finally, this paper contributes to the broader literature on misallocation by estimating its empirical importance for R&D expenditure of US listed firms and highlighting its potency to reduce economic growth.[4] The latter is especially important given that welfare tends to be more sensitive to changes in the growth rate compared to changes in static efficiency (Acemoglu and Akcigit, 2012; Peters, 2020). My paper has many commonalities with the broader finding in this literature. For example, similar to the findings Restuccia and Rogerson (2008) for static production, I find that ARDP is significantly correlated with R&D productivity. Furthermore, I find a limited role for financial frictions as in Midrigan and Xu (2014), while adjustment cost and labor market power seem important (Asker et al., 2014; Berger et al., 2022).

[2] See also Andrews et al. (2016); Philippon and Gutiérrez (2017); Aksoy et al. (2019); Brynjolfsson et al. (2019); Engbom (2020); Akcigit and Ates (2021); Peters and Walsh (2021); de Ridder (2021); Aghion et al. (2022); Liu et al. (2022); Lehr (2022).

[3] See Gancia and Zilibotti (2005) and Aghion et al. (2014) for literature reviews. Recent contributions along those lines include Akcigit and Kerr (2018); Peters (2020); Akcigit and Ates (2021); Akcigit et al. (2021); de Ridder (2021); Mezzanotti (2021); Aghion et al. (2022); Liu et al. (2022).

[4] Important contributions to the broader literature on misallocation include Whited and Wu (2006); Restuccia and Rogerson (2008); Hsieh and Klenow (2009); Banerjee and Moll (2010); Asker et al. (2014); Midrigan and Xu (2014); David et al. (2016); David and Venkateswaran (2019); De Loecker et al. (2020); Ottonello and Winberry (2020); Berger et al. (2022).

The remainder of this paper is structured as follows: Section 2 introduces the data and sample selection. Section 3 motivates the focus on ARDP, discusses its empirical measurement and presents motivational facts. Section 4 shows how one can map dispersion in ARDP to economic growth in a general class of endogenous growth models, discusses measurement of key parameters, and estimates the impact of ARDP dispersion on economic growth and the contribution of rising ARDP dispersion to the productivity slowdown. Section 5 discusses adjustment cost and labor market power as mechanisms for ARDP dispersion, presents associated evidence, and quantifies their contribution to the empirical dispersion in ARDP. Finally, 6 concludes and highlights avenues for future research.

2 Data

My main dataset combines balance sheet and income statement information for publicly listed firms with information on their patenting activities. I will detail the sample selection and aggregation process below.

Accounting data. I obtain accounting data on US listed firms from WRDS Compustat. I restrict my sample to firms outside the utilities and financial sector as is common in the literature. I impute missing R&D expenditure (`xrd`) with 0s. I focus on firms that are active beyond 1984.

Patent data. I create firm-level measures of innovation output using patent data. My main measure of innovation value uses the patent valuation measures in Kogan et al. (2017) aggregated to the firm-year level via the date of patent application. Kogan et al. (2017) measure the value created by patents in an event study approach by studying the trajectory of firm valuations around the patent announcement date. I use the updated version of their dataset, which includes data on patent valuations up to 2019.

Using the crosswalk between patents and firms established in Kogan et al. (2017), I complement this data with information patent impact via forward citations and via the text-based measure constructed in Kelly et al. (2021). For the citation based measures I count the citations that a patent receives within the first five years since the grant and divide this number by the average for all patents within a application year. This procedure allows me to have an impact measure that is comparable across time and can be constructed even for patents towards the end of my sample. I aggregate the measures up to the firm-year-level by summation.

Sample selection. I merge both datasets via `gvkey` and year and restrict my sample to innovation-intensive firms with at least

1. 2.5m in annual R&D expenditure on average,

2. 2.5 patents per year on average, and

3. 5 years with positive patenting and R&D expenditure.

Furthermore, I drop firm observations with less than 100 employees, 100k in sales, or 100k in capital. The final sample covers 80% of the R&D expenditure in WRDS Compustat. Finally, I restrict the sample to the 1975-2014 period.

Additional data. Furthermore, I calculate measures regarding inventor employment of firms by mapping inventors to firms via patents and identifying inventors from name×location combinations in the Berkes (2016) dataset. For further details, see text and Appendix A.

3 Facts on Average R&D Product Dispersion

This section motivates the use of ARDP, discusses measurement, and presents stylized facts on ARDP dispersion.

3.1 A Firm's Innovation Problem

It is useful to motivate the focus on ARDPs within the context of the firm's problem in a standard endogenous growth model (Romer, 1990; Aghion and Howitt, 1992; Klette and Kortum, 2004; Acemoglu and Cao, 2015). Consider a firm in continuous time with access to R&D projects that arrive at Poisson rate z_{it} and yield expected value $\mathbb{E}_t[\mathcal{V}_{it}]$. The arrival rate itself depends on the firm's innovation efficiency φ_{it} and the hired R&D labor ℓ_{it} via a decreasing returns to scale production function with scale elasticity $\frac{1}{1+\phi}$. Researchers are hired in a frictionless spot market at a wage W_t.

The firm's maximization problem with respect to innovation is thus given by

$$\max_{\ell_{it}} \left\{ z_{it}\mathbb{E}_t[\mathcal{V}_{it}] - W_t\ell_{it} \quad \text{s.t.} \quad z_{it} = \varphi_{it}\ell_{it}^{\frac{1}{1+\phi}} \right\}. \tag{1}$$

Note that this formulation encompasses many models in the literature by leaving the value of innovation unspecified. It thus includes models where the value is exclusively tied to profits such as Romer (1990) as well as models where part of the value is due to expanding a firm's innovation abilities such as Klette and Kortum (2004). Furthermore, it allows for systematic differences in the value across firms due to e.g. differential market power or R&D depreciation rates as in de Ridder (2021). Similarly, a firm might be structured as a collection of these problem as in Klette and Kortum (2004) in which case the optimization applies problem by problem. Similarly, this problem could apply to internal and external innovation separately as in Peters (2020).

The first order conditions to the firm's optimization problem together with the functional form of the production function require the expected average R&D product, defined as the expected value from innovation divided by total cost, to be equalized across firms or projects:

$$\text{Expected ARDP}_{it} \equiv \frac{z_{it}\mathbb{E}[\mathcal{V}_{it}]}{W_t \ell_{it}} = 1 + \phi. \tag{2}$$

This result follows as firms equalize marginal product to marginal cost, both of which are proportional to the total value created and total cost respectively. Note that this result does not hinge on the exact formulation of the value of innovation and can accommodate heterogeneous factor prices by replacing W_t with W_{it}. Furthermore, the formulation allows for arbitrary efficiency differences across firms.

In practice, firms might face a range of frictions such as adjustment cost, financial frictions, and market power that prevent them from achieving the unconstrained, competitive allocation. I summarize these friction in a single measure ξ_{it} such that

$$\frac{z_{it}\mathbb{E}[\mathcal{V}_{it}]}{W_t \ell_{it}} = \frac{1+\phi}{\xi_{it}}. \tag{3}$$

A large ξ_{it} is associated with low expected ARDP as the firm conducts more R&D than it would in absence of frictions.

How important are these frictions for growth? In Section 4 I show that under some conditions, we can express the economic growth rate of an economy as

$$g = g^{\max} \times \Xi,$$

where $g^{\max}$ is the frictionsless growth rate and Ξ is a function of the frictions faced by firms. Importantly, I show that the allocative efficiency term Ξ is decreasing in dispersion in ξ_{it} such that an economy with a higher level of volatility in ξ_{it} is characterized by lower growth. Motivated by this consideration, I will investigate the distribution of expected ARDPs across firms next.

3.2 Measuring Average R&D Products

I define the empirical counterpart of the expected ARDP, which I will call realized ARDP, for firm i at time t and over some window W as

$$\texttt{Realized ARDP}_{it}^{W} \equiv \frac{\sum_{w=0}^{W-1} \texttt{Patent valuations}_{it+w}}{\sum_{w=-1}^{W-2} \texttt{R\&D Expenditure}_{it+w}}. \tag{4}$$

Note that this formulation assumes that innovation projects on average are realized around one year after the associated expenditure occurred. Furthermore, I rely on the assumption that patent valuations capture the value created from innovation.

There are at least three immediate concerns when relating this measure to its theoretical counterpart. Firstly, realizations do not necessarily equal expectations. I address this concern by (i) focusing on firms with large R&D budgets and meaningful patenting activity via my sample selection, (ii) considering longer time horizons, using 5-year windows as my baseline and providing robustness using 10-year and 20-year windows, and (iii) only considering realized ARDPs with at least 50 patent valuations in the nominator. Jointly, these restrictions aim to approximate a law-of-large-number environment in which realization equal expectations. Secondly, I winsorize patent valuations at the 95th percentile to guard against outliers driving variation. This is important as it is a well established fact that patent outcomes are fat-tailed (Kogan et al., 2017; Akcigit and Kerr, 2018). Finally, I residualize log ARDPs with respect to industry $\times$ year fixed effects. This adjustment ensures, e.g., that variation in realized ARDPs is not driving by industry differences in patenting rates of innovation and also filters out potential technology differences across industries via ϕ as drivers of variation.

3.3 Average R&D Product Dispersion

With this measure in hand, I document the first stylized fact: ARDPs are dispersed. Panel A in Table 1 shows that the standard deviation of ARDP is around 0.92 in my sample. Furthermore, this value is tightly estimated as confirmed by the small standard errors.[5] To

[5] I calculate standard errors via GMM and the beta method. Standard GMM arguments ensure that the estimator for variance of ARDPs converges towards a normal distribution. The beta method then allows me to extend this insight to transformations of the variance including the standard deviation.

highlight the importance of above mentioned adjustments, I calculate "unadjusted" ARDPs that consider all observations with positive patenting, do not winsorize patent values, and do not take out industry × year fixed effects. The estimated standard deviation is $\frac{1.308-0.916}{0.916}=$ 42.7% larger for unadjusted ARDPs, confirming that my baseline estimate is conservative.

Table 1: Average Product Dispersion in Comparison

Variable	Sample	
	Unadjusted	Baseline
A. Average R&D Product		
SD	1.308	0.916
S.E. (SD)	(0.005)	(0.006)
B. Labor Revenue Product		
SD	0.680	0.568
S.E. (SD)	(0.010)	(0.014)
$\frac{\text{SD(LRP) - SD(ARDP)}}{\text{SD(LRP)}}$	92.4%	61.4%
C. Capital Revenue Product		
SD	0.821	0.701
S.E. (SD)	(0.010)	(0.015)
$\frac{\text{SD(KRP) - SD(ARDP)}}{\text{SD(KRP)}}$	59.2%	30.7%
Observations	27,140	11,773

Note: Labor and capital revenue products are defined as revenue divided by employment and capital respectively. All variables in logs and residualized with respect to NAICS3 × year. Unadjusted restricts sample to at least 5 patents over the 5-year period. Baseline restricts the sample to at least 50 patents over the 5-year period and winsorizes top 5% of patent valuation by year. Standard errors are calculated via GMM.

Naturally, the question arises as to whether this dispersion is large. One potential source of benchmarks are measures considered to be informative about static resource allocation. For example, Hsieh and Klenow (2009) estimate that the standard deviation log TFPR for US plants is around 0.45, while their theory suggests that TFPRs should be equalized across plants in a frictionsless benchmark. Similarly, David and Venkateswaran (2019) estimate the standard deviation of the average product of capital, defined as sales divided by capital stocks, to be around 0.46 for Compustat firms, while their theory suggest equalization, or at least very low dispersion, in absence of frictions. Compared to both values, the estimated standard deviation of ARDPs is about twice as large.

I offer two additional points of reference leveraging accounting data for my sample. Using revenue (`revt`), employment (`emp`), and beginning of period capital stock (lagged `ppent`) I calculate the average revenue product of labor and capital (APL/APK) at the 5-year horizon:

$$\texttt{APK} \equiv \frac{\sum_{w=0}^{4} \texttt{Revenue}_{it}}{\sum_{w=0}^{4} \texttt{Capital stock}_{it}} \qquad \text{and} \qquad \texttt{APL} \equiv \frac{\sum_{w=0}^{4} \texttt{Revenue}_{it}}{\sum_{w=0}^{4} \texttt{Employment}_{it}}. \tag{5}$$

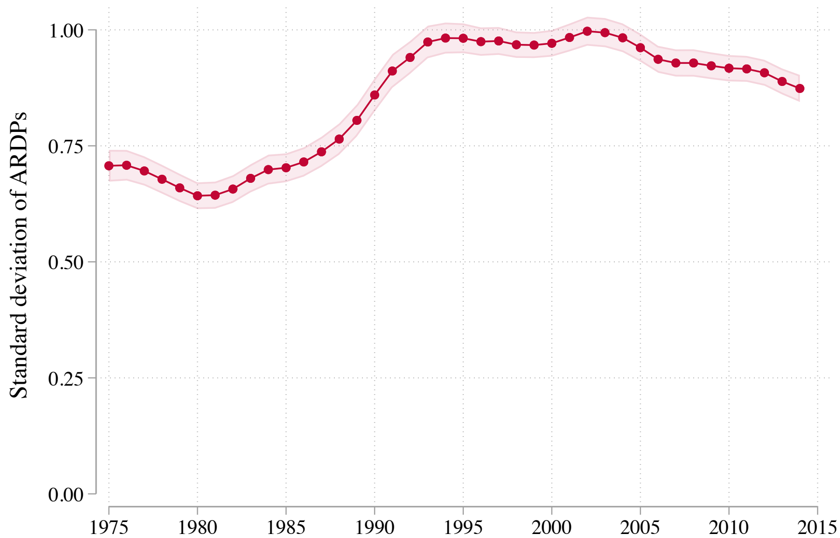
As discussed in e.g. Hsieh and Klenow (2009), both should be equalized across firms in a frictionless benchmark, similar to ARDPs, and thus serve as a natural reference point. As reported in in Panel B and C of Table 1, ARDP dispersion is 61.4% and 30.7% larger than dispersion in `MPL` and `MPK` respectively. I thus conclude that the dispersion is large.

Stylized Fact 1. *Realized ARDP dispersion is large.*

3.4 Evolution of Dispersion

Naturally, the question arises as to whether ARDP dispersion has changed over time. Indeed, as documented in Figure 1, it has increased over time. While dispersion was below 0.75 from 1975 to 1985, it has been almost entirely above 0.90 since 1995. Column (1) in Table 2 confirms that this increase has been significant. In the early sample from 1974-94, the standard deviation has been around 0.80 compared to 0.95 for the late period from 2004-14, a 19% increase. Both values are precisely estimated and the difference is highly significant.

Figure 1: Evolution of ARDP Dispersion



Notes: Figure plots the evolution of the standard deviation of ARDPs estimated using 5-year rolling windows around the central year. Standard errors are calculated via GMM.

A natural question is whether this increase is driven by firm selection or within firm. I attempt to shed some light on this question by focusing on firms with significant activity in the early and late sample. Column (2) shows that these firms have lower ARDP dispersion overall, however, the increase in dispersion has been even more pronounced going from 0.69 to 0.87, a 26% increase.

Table 2: Evolution of ARDP Dispersion

	Specification		
	Baseline	Continuing firms	Patent-weighted
A. Full sample			
SD	0.909	0.770	0.727
S.E. (SD)	(0.006)	(0.005)	(0.005)
Observations	11,804	3,057	11,804
B. Early sample (1974-94)			
SD	0.797	0.690	0.667
S.E. (SD)	(0.009)	(0.007)	(0.007)
Observations	4,494	1,479	4,494
C. Late sample (2004-14)			
SD	0.946	0.868	0.740
S.E. (SD)	(0.010)	(0.009)	(0.009)
Observations	3,894	856	3,894
$\frac{\text{SD Late} - \text{SD Early}}{\text{SD Early}}$	18.7%	25.8%	10.9%

Note: Standard errors are calculated via GMM. The F-test for equality of variances for the late vs early period is rejected with p-values below 0.001 in all cases.

Stylized Fact 2. *Realized ARDP dispersion has increased significantly from the 1974-94 to 2004-14 period.*

3.5 Benign Sources of ARDP Dispersion

A lingering concern at this point might be that the variation is not economically meaningful. On the one hand, the variation could be driven by measurement error or misspecification and thus be unconnected to real outcomes. On the other hand, they could reflect technology differences across firms or heterogeneous subsidies that make them uninformative about welfare-relevant frictions. I will address these concerns next.

Measurement error. Given the volatile nature of innovation, potential long time-horizons involved, and noise involved in measuring the value of individual patents, measurement error is a natural concern. In this context it could show up through multiple channels. For example, measurement error could arise simply due to noise in the patent valuations from Kogan et al. (2017). Alternatively, it could arise as there remains a distance between expectations and realization. Note that for both cases, we might reasonably expect the measurement error to show up as uncorrelated over time, i.e. as classical measurement error. I test prediction by estimating a simple auto-correlation model:

$$\texttt{ARDP}_{it} = \alpha_{j(i)\times t} + \beta\texttt{ARDP}_{it-5} + \varepsilon_{it}, \tag{6}$$

where $\alpha_{j(i)\times t}$ are industry $\times$ year fixed effects. If the measured ARDPs were pure measurement error, then we'd expect $\hat{\beta} = 0$.

Table 3 reports the associated estimates in column (1). ARDPs are highly persistent at the 5-year horizon with a precisely estimated auto-correlation coefficient around 0.7. Note that this finding also rules out timing mismatch, where there is a mismatch between the expenditure and created valuation due to the time-windows in consideration, as this source of measurement error would predict negative auto-correlation.[6]

[6] For example, suppose that for a particular innovation the cost are incurred 6 years instead of 1 year before the patent application. In this case my ARDP for the time-window from t-5 to t-1 is too low as I count R&D expenditure that only yields value after the period. On the other hand, my ARDP for the t to t+4 period is going to be too high as I only count the value of the innovation and not the cost. This phenomenon naturally leads to negative auto-correlation of ARDPs.

Table 3: ARDPs Are Significantly Auto-Correlated

	(1) OLS	(2) IV	(3) IV	(4) IV
		$\texttt{ARDP}_{it+5}$		
$\texttt{ARDP}_{it}$	0.711***	0.638***	0.699***	0.657***
	(0.016)	(0.031)	(0.165)	(0.031)
Instrument		Lagged ARDP	Sales-based ARDP	Employment-based ARDP
First stage F		676	10	23
Observations	6,996	4,443	6,996	6,996

Note: This table reports coefficient estimates for an AR(1) specification. All specifications include industry × year fixed effects. Standard errors clustered at the industry level.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

Using this regression specification we can go one step further to get a sense of the importance of measurement error using a simple instrumental variable approach. Suppose that ARDP is measured with some noise, then the OLS auto-correlation coefficient is going to be downwards biased due to measurement error. I can thus gage the extend of measurement error by comparing the OLS coefficient with an unbiased estimate.

Formally, suppose that the researcher only observed $\widetilde{\texttt{ARDP}}_{it} = \texttt{ARDP}_{it} + \nu_{it}$, where ν_{it} is a iid measurement error with variance σ_ν^2. Then we have

$$\beta^{OLS} \equiv \frac{\text{Cov}\left(\widetilde{\texttt{ARDP}}_{it}, \widetilde{\texttt{ARDP}}_{it+5}\right)}{\text{Var}\left(\widetilde{\texttt{ARDP}}_{it}\right)} = \underbrace{\frac{\text{Cov}\left(\texttt{ARDP}_{it}, \texttt{ARDP}_{it+5}\right)}{\text{Var}\left(\texttt{ARDP}_{it}\right)}}_{\equiv\beta} \times \underbrace{\frac{\text{Var}\left(\texttt{ARDP}_{it}\right)}{\text{Var}\left(\texttt{ARDP}_{it}\right) + \sigma_\nu^2}}_{<1 \text{ if } \sigma_\nu^2>0}. \tag{7}$$

Thus, if we have an unbiased estimate of β, we can gage the importance of measurement error using the ratio of β^{OLS} to β with values below 1 indicating the presence of measurement error.

To get an unbiased estimate I rely on an instrumental variable strategy using either lagged values of $\texttt{ARDP}_{it}$ as instrument or by constructing alternative measures using positive changes in the sales or employment instead of patent valuations. For the latter I replace patent valuations in the formula for ARDPs by year-by-year positive changes in employment `emp` or sales `sales` over the same 5-year window as the patent valuations:

$$\mathtt{ARDP}_{it}^{X} = \frac{\sum_{w=0}^{4} \max\{0, X_{it+w} - X_{it+w-1}\}}{\sum_{w=-1}^{3} \mathtt{R\&D\ Expenditure}_{it+w}} \qquad X \in \{\mathtt{Employment}, \mathtt{Sales}\}. \tag{8}$$

This measure is an instrument in so far as that it is not correlated with any measurement error in patent valuations. Lagged values are a valid instrument when adjusting for measurement error as long as it is not auto-correlated nor correlated with future ARDP other than via the current value of ARDP.

Columns (2)-(4) in Table 3 report the associated second stage results. Throughout, I find that the IV coefficients are smaller than the OLS coefficient, although not always signficantly so. This results suggests that measurement error is not a significant concern.

In Appendix B.1 I expand on these results by conducting a structural decomposition of the variation in ARDP via GMM linking it to permanent differences across firms, temporary innovations, and classical measurement error. The exercise confirms that classical measurement error is not a significant contributor to variation in ARDPs.

Misspecification. Another potential source of ARDP variation is misspecification in inputs or outputs, where the variables either do not capture all the value created or the inputs used. Starting with the latter, I investigate the possibility that R&D stock instead of flow is the appropriate input measure, where I define the concept iteratively as

$$\mathtt{R\&D\ Stock}_{it} = \mathtt{R\&D\ Expenditure}_{it} + (1-\delta) \times \mathtt{R\&D\ Stock}_{it-1}, \tag{9}$$

where $\delta = 1$ recovers the baseline input cost. I then construct an alternative ARDP as

$$\mathtt{ARDP}_{it}^{Stock} \equiv \frac{\sum_{w=0}^{4} \mathtt{Patent\ valuations}_{it+w}}{\sum_{w=-1}^{3} \mathtt{R\&D\ Stock}_{it+w}}. \tag{10}$$

I experiment with different values for δ and consistently find that dispersion in this alternative ARDP is larger for $\delta < 1$. Table 4 reports the associated results.

Table 4: Average R&D Product Dispersion With Alternative R&D Depreciation Rates

	R&D Depreciation Rate					
	5%	10%	25%	50%	75%	100%
SD of ARDP	1.112	1.071	1.003	0.957	0.935	0.923
Δ% to Baseline	20.6%	16.1%	8.7%	3.7%	1.4%	0.0%

Note: Standard deviations taken over natural logarithm of Average R&D Product measure. All values residualized by NAICS3× Year fixed effects. ARDPs are calculated at the 5-year horizon. Columns use alternative depreciation rates for R&D expenditure to get a measure of total innovation input.

As a second robustness check, I take inspiration in the growing literature on intangibles and investigate whether including overhead expenses can account for part of the variation (Peters and Taylor, 2017; Ewens et al., 2020). The argument would be that part of overhead expenses facilitate R&D and, thus, account for effective R&D expenditure. I explore this possibility by constructing an alternative measure of R&D input that adds a fraction α of S,G&A expenses (`xsga`) to R&D expenditure:

$$\texttt{R\&D Input}_{it} = \texttt{R\&D Expenditure}_{it} + \alpha \times \texttt{S,G\&A Expenditure}. \tag{11}$$

I then calculalte the dispersion of the alternative ARDPs for a range of α values. As shown in Table 5, I find that a intermediate value of α decreases ARDP dispersion moderately. For example, setting $\alpha = 10\%$ results in a reduces ARDP dispersion by a moderate 6.5%. On the other hand, setting $\alpha = 100\%$ leaves dispersion virtually unaffected.

Table 5: Average R&D Product Dispersion Adding S,G&A to R&D

	Percent of S,G&A Added to R&D					
	0%	5%	10%	25%	50%	100%
SD of ARDP	0.917	0.878	0.872	0.882	0.908	0.950
Δ% to Baseline	0.0%	-4.3%	-4.9%	-3.8%	-1.0%	3.6%

Note: Standard deviations taken over natural logarithm of Average R&D Product measure. All values residualized by NAICS3× Year fixed effects. ARDPs are calculated at the 5-year horizon. Columns add alternative percentages of S,G&A to R&D expenditure to get a measure of total innovation expenditure.

Finally, I investigate misspecification of outputs by experimenting with alternative formulations of value creation. Inspired by Bloom et al. (2020), I consider four potential measures capturing value creation: Changes in revenue, employment, market valuation, and labor productivity. For each of these variables I construct the measure of value created as non-negative changes in the outcome:

$$\texttt{Value Creation}_{it}^{Y} = \max\{0, Y_{it} - Y_{it-1}\}. \tag{12}$$

I then replace patent valuation with the respective outcome measure in my realized ARDP formula (4) and calculate dispersion for the resulting alternative ARDP.

Table 6: Average R&D Product Dispersion using Alternative Measures of Value Creation

ARDP Measure	Alternative Outcome Measure				
	—	Δ Revenue	Δ Employment	Δ Market Valuation	Δ $\frac{\text{Revenue}}{\text{Employment}}$
Valuation-based ARDP	0.916	0.913	0.904	0.959	0.909
Alternative ARDP	—	1.136	1.384	1.172	1.624
Correlation	—	0.502	0.389	0.680	0.086
Observations	11,839	11,682	10,870	6,749	11,580

Note: Standard deviations taken over natural logarithm of Average R&D Product measure. All values residualized by NAICS3× Year fixed effects. Valuation-based ARDP uses baseline ARDP measure for same sample as available for alternative ARDP measure. ARDPs are calculated at the 5-year horizon. Value created for alternative measures is calculated as the sum of positive year-to-year changes over the time-horizon.

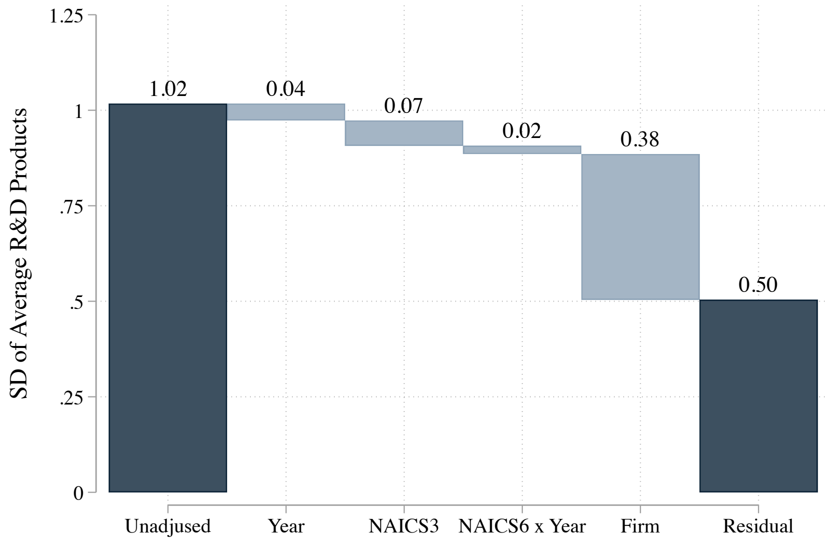
The associated results are presented in Table 6 and suggest that using the alternative outcomes measures yields significantly larger dispersion in ARDPs of at least 20%. Nonetheless, the alternative measures are significantly correlated with the baseline measure of ARDPs. For example, the correlation between the ARDP based on revenue changes has a correlation coefficient around 0.5 with the baseline ARDP. This finding confirms that the measure captures something that goes beyond patent valuations.

Technology differences in ϕ. The simplest extension to justify ARDP dispersion is to simply assume that scale elasticities are time and firm dependent such that

$$\frac{z_{it}\mathbb{E}[\mathcal{V}_{it}]}{W_t\ell_{it}} = 1 + \phi_{it}. \tag{13}$$

Importantly, this explanation implies that ARDP dispersion is optimal from a firm and planner perspective as firms operate on their first-order conditions. While this is a possibility that is difficult to fully rule out, I will argue that it provides an unsatisfactory explanation to the variation. For this I will provide three pieces of evidence. Firstly, if we think that differences in ϕ_{it} drive ARDP dispersion, we might expect that most of the variation is across (potentially narrow) industries that share production technologies and methods of operation. Figure 2 show that this is not the case. Even detailed industry × year fixed effects explain less than 13% of the overall variation compared to 37% of the variation attributable to long-run differences across firms, but within narrow industry × year categories, and about 50% of the variation being within firm and detailed industry × year fixed effects.

Figure 2: Decomposition of Variation in Realized ARDPs



Notes: This figure decomposes the standard deviation of ARDP. The first and last bar report the variation in the baseline and the fully residualized ARDP respectively. Bars two to five report the additional variation explained by adding the indicated fixed effect. See text and Appendix A for data details.

Secondly, I conduct an explicit test of this hypothesis building on the empirical literature estimating ϕ via changes in the subsidy rate for R&D. In particular, one can show that the unit-prices elasticity of R&D is given by $-\frac{1}{\phi}$ such that

$$\frac{\partial \ln \texttt{R\&D Expenditure}_{it}}{\partial \ln(1-\tau_{it})} = -\frac{1}{\phi}, \tag{14}$$

where τ_{it} is the subsidy rate. The empirical literature estimates this relationship using changes in subsidy rates as identifying variation (Bloom et al., 2002; Guceri, 2018; Guceri and Liu, 2019; Dechezleprêtre et al., 2019). I follow this approach using the state-level dataset on R&D subsidies from Lucking (2019) and estimate a simple interaction model:[7]

$$\ln\left(\sum_{w=0}^{4} \texttt{R\&D Expenditure}_{it+w}\right) = \alpha_i + \beta \ln(1-\tau_{it}) + \gamma \ln(1-\tau_{it}) \times \overline{ARPD}_i + \delta X_{it} + \varepsilon_{it}. \tag{15}$$

Here, X_{it} includes a set of size control including employment and capital stock that add precision to the estimate as well as year fixed effects. This specification is similar to Bloom

[7] In my preferred specification I merge via the state of residency of the majority of the inventors working for the firm. I also provide robustness merging instead via the state of the assignee and company headquarters. I consider the inventor location to be the most accurate measure of the unit cost faced by the firm as it directly captures the location of the inventive workforce.

et al. (2002).

While the coefficient β simply identifies the unit-cost elasticity of a firm with average ARDP, the coefficient γ investigates whether larger ARDPs are also associated with larger unit cost elasticities. If variation in ARDP was driven by differences in ϕ_{it}, then we would expect $\gamma > 0$ as firms with large ARDP have lower unit cost elasticities. Instead my estimates in Appendix Table B.2 suggest that there is either no effect at all or that $\gamma < 0$. Naturally, these estimates are subject to the general empirical concerns with estimating ϕ via changes in R&D subsidies, however, they nonetheless provide some evidence that technology differences are unlikely to be the dominant driver of ARDP differences.

Finally, I show that heterogeneity in ARDPs is linked to R&D growth. For this I regress the growth rate of 5-year R&D expenditures on current 5-year ARDP:

$$\Delta_{t,t+5} \ln \left(\texttt{R\&D Expenditure}_{it \to t+4}\right) = \alpha + \beta \texttt{ARPD}_{it} + \delta X_{it} + \varepsilon_{it}. \tag{16}$$

If heterogeneity in ARDP simply was concerned with technology differences, then we might not expect it to predict changes in R&D expenditures. However, as shown in Table 7, it is significantly positively correlated with growth in R&D expenditures. Notably, this holds even when controlling for past and current R&D expenditure itself as shown in column (2) and (3). Column (4) shows that this correlation also holds within firm.

Table 7: ARPDs and R&D Growth

	(1)	(2)	(3)	(4)
	$\Delta_{t,t+5}$`R&D Expenditure`$_{it\to t+4}$			
`Realized ARPD`$_{it\to t+4}$	0.333***	0.278***	0.278***	0.356***
	(0.035)	(0.026)	(0.029)	(0.045)
$\Delta_{t-5,t}$`R&D Expenditure`$_{it\to t+4}$		0.222***	0.215***	0.140***
		(0.024)	(0.030)	(0.031)
`R&D Expenditure`$_{it\to t+4}$			-0.032*	-0.279***
			(0.018)	(0.036)
Firm FEs				✓
R2	0.421	0.457	0.463	0.718
Within R2	0.246	0.269	0.276	0.323
Observations	10,631	9,336	9,336	9,283

Note: This table reports coefficient estimates for . See text and Appendix for variable descriptions. All variables in logs. All regressions control for NAICS3 $\times$ year fixed effects and cluster standard errors at the NAICS3 level.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

In conclusion, the evidence presented above suggest that technology differences via ϕ are unlikely to be a substantial driver of ARDP heterogeneity.

Taxes and Subsidies. Another possibility is that some of the variation is driven by differences in subsidies across firms. Formally, let τ_{it} be the subsidy rate, then we have

$$\frac{z_{it}\mathbb{E}[\mathcal{V}_{it}]}{W_t\ell_{it}} = (1+\phi)(1-\tau_{it}). \tag{17}$$

It thus follows immediately that heterogeneity in τ_{it} will result in heterogeneity in ARDPs as well, which might be an important factor as policy makers have long recognized the importance of R&D to economic growth and there is a long history of public financial support (Bryan and Williams, 2021). Note, however, that most of the support has been either delivered via state-level subsidies or specific federal R&D grants, which tend to be focused on small and medium sized enterprises. Given that my sample is composed of mostly large firms, it is unclear ex-ante how important variation in τ_{it} might be for ARDP dispersion.

I investigate the importance of subsidies for ARDP dispersion with two complementary approaches. Firstly, I adjust R&D expenditure for investment tax credits, which recovers the prediction of no ARDP dispersion in the model. Secondly, I create a range of measures of government exposure and investigate how much variation in ARDP they can explain.

Firms in Compustat report investment tax credits in their income statement (`itci`), which lessen their tax burden and thus can be interpreted as subsidies to investment. These subsidies could be due to capital or R&D investments and, thus, are not necessarily attributable to R&D. I investigate two different approaches to allocating them by either allocating a fixed share to R&D or using the share of R&D in total investment:[8]

$$\texttt{TC}_{it}^{Fixed,\alpha} \equiv \alpha \times \texttt{itci} \qquad \text{and} \quad \texttt{TC}_{it}^{Prop.} \equiv \frac{\texttt{xrd}_{it}}{\texttt{xrd}_{it} + \texttt{capxv}_{it}} \times \texttt{itci}_{it}. \tag{18}$$

For each of these measures of tax credit I then define the adjusted ARDP as

$$\texttt{ARDP}_{it}^{Tax} =\equiv \frac{\sum_{w=0}^{4} \texttt{Patent valuations}_{it+w}}{\sum_{w=-1}^{3} (\texttt{R\&D Expenditure}_{it+w} - \texttt{TC}_{it+w})}. \tag{19}$$

Appendix Table B.3 reports the associated result. In short, subtracting R&D tax credits does not move ARDP dispersion. The reason for this is simple: They are not large enough. The average company in my sample expends around 1.7 billion on R&D, while the average investment tax credit is around 1.6 million.

Naturally, tax credits are not the only tool through which governments subsidize R&D. Other measures include direct contracting or R&D grants, which are not captured on a company's balance sheet. I investigate this possibility using an indirect approach by exploiting industry-level or firm-level evidence of public involvement as proxies for implies subsidies. At the firm-level, I create an indicator for public involvement taking the value of 1 if any of the firm's patents in the period of consideration acknowledges public support in the patent text (Fleming et al., 2019). At the industry-level I calculate two complementary measures. Firstly, I calculate the share of industry R&D funded federally as calculated from the NSF National Patterns. Secondly, I calculate the share of patents associated with firms in a particular industry and time-window that acknowledge public support. I transform both measures into indicator variables using a threshold of 10% and 2% respectively, which ensure that I have a meaning number of affected firms.

Using this measure of government involvement I then estimate a simple regression equation via OLS:

[8] I measure investment using capital expenditure `capxv`. Total investment is defined as the some of R&D expenditure and capital investment: `rnd` + `capcv`.

$$\ln \texttt{ARDP}_{it} = \alpha + \beta \texttt{Government involvement}_{it} + \varepsilon_{it}. \tag{20}$$

If government support was a strong driver of returns via subsidies, we would expect $\beta < 0$ as I am over-counting the true R&D cost to the firm. In contrast, my results, as reported in Appendix Table B.4, suggest little to no correlation of government interest with ARDP and R^2s remain consistently below 1%. In conclusion, government subsidies, as far as they are captured by my investigation, do not appear to be a significant driver of ARDP dispersion.

Jointly, my findings in this section let me conclude that benign mechanisms are not a significant driver for ARDP dispersion.

Stylized Fact 3. *Realized ARDP dispersion is not driven by measurement error, misspecification, technology differences, or subsidies.*

4 ARDP Dispersion and Economic Growth

Having established ARDP dispersion as a stylized fact, we can then study its impact on economic growth. For this purpose, I will develop a general endogenous growth model linking ARDP dispersion to economic growth.

4.1 A Bare-bones Endogenous Growth Model

Consider a simple economy, where aggregate output is a function of the current state of technology A_t and labor input L:[9]

$$Y_t = A_t L. \tag{21}$$

The growth rate of productivity $g_t \equiv \frac{\dot{A}_t}{A_t}$ is determined by the innovation efforts of a unit mass of innovating firms via

$$g_t = \int_0^1 z_{it} \times \mathbb{E}_t[\lambda_{it} - 1] di, \tag{22}$$

where z_{it} is the arrival rate of innovations and $\mathbb{E}_t[\lambda_{it} - 1]$ is their expected productivity improvement. Firms have access to innovation production function

[9]Note that this formulation can accommodate static inefficiency as a wedge between first best $A_t^{\max}$ and realized A_t. I will ignore this wedge for the purpose of this paper.

$$z_{it} = \varphi_{it} \ell_{it}^{\frac{1}{1+\phi}}, \tag{23}$$

where φ_{it} is the firm's efficiency at producing ideas and ℓ_{it} its R&D labor input, which is hired at wage W_t. Firms receive expected value $\mathbb{E}_t[\mathcal{V}_{it}]$ from innovation and I will define R&D productivity γ_{it} as the product of R&D efficiency and the value created: $\gamma_{it} \equiv \varphi_{it}\mathbb{E}_t[\mathcal{V}_{it}]$. Firm's face idiosyncratic R&D wedges $\frac{1}{\xi_{it}}$ that enter multiplicative to total R&D cost. The firm's problem is thus given by

$$\max_{\ell_{it}} \left\{ \gamma_{it} \ell_{it}^{\frac{1}{1+\phi}} - \left(\frac{1}{\xi_{it}}\right) W_t \ell_{it} \right\}. \tag{24}$$

The firm order conditions of the firm's problem result in the familiar expression

$$\frac{z_{it}\mathbb{E}_t[\mathcal{V}_{it}]}{W_t \ell_{it}} = \frac{1+\phi}{\xi_{it}} \tag{25}$$

Finally, there is a fixed mass of R&D workers $\mathcal{L}$, which pins down the wage of R&D workers W_t via the labor market clearing condition

$$\mathcal{L} = \int_0^1 \ell_{it} di. \tag{26}$$

Proposition 1. *Suppose that the ratio between expected productivity impact and private value created is constant*

$$\zeta \equiv \frac{\mathbb{E}_t[\lambda_{it} - 1]}{\mathbb{E}_t[\mathcal{V}_{it}]}. \tag{27}$$

Then, we can express the economic growth rate in the model described above as

$$g = g^{\max} \times \Xi, \tag{28}$$

where $g^{\max}$ is the growth-rate in absence of ARDP dispersion via ξ_{it} and allocative efficiency Ξ is given by

$$\Xi^{\phi} \equiv \frac{\left(\int_0^1 \omega_{it} \times \xi_{it}^{\frac{1}{\phi}} di\right)^{\phi}}{\left(\int_0^1 \omega_{it} \times \xi_{it}^{\frac{1+\phi}{\phi}} di\right)^{\frac{\phi}{1+\phi}}}, \quad \textit{with} \quad \omega_{it} \equiv \frac{\gamma_{it}^{\frac{1+\phi}{\phi}}}{\int_0^1 \gamma_{it}^{\frac{1+\phi}{\phi}} di}. \tag{29}$$

The formula for $g^{\max}$ is given by

$$g^{\max} = \mathcal{L}^{\frac{1}{1+\phi}} \times \zeta \times \left(\int_0^1 \gamma_{it}^{\frac{1+\phi}{\phi}} di \right)^{\frac{\phi}{1+\phi}}. \tag{30}$$

Corollary 1. *Allocation efficiency Ξ achieves its maximum of 1 if there is no dispersion in ξ_{it}. Otherwise, $\Xi \in (0,1)$.*

Corollary 2. *If ω_{it} and ξ_{it} are jointly log-normal, then allocation efficiency reduces to*

$$\Xi \overset{log.\ norm.}{=} \exp\left(-\frac{1}{2\phi} \times \sigma^2(\ln \xi_{it}) \right). \tag{31}$$

Naturally, the question arises as to the class of models covered when imposing $\zeta_{it} = \zeta$. One can show that this class includes standard expanding variety models and workhorse Schumpeterian growth models including those covered in Aghion et al. (2014). It does not, however, include models in the recent literature including de Ridder (2021) and Aghion et al. (2022). The former introduces heterogeneity in ζ_{it} by assuming that firms have differential static productivity, which gives rise to heterogeneous markups and R&D depreciation rates. This extension gives rise to variation in the value of innovation to firm, but not to the growth maximizing planner, which results in variation in ζ_{it}. Similarly, one can achieve variation in ζ_{it} by assuming heterogeneous step-sizes across firms in standard Schumpeterian growth models as covered in Aghion et al. (2014) and Acemoglu and Cao (2015). The variation in ζ_{it} in this case reflects the fact that firms undervalue breakthrough innovations relative to the planner due to their profit function.

In practice and in light of the discussion above, we might be skeptical of a constant ζ across firms and might want to consider a more general case with ζ_{it}. In this case it will be useful to define social R&D productivity $\tilde{\gamma}_{it}$ as $\tilde{\gamma}_{it} \equiv \gamma_{it} \times \zeta_{it}$. The following Proposition considers this more general case.

Proposition 2. *Denote the ratio between expected productivity impact and private value as ζ_{it}:*

$$\zeta_{it} \equiv \frac{\mathbb{E}_t[\lambda_{it} - 1]}{\mathbb{E}_t[\mathcal{V}_{it}]}. \tag{32}$$

We can express the economic growth rate in the model described above as

$$g = \tilde{g}^{\max} \times \tilde{\Xi}, \tag{33}$$

where $\tilde{g}^{\max}$ is the growth-rate in absence of social ARDP dispersion via ξ_{it}/ζ_{it} and allocative efficiency $\tilde{\Xi}$ is given by

$$\tilde{\Xi}^{\phi} \equiv \frac{\left(\int_0^1 \tilde{\omega}_{it} \times (\xi_{it}/\zeta_{it})^{\frac{1}{\phi}} di\right)^{\phi}}{\left(\int_0^1 \tilde{\omega}_{it} \times (\xi_{it}/\zeta_{it})^{\frac{1+\phi}{\phi}} di\right)^{\frac{\phi}{1+\phi}}}, \quad \textit{with} \quad \tilde{\omega}_{it} \equiv \frac{\tilde{\gamma}_{it}^{\frac{1+\phi}{\phi}}}{\int_0^1 \tilde{\gamma}_{it}^{\frac{1+\phi}{\phi}} di}. \tag{34}$$

The formula for $\tilde{g}^{\max}$ is given by

$$\tilde{g}^{\max} = \mathcal{L}^{\frac{1}{1+\phi}} \times \left(\int_0^1 \tilde{\gamma}_{it}^{\frac{1+\phi}{\phi}} di\right)^{\frac{\phi}{1+\phi}}. \tag{35}$$

Corollary 3. *Allocation efficiency $\tilde{\Xi}$ achieves its maximum of 1 if there is no dispersion in ξ_{it}/ζ_{it}. Otherwise, $\tilde{\Xi} \in (0,1)$. Note that this implies that dispersion in ξ_{it} is efficient as long as decreases dispersion in ξ_{it}/ζ_{it}. In particular, $g = g^{\max}$ when $\xi_{it} \propto \zeta_{it}$.*

Corollary 4. *If ω_{it}, ζ_{it}, and ξ_{it} are jointly log-normal, then allocation efficiency reduces to*

$$\tilde{\Xi} \overset{log.\ norm.}{=} \exp\left(-\frac{1}{2\phi} \times \left(\sigma^2(\ln \xi_{it}) + \sigma^2(\ln \zeta_{it}) - 2\rho(\ln \xi_{it}, \ln \zeta_{it}) \cdot \sigma(\ln \zeta_{it}) \cdot \sigma(\ln \xi_{it})\right)\right). \tag{36}$$

This formulation clarifies the circumstances under which heterogeneity in ζ_{it} will increase or decrease allocative efficiency. In particular, we have that Ξ will be larger than $\tilde{\Xi}$ iff

$$\rho(\ln \xi_{it}, \ln \zeta_{it}) > \frac{1}{2} \frac{\sigma(\ln \zeta_{it})}{\sigma(\ln \xi_{it})}. \tag{37}$$

Naturally, this can never hold if $\ln \zeta_{it}$ twice as dispersed as ξ_{it} or if $\rho(\ln \xi_{it}, \ln \zeta_{it})$ is negative.

Appendix D reports additional results relaxing the assumption of a constant mass of firms. I find that my results provide an lower bound for the growth rate impact of ξ_{it} dispersion as long as the mass of firms is not increasing in ARDP dispersion.

4.2 Mapping data to model

In order to understand the quantitative implications of the facts documented in Section 3, I will combine data and model. I will create two measures of allocation efficiency linking them to the insights in both Propositions above. I will estimate "private" allocative efficiency Ξ_t for a given year t as

$$\widehat{\Xi}_t^{\phi} \equiv \frac{\left(\sum \widehat{\omega}_{it} \times \widehat{\xi_{it}^{\frac{1}{\phi}}}\right)^{\phi}}{\left(\sum \widehat{\omega}_{it} \times \widehat{\xi_{it}^{\frac{1+\phi}{\phi}}}\right)^{\frac{\phi}{1+\phi}}}, \quad \text{with} \quad \widehat{\omega}_{it} \equiv \frac{\widehat{\gamma}_{it}^{\frac{1+\phi}{\phi}}}{\sum \widehat{\gamma}_{it}^{\frac{1+\phi}{\phi}}}. \tag{38}$$

Summations are taken above all active firms within a year. Similarly, I will estimate "public" allocative efficiency $\tilde{\Xi}_t$ as

$$\widehat{\tilde{\Xi}}_t^{\phi} \equiv \frac{\left(\sum \widehat{\tilde{\omega}}_{it} \times (\widehat{\xi}_{it}/\widehat{\zeta}_{it})^{\frac{1}{\phi}}\right)^{\phi}}{\left(\sum \widehat{\tilde{\omega}}_{it} \times (\widehat{\xi}_{it}/\widehat{\zeta}_{it})^{\frac{1+\phi}{\phi}}\right)^{\frac{\phi}{1+\phi}}}, \quad \text{with} \quad \widehat{\tilde{\omega}}_{it} \equiv \frac{\widehat{\tilde{\gamma}}_{it}^{\frac{1+\phi}{\phi}}}{\sum \widehat{\tilde{\gamma}}_{it}^{\frac{1+\phi}{\phi}}}. \tag{39}$$

Based on annual estimates I will then create summary measures by taking simple average across sample years.

Two implement these formulas I need two ingredients: Firstly, I need to determine a value for ϕ. I will follow the literature in setting $\phi = 1$ and provide robustness checks for alternative values.[10] Secondly, I need estimates for $\{\hat{\xi}_{it}, \hat{\zeta}_{it}, \hat{\gamma}_{it}, \hat{\tilde{\gamma}}_{it}\}$. Importantly, the formulas are insensitive to scaling by construction and I will, thus, only need to measure these parameters up to scale.

Measuring $\hat{\xi}_{it}$ is straight-forward and can be done directly from the ARDPs. I will follow the approach taken in Section 3 and focus on 5-year windows for measurement with the same sample restrictions.

$$\hat{\xi}_{it} \equiv \frac{\sum_{w=-1}^{3} \texttt{R\&D expenditure}_{it+w}}{\sum_{w=0}^{4} \texttt{Patent valuations}_{it+w}} = \frac{1}{\texttt{ARDP}_{it}} \tag{40}$$

Measuring ζ_{it} is a challenge as we do not observe the productivity impact of a patent or innovation in general. I will attempt to make progress on this issue by assuming the that productivity impact of proportional to patent forward citations with additional robustness using the text-based patent impact measure from Kelly et al. (2021).

$$\hat{\zeta}_{it} \equiv \frac{\sum_{w=0}^{4} \texttt{Forward-citations}_{it+w}}{\sum_{w=0}^{4} \texttt{Patent valuations}_{it+w}} \tag{41}$$

Note that interpreting forward citations as productivity impact is broadly in line with the empirical literature on patents (Bryan and Williams, 2021). Similarly, Kelly et al. (2021) de-

[10] See e.g. Acemoglu et al. (2018); de Ridder (2021); Akcigit and Kerr (2018) for other papers using this parametrization and Bloom et al. (2002); Blundell et al. (2002); Wilson (2009); Guceri (2018); Guceri and Liu (2019) for empirical papers estimating this value to be around 1.

velop their impact measure explicitly with the goal of capturing productivity gains. Nonetheless, neither approach is fully satisfactory and I will rely on it mainly as a robustness check as they only affect the formula for the "public" allocative efficiency.

To measure γ_{it} I will rely on the model first order conditions that imply

$$\gamma_{it} \propto \frac{1}{\xi_{it}} \times (W_t \ell_{it})^{\frac{\phi}{1+\phi}}. \tag{42}$$

Consequently, I will set

$$\hat{\gamma}_{it} \equiv \frac{1}{\hat{\xi}_{it}} \times \left(\sum_{w=-1}^{3} \texttt{R\&D expenditure}_{it+w} \right)^{\frac{\phi}{1+\phi}}. \tag{43}$$

Intuitively, in absence of ARDP dispersion, this measure simply states that firms conducting more R&D do so due to high productivity. Finally, "social" R&D productivity is simply the product of private productivity and the R&D spillover term and I will measure it accordingly:

$$\widehat{\hat{\gamma}}_{it} \equiv \hat{\gamma}_{it} \times \hat{\zeta}_{it}. \tag{44}$$

4.3 Growth Opportunities from ARDP Dispersion

Before presenting the estimates of allocative efficiency, I want to briefly caveat this discussion. The exercise assumes that the unaccounted for variation in ξ_{it} is driven by arbitrary, inefficiency sources. Naturally, there might be potentially efficient sources of variation that I have not identified in Section 3. In this sense, the estimates are thus a lower bound for allocative efficiency within the data. On the other hand, my stringent selection of observations might yield an upwards bias in allocative efficiency if it underestimates true variation in ARDPs. Hence, there remains some uncertainty around the estimated values.

The estimated values for $\hat{\Xi}$, as reported in Panel A Table 8, suggests a substantial gap between the growth rate g and its first-best value $g^{\max}$. For the entire sample period from 1975-2014 I find an average private $\hat{\Xi}$ of 76.1%, which suggests a potential improvement in the annual growth rate around 30%.[11] Panel B translates these values into growth rate gains against a baseline of $g = 1.5\%$, which suggest a potential increase in the annual growth rate around 0.47 percentage points.

[11]In particular, we have $\frac{g^{\max}-g}{g} = \frac{1-\hat{\Xi}}{\hat{\Xi}} = \frac{1-0.761}{0.761} \approx 30\%$.

Table 8: R&D Return Dispersion Matters for Economic Growth

Measure	Private	Public	
	Adjusted	Citations	Impact
Allocative Efficiency $\hat{\Xi}$	76.1%	68.6%	72.8%
Growth Cost	0.47 p.p.	0.69 p.p.	0.56 p.p.
Consumption Equivalent	60.3%	99.0%	75.0%

Note: "Growth Cost" report $g^{\max} - g$ assuming a baseline growth rate g of 1.5% per annum. "Consumption Equilvalent" translates this number into a consumption equivalent value for welfare on a Balanced Growth Path with logarithmic preferences over consumption using the formula $\Delta_C = \exp\left(\frac{g}{\rho}\Delta_g\right) - 1$. See Appendix C.3 for details on the welfare calculations.

The estimated values of "public" allocative efficiency tend to be lower, implying larger potential gains from moving towards first-best. The private estimate is thus conservative relative the public estimates. Using citations and impact to measure productivity impact implies a potential growth rate improvement of 0.69 pp. and 0.56 p.p. respectively, which are 46% and 19% larger than their "private" counterpart respectively. From the earlier discussion we know that, at least in the case of log-normality, the relationship between the private and public allocative efficiency measure depends on the correlation between $\hat{\xi}_{it}$ and $\hat{\zeta}_{it}$. Indeed, as shown in Appendix B.5, both variables are positively correlated or, in other words, $\hat{\zeta}_{it}$ is negatively correlated with $\texttt{ARDP}_{it}$. On the other hand, $\hat{\zeta}_{it}$ is much more dispersed such that the joint dispersion in $\hat{\xi}_{it}/\hat{\zeta}_{it}$ is still larger than the dispersion in $\hat{\xi}_{it}$ alone.

Robustness. I present two robustness checks in Appendix Table E.1. Firstly, larger values for ϕ reduce the losses from ARDP dispersion as they make firms' input choices less responsive to ξ_{it}. Nonetheless, the estimates remain large even when considering values far above the literature consensus. Secondly, replacing ω_{it} with uniform weights across active firms decreases allocative efficiency estimates, which suggest that there is less dispersion in ARDPs among the most productive firms.

4.4 Allocative Efficiency and the Growth Slow Down

Finally, I investigate the impact of the rising ARDP dispersion documented in Section 3. This finding seems especially important as Syverson (2017) and Gordon (2016) find that economic growth has slowed down in recent years. The former paper finds that labor productivity growth has slowed down from 1.54% p.a. in 1974-94 to 1.27% p.a. in 2005-2015. The question thus arises whether at least part of the rise in ARDP dispersion could be linked to declining growth.

Formally, I estimate $\hat{\Xi}$ for the 1974-94 and 2004-2014 period separately and calculate the contribution of changing allocative efficiency as

$$\text{Contribution of } \Xi \equiv 1 - \frac{(g^{\max}_{2005-15} - g^{\max}_{1974-94}) \times \hat{\Xi}_{1974-94}}{g_{2005-15} - g_{1974-94}} = \frac{\frac{\hat{\Xi}_{2004-14} - \hat{\Xi}_{1974-94}}{\hat{\Xi}_{2004-14}}}{\frac{g_{2005-15} - g_{1974-94}}{g_{2005-15}}} \tag{45}$$

Table 9 reports the associated results and finds that changing allocative efficiency has been an important contributor to declining growth. For the private measure I find that allocative efficiency has declined from around 80% to 72%, a 10% decline, which implies that declining allocative efficiency can account for around 50% of the growth slowdown. The public measures of allocative efficiency declined faster in absolute and relative terms and, thus, the model attributes and even larger fraction of the decline to changing allocative efficiency. Thus, as for the average growth rate impact, I find that the private values are a conservative estimate of the contribution of declining allocative efficiency to the growth slowdown.

Table 9: Contribution of Declining Allocative Efficiency to Growth Slowdown

Measure	Private	Public	
	Adjusted	Citations	Impact
Allocative Efficiency 1974-94	79.9%	73.4%	74.9%
Allocative Efficiency 2004-14	71.8%	61.5%	64.5%
Contribution to Declining Growth	52.8%	90.7%	75.5%

Note: This tables reports the cost of changing allocation efficiency comparing the 1975-1994 to the 2004-2014 period in percentage terms. The values are calculated as $\frac{\Xi_{74-94}/\Xi_{04-14}-1}{1.54/1.27-1}$.

Appendix Table E.1 reports estimates for alternative values of ϕ. As with my main results, I find that even values far above the empirical consensus find a meaningful role for declining allocative efficiency.

5 Mechanisms for ARDP Dispersion

A lingering question up to this point is which frictions drive dispersion in ARDPs. In this section I provide evidence that at part of the dispersion is driven by adjustment cost and labor market power. For this purpose, I will firstly provide some reduced for evidence in favor of both mechanisms and then estimate their contribution to ARDP dispersion within a quantitative Schumpeterian growth model.

5.1 Reduced Form Evidence

Productivity. Before investigating specific channels of misallocation, I want to provide some suggestive evidence that frictions might matter in general. For this purpose I will investigate the relationship between ARDP and R&D productivity γ_{it} as estimated via (43). This relationship is informative as many potential frictions, including adjustment cost and financial frictions, predict a positive correlation between this measure and ARDPs. The intuition is simple: A positive correlation indicates that the most productive firms don't do enough R&D, which in turns is a standard predictions of models with frictions.

I present two related pieces of evidence. Firstly, I show that R&D productivity is positively correlated with ARDPs. A challenge when establishing this correlation is measurement error as confounding factor. In particular, one might be concerned about measurement error in patent valuation, which would bias the estimated correlation upwards. I address this concern using an instrument variable strategy, using previous R&D productivity as an instrument. While this approach addresses any short-run measurement error concerns, it does not address potential time-invariant measurement error due to alternative accounting conventions or similar. I address this concern by adding firm fixed effects to the IV specification and show that the estimated correlation remains strong. Finally, I use changes in revenue instead of patent valuation as a value measure and confirm a strong, positive correlation.

Table 10: ARDPs and R&D Productivity

	(1)	(2)	(3)	(4)
	OLS	IV	IV	IV
	Realized ARPD$_{it}$			
`R&D Productivity`$_{it}$	0.624***	0.483***	0.534***	0.310**
	(0.028)	(0.029)	(0.078)	(0.123)
Firm FEs			✓	✓
Value Measure	Patent valuations	Patent valuations	Patent valuations	Δ Revenue
First Stage F statistic		16,970	53	65
Observations	11,839	7,623	7,573	7,459

Note: All regressions control for industry× year and firm fixed effects. Standard errors clustered at the 3-digit NAICS level.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

One of the challenges with the evidence above it that it provide a static correlation, while many frictions might operate more dynamically. For example, adjustment cost slow down the speed of adjustment between states, but do not necessarily affect the long-run allocation. Similarly, with many financial frictions, we might expect that they are temporary, while a firm is on the expansion, instead of permanent (Moll, 2014).

I test for these more dynamic effects by regression productivity growth on ARDPs. As shown in Column (1) of Table 11, I find a strong positive correlation between productivity growth and ARDP. I show that this correlation is not driven by correlated measurement error in patent valuations by replacing my main measure of ARDP with alternative measures using changes in revenue, employment, market valuation, and labor productivity instead of patent valuations in columns (2)-(5). This evidence is in line with an interpretation of ARDP as potentially temporary frictions in the sense that firms experiencing rapid productivity growth conduct insufficient R&D to equalize ARDPs.

Table 11: ARDPs and R&D Productivity Growth

	(1)	(2)	(3)	(4)	(5)
			$\texttt{Realized ARPD}_{it}$		
$\Delta_{t-5,t}\texttt{R\&D Productivity}_{it}$	0.503***	0.381***	0.381***	0.528***	0.499***
	(0.013)	(0.048)	(0.040)	(0.023)	(0.027)
Value Measure	Patent valuations	Δ Revenue	Δ Employment	Δ Market Valuation	Δ Labor Productivity
Observations	10,457	10,337	9,676	6,018	10,378

Note: All regressions control for industry× year and firm fixed effects. Standard errors clustered at the 3-digit NAICS level. Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

Having establish a general pattern in support of frictions as drivers of ARDPs, I next turn to two particular frictions: Adjustment cost and labor market power.

Adjustment cost. A natural candidate for variation in ARDPs are adjustment cost, which give rise to a first order condition of the form

$$\frac{z_{it}\mathbb{E}[\mathcal{V}_{it}]}{W_t \ell_{it}} = (1+\phi)(1+\lambda_{it}), \tag{46}$$

where λ_{it} captures the degree to which adjustment cost are preventing the firm from achieving its frictionless R&D input level. In general, λ_{it} will be positive for firms that have recently experiences an increase in its effective R&D productivity, either via the R&D efficiency or via the value of its innovation, and vice versa.

It is difficult to test for adjustment cost directly, but I offer three pieces of suggestive evidence. Firstly, I show that R&D expenditure growth is positively auto-correlated both at the 1-year and 5-year level as showing in Panel A of Table 12. Positive auto-correlation of growth rates is a moment that if often considered to be suggestive of adjustment cost as models without adjustment cost and AR(1) productivity processes predict negative auto-correlation.

Table 12: ARPDs and Adjustment Cost

	(1)	(2)
	Δ`R&D Expenditure`$_{it}$	
Previous Δ`R&D Expenditure`$_{it}$	0.136***	
	(0.021)	
Previous Δ`R&D Expenditure`$_{it}$		0.253***
		(0.020)
	`R&D Expenditure`$_{it}$	
Previous `R&D Expenditure`$_{it}$	0.831***	0.755***
	(0.020)	(0.029)
`R&D Productivity`$_{it}$		0.201***
		(0.026)
Observations	10,457	10,457

Note: This table reports coefficient estimates for . See text and Appendix for variable descriptions. The outcome variable is in logs. All regressions control for NAICS3 $\times$ year fixed effects.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

Secondly, I show that R&D expenditure is positively auto-correlated even conditional on current R&D productivity. Within the frictionless model, R&D productivity γ_{it} is the only firm-level determinant of R&D expenditure. It thus follows that previous and current R&D expenditure should not be correlated conditional on R&D productivity in absence of adjustment cost.

Panel B in Table 12 shows that previous R&D expenditure is indeed a strong predictor of current R&D expenditure, even conditional on R&D productivity. This evidence is thus another indication that adjustment cost might be important.

Heterogeneous Labor Market Power. Another potential mechanism that has received some attention in recently is labor market power Berger et al. (2022); Seegmiller (2021). In particular, if we allow the wage faced by a firm to respond to its labor demand, then we have

$$\frac{z_{it}\mathbb{E}[\mathcal{V}_{it}]}{W_t \ell_{it}} = (1+\phi)(1+1/\epsilon_{it}), \tag{47}$$

where $\epsilon_{it} \equiv \frac{\partial \ell_{it}}{\partial W_{it}}$ is the labor supply elasticity. This formulation clarifies that ARDP dispersion could be explained by firm-level variation in the labor supply elasticity. Importantly, Seegmiller (2021) documents just that. In his paper, he shows that firms with high labor productivity face more inelastic labor supply, especially so for high-skill workers. Motivated by this evidence conduct three sets of tests. Firstly, I follow Seegmiller (2021) and investigate whether firms high high labor productivity also have larger ARDPs. I construct two measures of labor productivity. Firstly, I follow the approach taken in Seegmiller (2021) and Donangelo et al. (2019) and define firm-level labor productivity as value added per worker:[12]

$$\texttt{Labor productivity}_{it} = \frac{\sum_{w=0}^{4} \texttt{Value Added}_{it+w}}{\sum_{w=0}^{4} \texttt{Employment}_{it+w}} \tag{48}$$

Secondly, I define a measure of inventor productivity as the value created divided by associated number inventors at the firm level. This measure thus directly speaks to the productivity that is crucial from an R&D perspective.

$$\texttt{Inventor productivity}_{it} = \frac{\sum_{w=0}^{4} \texttt{Patent Valuations}_{it+w}}{\sum_{w=0}^{4} \texttt{Inventors}_{it+w}}. \tag{49}$$

Columns (1) and (3) in Table 13 show that labor and inventor productivity is positively correlated with ARDPs, which is inline with lower labor supply elasticities for very productivity firms. Importantly, inventor productivity can explain a large share of the variation in ARDP as witness by the within-R^2 above 30%. A natural concern at this point is correlated measurement error via patent valuations, which could lead to an artificially large coefficient and R^2. I address this concern using two complementary approaches. Firstly, I use lagged inventor productivity as an instrument for current inventor productivity in column (4). The coefficient about 20% lower, but remains highly significant. Secondly, I use changes in value added as an alternative measure of value creation and confirm a highly significant relationship with ARDP again in column (5).

[12] I calculate value-added as in Donangelo et al. (2019), who define as value added as operation income before depreciation, interest, and taxes plus changes in inventory and labor expenditure. The latter is only available for a small share of observations and I impute it for the remaining firms using the methodology proposed in Donangelo et al. (2019).

Table 13: ARPDs and Labor Productivity

	(1)	(2)	(3)	(4)	(5)
	OLS	OLS	OLS	IV	OLS
			Realized ARPD$_{it}$		
Labor productivity	0.329**				
	(0.132)				
Labor share		-0.426***			
		(0.082)			
Inventor productivity			0.569***	0.455***	0.068***
			(0.060)	(0.083)	(0.018)
Value Definition			Patent Valuation	Patent Valuation	Δ Value Added
R2	0.299	0.312	0.517	0.314	0.286
Within R2	0.038	0.046	0.338		0.009
Observations	10,711	9,903	10,471	6,540	9,651

Note: This table reports coefficient estimates for . See text and Appendix for variable descriptions. All variables are in logs.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

In addition, I provide evidence that firms employing a larger share of researchers nationally and locally have higher ARDPs. I calculate the national inventor share as the share of active US-based inventors that work for a particular firm in a 5-year window from t to t+4 as witnessed by their patenting activities. Since I do not observe employment directly in the patent data, I instead attribute inventors to firms based on the share of their patents that is associated with the firm. I similarly calculate a local inventor share by first mapping the local of all inventors to commuting zones (CZ) via their FIPS code and calculating the CZ specific inventor share. I then take an average over this measure for commuting zones in which the firm has significant inventor employment.

Table 14 reports simple OLS regressions of these measures on ARDPs with left and right hand side variables in logs. Both the national and local inventor share are significantly associated with larger ARDPs. When including both at the same time in column (3) I find that the national share appears more important than the local share. This finding is qualitatively in line with the idea that the relevant market for high-skill inventors is national and not local. On the other hand, it is difficult to rule out that this relationship is not driven by pure productivity differences across firms that also correlate with the inventor employment shares. Note that this would still be in line with labor market power as long as it is linked to R&D productivity as documented in Seegmiller (2021).

Table 14: ARPDs and Inventor Concentration

	(1)	(2)	(3)
	OLS	OLS	OLS
		Realized ARPD$_{it}$	
National innovator share	0.199***		0.168***
	(0.034)		(0.040)
Local innovator share		0.108***	0.033*
		(0.020)	(0.016)
Within R2	0.058	0.036	0.060
Observations	10,471	10,471	10,471

Note: This table reports coefficient estimates for . See text and Appendix for variable descriptions. All variables are in logs.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

5.2 Quantitative Evidence

Model Setup. To estimate the contribution of adjustment cost and labor market power I will resort to a quantitative Schumpeterian growth model similar to Klette and Kortum (2004). In the interest of space I will focus on the main equations here and leave the model details to Appendix D. Time is discrete and indexed by t. There is a unit mass of firms, which innovate with probability z_{it} and improve upon existing technology in a random product line by $\lambda_{it} - 1$ if innovation is successful. The aggregate growth rate is given by

$$g = \int_0^1 z_{it} \times \ln(\lambda_{it}) di. \tag{50}$$

I will assume that $\lambda_{it} \equiv \lambda^{\Delta_{it}}$ is an i.i.d. random variable that is unknown to the firm before R&D expenditure is occurred and is distributed according to

$$P(\Delta) = \frac{(1-p)^{\Delta-1}p}{1-(1-p)^{\bar{\Delta}}} \qquad \text{for} \quad \Delta = 1, ..., \bar{\Delta}. \tag{51}$$

This formulation allows me to introduce some ex-ante uncertainty about the quality of the research output. Setting $\bar{\Delta} = 1$ recovers the deterministic case.

Firms produce innovation using R&D workers ℓ_{it} and differ in their R&D efficiency φ_{it}, which follows an AR(1) process in logs:

$$\ln \varphi_{it} = (1-\rho)\mu + \rho \ln \varphi_{it-1} + \varepsilon_{it} \qquad \varepsilon_{it} \sim N(0, \sigma^2). \tag{52}$$

The innovation production function is given by

$$z_{it} = \varphi_{it} \ell_{it}^{\frac{1}{1+\phi}}. \tag{53}$$

The normalized value from successful innovation is a function of the associated expected, discounted profits only such that

$$\mathbb{E}_\lambda[\mathcal{V}] = \frac{\mathbb{E}_\lambda[\pi(\lambda)]}{1-\beta(1-z)} \quad \text{with} \quad \pi = 1 - 1/\lambda. \tag{54}$$

Here $z \equiv \int_0^1 z_{it} di$ is the aggregate innovation rate, which adds to the discounting as all innovations replace existing products, and β is the household discount factor.

Finally, the cost of R&D are subject to two distortions: Adjustment cost and monopsony power. I model the latter in a reduced form by assuming that the individual wage faced by a firm W_{it} is given by

$$W_{it} = W_t \left(1 + \frac{\ell_{it}}{\bar{\ell}}\right)^{\xi^{LMP}}, \tag{55}$$

where W_t is a common factor that ensures overall labor market clearing. Note that $\xi^{LMP} = 0$ is the case of no labor market power. I choose this formulation as it implies that the elasticity of the wage with respect to own labor demand is increasing:

$$\frac{\partial \ln W_{it}}{\partial \ln \ell_{it}} = \xi^{LMP} \times \frac{\ell_{it}}{\bar{\ell} + \ell_{it}}. \tag{56}$$

This property allows me to fit the estimates in Seegmiller (2021) as firms with higher ℓ_{it} will also have higher R&D productivity.

I allow for quadratic adjustment costs in R&D labor, following the literature on capital adjustment costs:

$$AC(\ell_{it}, \ell_{it-1}) = \xi^{Adj.} W_{it} \ell_{it} \left(\frac{\ell_{it} - (1-\delta)\ell_{it-1}}{\ell_{it}}\right)^2. \tag{57}$$

Here, δ captures a natural outflow of R&D workers due to exogenous causes.

The cost of R&D is thus given by

$$C(\ell_{it}, \ell_{it-1}) = W_t \ell_{it} \left(1 + \xi^{Adj.} \left(\frac{\ell_{it} - (1-\delta)\ell_{it-1}}{\ell_{it}}\right)^2\right) \left(1 + \frac{\ell_{it}}{\bar{\ell}}\right)^{\xi^{LMP}}. \tag{58}$$

The firm's R&D value function maximization problem is given by

$$V(\varphi_{it}, \ell_{it-1}) = \max_{\ell_{it}} \{-C(\ell_{it}, \ell_{it-1}) + \beta z(\varphi_{it}, \ell_{it})\mathbb{E}_\lambda[\mathcal{V}] + \beta \mathbb{E}_t[V(\varphi_{it+1}, \ell_{it})]\}. \tag{59}$$

Note that this formulation does not include the flow profits from product lines in which the firm is currently the incumbent. I show in Appendix D that these are irrelevant for the choice of ℓ and I thus do not consider them in the value function maximization.

Finally, I close the model with the R&D labor supply constraint, which pins down W:

$$\mathcal{L} = \int_0^1 \ell_{it} di. \tag{60}$$

Recursive form. On a Balanced growth path, we can write down the model in recursive form, which allows me to solve it numerically. R&D cost become

$$C(\ell', \ell) = W\ell' \left(1 + \xi^{Adj.} \left(\frac{\ell' - (1-\delta)\ell}{\ell'}\right)^2\right) \left(1 + \frac{\ell'}{\bar{\ell}}\right)^{\xi^{LMP}}, \tag{61}$$

where W is the appropriately normalized, time-invariant wage. The value function optimization is given by

$$V(\varphi, \ell) = \max_{\ell'} \{-C(\ell, \ell') + \beta z(\varphi, \ell')\mathbb{E}_\lambda[\mathcal{V}] + \beta \mathbb{E}_t[V(\varphi', \ell')]\}. \tag{62}$$

Furthermore, the labor supply constraint expressed via the distribution over the state-space $f(\ell, \varphi)$ is given by

$$\mathcal{L} = \int \int \ell'(\varphi, \ell) f(\varphi, \ell) d\ell d\varphi, \tag{63}$$

where $f(\ell, \varphi)$ is the solution to

$$f(\ell', \varphi') = \int \int f(\ell, \varphi) \mathbb{I}\{\ell' = \ell(\varphi, \ell)\} d\ell f(\varphi'|\varphi) d\varphi. \tag{64}$$

Here, $f(\varphi'|\varphi)$ is the conditional density of the AR(1) process and $\ell(\varphi, \ell)$ is the equilibrium policy function mapping φ and ℓ into ℓ'.

Definition 1. *A Balanced Growth Path equilibrium is a value function $V(\varphi, \ell)$, policy func-*

tion $\{\ell'(\varphi,\ell)\}$, distribution $f(\varphi,\ell)$ and wage W such that the value and policy function solve the recursive problem (62) *subject to cost function* (61)*, the labor market clearing condition* (63) *holds, and the distribution satisfies* (64) *evaluated using the policy function.*

Solving and simulating the model. In practice I will discretize the state-space to solve the model numerically. For this I create a grid for φ and ℓ and discretize the AR(1) process for $\ln\varphi$ using standard methods. I detail the solution method in the Appendix.

When simulating the model, I interpret each successful innovation as a patent. To match the average number of patents per firm in my sample, I will simulate firms as collections of N_P innovation lines with identical productivity process and independent realization of inventions. I choose N_P such that the simulated sample matches the average number of patents in my sample. In particular, I set

$$N_P \approx \frac{\text{Avg. \# of patents per year in sample}}{z}, \tag{65}$$

which ensures that the expected number of patents for the simulated firm fits the data.

Estimation. The model has 11 parameters: $\{\beta, p, \lambda, \bar{\Delta}, \mu, \rho, \sigma, \phi, \xi^{Adj.}, \delta, \xi^{LMP}, \bar{\ell}\}$. I set μ such that the model always hits a target growth rate g^{Target} exactly. Of the remaining parameters I set 6 exogenously and estimate five.

Table 15 reports the exogenous parameters. In particular, I set β in line with the literature. To calibrate the ex-post uncertainty I set $p = 0.25$ and $\bar{\Delta} = 10$. I then calibrate λ such that $\mathbb{E}[\lambda] = 1.25$, i.e. the average innovation improves upon existing technology by 25%, which is similar to the calibration in Terry (2017).[13] I follow the literature and set $\phi = 1$ (Terry, 2017; Acemoglu et al., 2018; de Ridder, 2021).

Finally, I calibrate δ using data from the LEHD's Quarterly Workforce Indicators. In particular, I calculate the total employment, separation, and net-job gains for workers with at least a bachelors degree for the 2014Q1-2019Q4 period for all states and industries (4-digit NAICS) individually. For each combination I then calculate the ratio of separations minus net job losses divided by employment minus net job gains. I only factor in net job losses or gains when they are positive respectively. The ratio of both is then my estimate for δ. I average across all industries and years weighted by employment. My final estimate is 11.7% and I set $\delta = 0.12$ accordingly.

[13] In future work I will also estimate these parameters by targeting the average markup in my sample as well as the variance and skewness of patent valuations from Kogan et al. (2017) for the firms that ultimately end up in my sample.

Table 15: Exogenous parameters

Parameter	Value	Notes
β	0.97	See Acemoglu et al. (2018).
p	0.25	Set to roughly match dispersion
$\bar{\Delta}$	10	in patent valuations.
$\mathbb{E}[\lambda]$	1.25	See Terry (2017)
ϕ	1	See Acemoglu et al. (2018).
δ	0.12	See text.

I estimate the remaining five parameters $\{\rho, \sigma, \xi^{Adj.}, \xi^{LMP}, \bar{\ell}\}$ via SMM using five moments: The variance of R&D growth, the auto-correlation of R&D productivity, the auto-correlation of R&D growth, and the top and bottom quartile of the labor supply elasticities estimated in Seegmiller (2021). The variance of R&D growth together with the auto-correlation of R&D productivity is informative about σ^2 and ρ. The auto-correlation of R&D growth in turn informs the degree of adjustment costs. Finally, the labor supply elasticities allow me to pin down the parameters of the wage function.

I simulate the first three moments in the model following the same steps as in the data. I aggregate annual observations to 5-year windows and then calculate the relevant moments.

I used the total squared distance normalized by target values as my objective function and solve

$$\min \sum_{i=1}^{5} \left(\frac{\text{Target moment}_i - \text{Simulated moment}_i}{\text{Target moment}_i} \right)^2 . \tag{66}$$

Model fit. Table 16 reports the estimated parameters.[14] Importantly, the estimates suggest significant adjustment cost and labor market power. The associated moments and its targets are reported in Table 17. The model overall fits the targets reasonably well, however, the auto-correlation of R&D growth is somewhat too low, while the wage elasticity for top firms is too large.

[14] I'm still working on calculating the standard errors.

Table 16: Parameter Estimates

Parameter	Description	Estimate
σ	Std. dev. of shocks	0.321
ρ	Auto-correlation of R&D productivity	0.963
$\xi^{Adj.}$	Adjustment cost	1.415
ξ^{LMP}	LMP exponent	10.249
$\bar{\ell}$	LMP shifter	0.269

Table 17: Model Fit

	Target	Model
SD(Δ R&D)	0.354	0.322
Auto-correlation of R&D	0.715	0.703
Auto-correlation of R&D Growth	0.215	0.160
Wage elasticity, bottom quartile	0.490	0.417
Wage elasticity, top quartile	1.361	1.558

Results. How important are adjustment cost and labor market power jointly in explaining ARDP dispersion? To answer this question I simulate the model and repeat the data construction steps from the data exercise. I then compare my findings to the simulated data when setting $\xi^{Adj.} = \xi^{LMP} = 0$. Table 18 reports the results for ARDP dispersion and auto-correlation. The model with frictions has an ARDP dispersion around 0.22 compared to 0.09 for the frictionless model. Thus, adjustment cost and labor market power jointly around 0.13 to the standard deviation or around 14% of the empirical value. Notably, adding either friction alone already account for around 10% of the empirical standard deviation. Thus, both frictions interact and the joint contribution is smaller than the sum of its parts.

Another important empirical feature is the auto-correlation of ARDPs. The second row of Table 18 highlights that this is also a characteristic feature of the estimated model, while it is non-existent in the model without frictions. Column (3) and (4) further highlight that labor market power is particularly important in explaining this feature of the data. Both mechanism, thus, not only explain the part of the variation in ARDPs, but also its auto-correlation.

Table 18: ARDP Properties from Simulation

	Baseline	No Frictions	No Adj. Cost	No LMP
Std. dev. of log ARDP	0.223	0.087	0.190	0.217
Auto-correlation of ARDP	0.624	0.011	0.636	0.161

Finally, Table 19 reports a set of untargeted regression coefficients that are informative about the mechanisms. The first row reports the results for a simple OLS regression of log ARDP on log productivity. In line with the data, the model with frictions find a strong coefficient. In contrast, the model without frictions finds no correlation. This finding is also supported using an IV approach leveraging past productivity as an instrument for current productivity. Comparing columns (3) and (4) then highlights that LMP is key for this relationship. While the OLS coefficient in the model with only adjustment cost is positive, it is not in the IV specification. In contrast, the coefficient in the model with only labor market power is virtually unaffected by the IV approach. On the other hand, the simulated data also supports the finding of auto-correlation of R&D conditional on productivity. Columns (3) and (4) suggest that this finding is driven by adjustment cost. Note, however, that this auto-correlation is much weaker than in the data.

Table 19: Untargeted Regression Coefficients

	Baseline	No Frictions	No Adj. Cost	No LMP
ARDP on Productivity (OLS)	0.329	0.005	0.285	0.137
ARDP on Productivity (IV)	0.288	-0.008	0.273	-0.058
Conditional auto-correlation of R&D	0.122	0.018	0.034	0.296

5.3 Other mechanisms.

Before concluding I want two highlight two additional pieces of analysis. Firstly, I investigate the extend of financial frictions as a driver of ARDP dispersion in Appendix B.5. My findings suggest that they are not an important driver, similar to the findings in Midrigan and Xu (2014) for capital allocation. Given that my sample consistent of large, listed US firms, this finding is not particularly surprising as we would not expect them to be constrained.

Secondly, not all dispersion in ARDPs might be harmful in the first place. In particular, part of ARDP dispersion is driven by companies creating a lot of public per private value.

As an economy, we might be better off if firms with large positive externalities of their R&D do too much of it from a private value maximization perspective. I investigate this possibility in Appendix B.5 and indeed find low ARDPs for firms creating a lot of public value per private value, which is in line with them conducting more R&D than warranted from a private perspective. This finding suggests that part of the ARDP dispersion might be beneficial from a public perspective. Note, however, that a full account of public value tends to decrease my estimates of Ξ in the previous section. The reason for this finding is that the value-impact wedge ζ_{it} is so dispersed that it the total dispersion in ζ_{it}/ξ_{it} is larger than the dispersion in $1/\xi_{it}$ even though ζ_{it} and ξ_{it} are positively correlated.

6 Conclusion

This paper documents that ARDPs, defined as the value created from innovation divided by the associated cost, are highly dispersed among US listed, innovation-intensive firms and that this dispersion has increased over time. This finding is surprising ex-ante as workhorse models of endogenous growth theory predict ARDP equalization across firms. After ruling out benign drivers of ARDP dispersion such as measurement error, misspecification, and technology differences across firms, I argue that ARDP dispersion can be interpreted as misallocation with ARDPs measuring frictions.

Following this interpretation, I develop a growth rate decomposition for a general class of endogenous growth models that allows me to directly link the empirical distribution of ARDPs to economic growth. The model growth rate is the product of a first-best growth rate and an adjustment factor for frictions that is decreasing in ARDP dispersion. Combing data and model I estimate the hypothetical first-best growth rate is 30% larger than its actual value for the US, as 0.47 p.p. increase in annual growth against a baseline of 1.5% p.a. Furthermore, I find that increasing ARDP dispersion and the associated decline in allocative efficiency can account for up to 50% of the declining long-run growth rate documented in Syverson (2017).

To shine more light on the origins of ARDP dispersion, I investigate two mechanisms directly: Adjustment cost and labor market power. I find robust reduced form evidence for both. I document positive auto-correlation R&D growth rates as well as for R&D expenditure conditional on R&D productivity. Both are signs of adjustment cost in standard models. Furthermore, I find that ARDPs are positively correlated with inventor productivity, which has been linked to labor market power in Seegmiller (2021), and concentration in the market for inventors. To estimate the contribution of both mechanisms, I propose a quantitative Schumpeterian growth model and estimate it via SMM. The estimated parameters suggest a robust role for both mechanisms. I find that the mechanisms jointly explain about 14% of the overall and 26% of the within firm dispersion in empirical ARDPs.

Naturally, this finding still leaves a large share of the variation unexplained leaving room for future research to investigate this dispersion further. My findings suggest that mechanism linking ARDP dispersion to firm productivity are particularly fruitful avenue as both variables are strongly positively correlated, suggesting that the most productive firms are not conducting enough R&D. One such channel could be a concern for future regulation as firms target a smaller than optimal size to avoid public scrutiny. Similarly, one could imagine a human capital based mechanism where researchers at highly innovative firms receive large outside offers by competitors aiming to emulate the firm, reducing the firm's incentive to expand their R&D. A final, and equally plausible, possibility is non-homotheticicities making the R&D production function more concave for high levels of R&D.

References

Acemoglu, Daron and Dan Cao, "Innovation by entrants and incumbents," *Journal of Economic Theory*, 2015, *157*, 255–294.

_ **and David Autor**, "Skills, Tasks and Technologies: Implications for Employment and Earnings," in "Handbook of Labor Economics," Vol. 4, Elsevier, 2011, pp. 1043–1171.

_ **and Ufuk Akcigit**, "Intellectual property rights policy, competition and innovation," *Journal of the European Economic Association*, 2012, *10* (1), 1–42.

_ **, _ , Harun Alp, Nicholas Bloom, and William R. Kerr**, "Innovation, Reallocation, and Growth," *American Economic Review*, 2018, *108* (11), 3450–3491.

Aghion, Philippe and Peter Howitt, "A Model of Growth Through Creative Destruction," *Econometrica*, 1992, *60* (2), 323–351.

_ **, Antonin Bergeaud, Timo Boppart, Peter J. Klenow, and Huiyu Li**, "A Theory of Falling Growth and Rising Rents," 2022.

_ **, Ufuk Akcigit, and Peter Howitt**, *What Do We Learn From Schumpeterian Growth Theory?*, Vol. 2, Elsevier B.V., 2014.

Akcigit, Ufuk and Sina T. Ates, "Ten Facts on Declining Business Dynamism and Lessons from Endogenous Growth Theory," *American Economic Journal: Macroeconomics*, 2021, *13* (1), 257–298.

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

_ **, Douglas Hanley, and Nicolas Serrano-Velarde**, "Back to Basics: Basic Research Spillovers, Innovation Policy, and Growth," *Review of Economic Studies*, 2021, *88* (1), 1–43.

Aksoy, Yunus, Henrique S Basso, Ron P Smith, and Tobias Grasl, "Demographic structure and macroeconomic trends," *American Economic Journal: Macroeconomics*, 2019, *11* (1), 193–222.

Andrews, Dan, Chiara Criscuolo, and Peter N. Gal, "The Best Versus the Rest: The Global Productivity Slowdown, Divergence Across Firms and the Role of Public Policy," *OECD Productivity Working Papers*, 2016, *November* (05).

Asker, John, Allan Collard-Wexler, and Jan De Loecker, "Dynamic Inputs and Resource (Mis)Allocation," *Journal of Political Economy*, 2014, *122* (5), 1013–1063.

Banerjee, Abhijit V and Benjamin Moll, "Why Does Misallocation Persist?," *American Economic Journal: Macroeconomics*, 2010, *2* (1), 189–206.

Berger, David W., Kyle Herkenhoff, and Simon Mongey, "Labor Market Power," *American Economic Review*, 2022, *forthcomin.*

Berkes, Enrico, "Comprehensive Universe of U.S. Patents (CUSP): Data and Facts," *Working paper*, 2016, pp. 1–27.

Bloom, Nicholas, Charles I. Jones, John Van Reenen, and Michael Webb, "Are Ideas Getting Harder to Find?," *American Economic Review*, 2020, *110* (4), 1104–1144.

_ , Rachel Griffith, and John Van Reenen, "Do R & D tax credits work? Evidence from a panel of countries 1979-1997," *Journal of Public Economics*, 2002, *85* (1), 1–31.

Blundell, Richard, Rachel Griffith, and Frank Windmeijer, "Individual effects and dynamics in count data models," *Journal of Econometrics*, 2002, *108* (1), 113–131.

Bryan, Kevin A. and Heidi L. Williams, *Innovation: market failures and public policies*, Vol. 5, Elsevier B.V., 2021.

Brynjolfsson, Erik, Daniel Rock, and Chad Syverson, "Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics," in "The Economics of Artificial Intelligence: An Agenda" 2019, pp. 23–57.

David, Joel M. and Venky Venkateswaran, "The sources of capital misallocation," *American Economic Review*, 2019, *109* (7), 2531–2567.

_ , Hugo A. Hopenhayn, and Venky Venkateswaran, "Information, Misallocation, and Aggregate Productivity," *The Quarterly Journal of Economics*, may 2016, *131* (2), 943–1005.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger, "The Rise of Market Power and the Macroeconomic Implications," *The Quarterly Journal of Economics*, 2020, *135* (2), 561–644.

de Ridder, Maarten, "Market Power and Innovation in the Intangible Economy," 2021.

Dechezleprêtre, Antoine, Elias Einiö, Ralf Martin, Nguyen Kieu-Trang, and John Van Reenen, "Do Tax Incentives for Research Increase Firm Innovation ? An RD Design for R&D," 2019.

Donangelo, Andres, François Gourio, Matthias Kehrig, and Miguel Palacios, "The cross-section of labor leverage and equity returns," *Journal of Financial Economics*, 2019, *132* (2), 497–518.

Engbom, Niklas, "Aging , Labor Market Dynamics , and Growth," 2020, pp. 1–71.

Ewens, Michael, Ryan Peters, and Sean Wang, "Measuring Intangible Capital with Market Prices," 2020, (September 2018).

Fleming, L., H. Greene, G. Li, M. Marx, and D. Yao, "Government-funded research increasingly fuels innovation," *Science*, 2019, *364* (6446), 1139–1141.

Gancia, Gino and Fabrizio Zilibotti, "Horizontal Innovation in the Theory of Growth and Development," *Handbook of Economic Growth*, 2005, *1* (SUPPL. PART A), 111–170.

Gordon, Robert J, *The Rise and Fall of American Growth*, Princeton University Press, 2016.

Guceri, Irem, "Will the real R&D employees please stand up? Effects of tax breaks on firm-level outcomes," *International Tax and Public Finance*, 2018, *25* (1), 1–63.

_ **and Li Liu**, "Effectiveness of fiscal incentives for R & D: Quasi-experimental evidence," *American Economic Journal: Economic Policy*, 2019, *11* (1), 266–291.

Hsieh, Chang-Tai and Peter J. Klenow, "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 2009, *124* (4), 979–1014.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy, "Measuring Technological Innovation over the Long Run," *American Economic Review: Insights*, sep 2021, *3* (3), 303–20.

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

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, "Technological innovation, resource allocation, and growth," *The Quarterly Journal of Economics*, 2017, pp. 665–712.

Lehr, Nils H, "Innovation in an Aging Economy," 2022.

Liu, Ernest, Atif Mian, and Amir Sufi, "Low Interest Rates, Market Power, and Productivity Growth," *Econometrica*, 2022, *90* (1), 193–221.

Lucking, Brian, "Do R&D Tax Credits Create Jobs?," 2019.

Mezzanotti, Filippo, "Roadblock to Innovation: The Role of Patent Litigation in Corporate R&D," *Management Science*, 2021, (June).

Midrigan, Virgiliu and Daniel Yi Xu, "Finance and misallocation: Evidence from plant-level data," *American Economic Review*, 2014, *104* (2), 422–458.

Moll, Benjamin, "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?," *American Economic Review*, 2014, *104* (10), 3186–3221.

Ottonello, Pablo and Thomas Winberry, "Financial Heterogeneity and the Investment Channel of Monetary Policy," *Econometrica*, 2020, *88* (6), 2473–2502.

Peters, Michael, "Heterogeneous Markups, Growth, and Endogenous Misallocation," *Econometrica*, 2020, *88* (5), 2037–2073.

_ **and Conor Walsh**, "Population Growth and Firm Dynamics," 2021.

Peters, Ryan H. and Lucian A. Taylor, "Intangible capital and the investment-q relation," *Journal of Financial Economics*, 2017, *123* (2), 251–272.

Philippon, Thomas and Germán Gutiérrez, "Investmentless Growth: An Empirical Investigation," *Brookings Papers on Economic Activity*, 2017.

Restuccia, Diego and Richard Rogerson, "Policy distortions and aggregate productivity with heterogeneous establishments," *Review of Economic Dynamics*, 2008, *11* (4), 707–720.

Romer, Paul M, "Endogenous Technological Change," *Journal of Political Economy*, 1990, *98* (5).

Seegmiller, Bryan, "Valuing Labor Market Power : The Role of Productivity Advantages," 2021, (November).

Syverson, Chad, “Challenges to Mismeasurement Explanations for the US Productivity Slowdown,” *Journal of Economic Perspectives*, may 2017, *31* (2), 165–186.

Terry, Stephen J., “The Macro Impact of Short-Termism,” 2017.

_ , Toni M. Whited, and Anastasia A. Zakolyukina, “Information Versus Investment,” 2021.

Whited, Toni M and Guojun Wu, “Financial Constraints Risk,” *The Review of Financial Studies*, 2006, *19* (2), 531–559.

Wilson, Daniel J., “Beggar Thy Neighbor? The In-State, Out-of-State, and Aggregate Effects of R&D Tax Credits,” *The Review of Economics and Statistics*, 2009, *91* (2), 431–436.

Appendix

A Data Appendix

B Empirical Appendix

B.1 Measurement Error

Methodology. I investigate the extend of measurement error parametrically using an auto-correlation framework. I first present the framework and GMM estimation procedure followed by the results for ARDP and R&D productivity.

Let $y_{i,t}$ be the variable of interest index by firm i and time t. The process follows an AR(1) process with persistence ρ, firm fixed effects μ_i, and innovation variance σ_ε^2:

$$y_{it} = (1-\rho)\mu_i + \rho y_{it} + \sigma_\varepsilon \varepsilon_{it} \quad \text{with} \quad \varepsilon_{it} \overset{i.i.d.}{\sim} N(0,1). \tag{B.1}$$

Instead of observing the process directly, the researchers only observes $\tilde{y}_{it}$, which includes iid normal measurement error ν_{it} with variance σ_ν^2:

$$\hat{y}_{it} = y_{it} + \sigma_\nu \nu_{it} \quad \text{with} \quad \nu_{it} \overset{i.i.d.}{\sim} N(0,1). \tag{B.2}$$

Given this setup, we are interested in estimating in the parameter vector $\beta \equiv \{\rho, \sigma_\mu, \sigma_\varepsilon, \sigma_\nu\}$, where σ_μ^2 is the cross-sectional variance of μ_i. Once we have our estimates, we are able to perform a complete variance decomposition:

$$Var(\tilde{y}_{i,t}) = \underbrace{\sigma_\mu^2 + \frac{\sigma_\varepsilon^2}{1-\rho^2}}_{=Var(y_{i,t})} + \sigma_\nu^2.$$

The proposed estimation strategy below relies on the auto-covariance structure of y_{it}, which requires me to make the following assumption:

Assumption 1. *$y_{i,t}$ is auto-correlated and stationary: $|\rho| \in (0,1)$.*

It will be useful to define the following population moments:

$$
\begin{aligned}
m_1 &\equiv \mathrm{Cov}(\tilde{y}_{i,t}, \Delta\tilde{y}_{i,t}) = \frac{1}{1+\rho}\sigma_\varepsilon^2 + \sigma_\nu^2 \\
m_2 &\equiv \mathrm{Cov}(\tilde{y}_{i,t}, \Delta\tilde{y}_{i,t-w}) = \frac{\rho}{1+\rho}\sigma_\varepsilon^2 \\
m_3 &\equiv \mathrm{Cov}(\tilde{y}_{i,t}, \Delta\tilde{y}_{i,t-w}) = \frac{\rho^2}{1+\rho}\sigma_\varepsilon^2 \\
m_4 &\equiv \mathrm{Cov}(\tilde{y}_{i,t}, \tilde{y}_{i,t-1}) = \sigma_\mu^2 + \frac{\rho}{1-\rho^2}\sigma_\varepsilon^2.
\end{aligned}
$$

Under this conditions, one can solve for the parameters of interest using the population auto-covariance structure:

Proposition 3. *Under Assumption 1, we can solve for $\{\rho, \sigma_\mu, \sigma_\varepsilon, \sigma_\nu\}$ using the population auto-covariance structure of y_{it} and $\Delta y_{it} \equiv y_{it} - y_{it-1}$:*

$$
\beta \equiv \begin{bmatrix} \rho \\ \sigma_\varepsilon^2 \\ \sigma_\mu^2 \\ \sigma_\nu^2 \end{bmatrix} = \begin{bmatrix} \frac{m_3}{m_2} \\ \frac{(m_2)^2}{m_3} + m_2 \\ m_4 - \frac{(m_2)^2}{m_2 - m_3} \\ m_1 - \frac{(m_2)^2}{m_3} \end{bmatrix}
$$

.

Furthermore, under standard LLN assumptions, the empirical counterpart $\hat{\beta}$ converges to a normal distribution according to

$$
\sqrt{N}(\hat{\beta} - \beta) \to N(0, \Sigma) \qquad \textit{with} \quad \Sigma = \left(\frac{\partial\beta}{\partial m}\right)\Omega\left(\frac{\partial\beta}{\partial m'}\right), \tag{B.3}
$$

where Ω is the asymptotic variance of the moment vector $m = \{m_1, m_2, m_3, m_4\}$. We can estimate $\hat{\Omega}$ with its within sample counterpart.

Proof. For the first part, it is straightforward to show within this framework that the following equations hold

$$
\begin{aligned}
Cov(\tilde{y}_{i,t}, \Delta\tilde{y}_{i,t}) &= \frac{1}{1+\rho}\sigma_\varepsilon^2 + \sigma_\nu^2 \\
Cov(\tilde{y}_{i,t}, \Delta\tilde{y}_{i,t-w}) &= \frac{\rho^w}{1+\rho}\sigma_\varepsilon^2 \quad \text{for} \quad w = 1, 2, ... \\
Cov(\tilde{y}_{i,t}, \tilde{y}_{i,t-1}) &= \sigma_\mu^2 + \frac{\rho}{1-\rho^2}\sigma_\varepsilon^2.
\end{aligned}
$$

Rearranging terms yields the formulas in the proposition.

The second part of the proposition follows from standard GMM results and the delta method. In particular, one can show that the moment vector converges towards a normal distribution. This property extends to β since β is a function of the moment vector only via the Delta-method. See e.g. the Appendix in Terry et al. (2021) for additional details on this result. The derivative of β is given by

$$\frac{\partial \beta}{\partial m} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ -\frac{m_3}{(m_2)^2} & 2\frac{m_2}{m_3}+1 & m_2\left(\frac{m_2-2m_3}{(m_2-m_3)^2}\right) & -2\frac{m_2}{m_3} \\ \frac{1}{m_2} & -\left(\frac{m_2}{m_3}\right)^2 & -\left(\frac{m_2}{m_2-m_3}\right)^2 & -\left(\frac{m_2}{m_3}\right)^2 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

□

Results. Table B.1 reports the results for ARDP, calculated either using valuations or alternatively scientific impact measures, and R&D productivity. For all three measures of ARDP I find (1) strong auto-correlation, (2) moderate importance of firm fixed effects, and (3) no significant measurement error. R&D productivity is significantly auto-correlated as well, but have a larger share of variance explained by permanent firm differences. I do not find evidence in favor of measurement error for R&D productivity. Note that the sample sizes are small due to the long-lags necessary to implement the GMM procedure.

Table B.1: GMM results for AR(1) with Noise Component

Parameter	ARDP			Productivity
	Baseline	Citations	Impact	
ρ	0.527***	0.645***	0.542***	0.624***
	(0.086)	(0.084)	(0.083)	(0.120)
σ^2_ε	0.616***	0.493***	0.430***	0.633***
	(0.125)	(0.048)	(0.077)	(0.110)
σ^2_μ	0.093***	0.186	0.175***	0.496***
	(0.031)	(0.135)	(0.054)	(0.124)
σ^2_ν	-0.118	-0.026	0.008	-0.088
	(0.080)	(0.090)	(0.109)	(0.070)
Observations	4,292	4,292	2,594	4,292

Note: Standard errors clustered at the 3-digit industry level and reported in brackets.

B.2 Misspecification

B.3 Estimating Unit Cost Elasticities

Table B.2: Are Firm with Large Avg. R&D Products More Sensitive to User Cost?

	(1)	(2)	(3)	(4)
		$\texttt{R\&D Expenditure}_{t,t+4}$		
R&D User Cost	-2.168*	-1.050**	-1.042**	-1.110**
	(1.201)	(0.414)	(0.441)	(0.431)
R&D User Cost × Long-run R&D Product	-0.469***	-0.070	-0.065	-0.089
	(0.133)	(0.065)	(0.101)	(0.087)
Firm FEs	✓	✓	✓	✓
Size controls		✓	✓	✓
User Cost Link	Inventor	Inventor	Assignee	Head-quarters
Observations	16,920	16,920	16,920	16,920

Note: R&D returns are adjusted for outliers and only reported if firm has at least 25 patents. All regressions control for state and year fixed effects. Standard errors clustered at the user cost state.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

B.4 Tax Credits and Subsidies

Table B.3: ARDP Dispersion With Investment Tax Credits

Measure	Fixed Share			Dynamic Share
	0%	50%	100%	$\frac{\texttt{R\&D}}{\texttt{Investment+R\&D}}$
SD	0.933	0.933	0.934	0.933
S.E. (SD)	(0.006)	(0.006)	(0.006)	(0.006)
Observations	11,804	11,803	11,803	11,803

Note: Standard deviations taken over natural logarithm of Average R&D Product measure. All values residualized by NAICS3× Year fixed effects. Columns (2)-(4) add a fixed fraction of the investment tax credit to R&D expenditure, while column (5) allocates based on the share of R&D in R&D and capital investment.

Table B.4: ARPDs and Government Involvement

	(1)	(2)	(3)	(4)
	ln **ARPD**			
Government involvement	-0.147	-0.092	0.088	0.164
	(0.096)	(0.105)	(0.131)	(0.106)
Fixed Effects	Year	Year	Year	NAICS3 × Year
Exposure Type	Subsidies	Interest	Interest	Interest
Exposure Level	Industry	Industry	Observation	Observation
Within R2	0.003	0.001	0.002	0.007
Observations	12,228	12,228	12,228	12,031

Note: Average R&D Products in logs. Government involvement via subsidies is an indicator for being in an industry with at least 10% of R&D conducted paid for by government. Government involvement via patents at the industry level is an indicator for at least 2% of patents reporting a government interest statement. Government involvement via patents at the observation level is an indicator for at least one patent acknowledging government support.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

B.5 Mechanisms

This section presents additional findings on correlates of ARDP.

Financial Frictions. A channel that has been highlighted in the literature on dynamic resource misallocation has been financial frictions (Whited and Wu, 2006; Midrigan and Xu, 2014). In this context one could imagine that firms are constrained in their current R&D expenditure by cash-flow considerations and limited availability of external finance. In this case, the first-order conditions become

$$\frac{z_{it}\mathbb{E}[\mathcal{V}_{it}]}{W_t \ell_{it}} = (1+\phi)(1+\lambda_{it}^{FF}), \tag{B.4}$$

where λ_{it}^{FF} measures the tightness of financial constraints. This formulation immediately predicts that more constrained firms should have higher returns all else equal. The challenge then arises to measure financial constraints. I will investigate this prediction using a range of alternative predictors of financial frictions. Firstly, I will rely on the two general measures that could be reflective of financial friction, but might also reflect other, unrelated frictions: the average revenue product of capital (APK) and (physical) investment Q. Following the

David and Venkateswaran (2019) and Whited and Wu (2006), I define APK as the ratio of revenue to physical capital stock and investment Q as the ratio of a firm's market valuation to its physical capital stock:

$$\texttt{APK}_{it} \equiv \frac{\sum_{w=0}^{4} \texttt{revt}_{it+w}}{\sum_{w=-1}^{3} \texttt{ppent}_{it+w}} \quad \text{and} \quad \texttt{Q}_{it} = \frac{\sum_{w=0}^{4} \texttt{mkvalt}_{it+w}}{\sum_{w=0}^{4} \texttt{ppegt}_{it+w}}, \tag{B.5}$$

where I define $\texttt{mkvalt}_{it} \equiv \texttt{prcc}_f \times \texttt{csho} + (\texttt{dltt}_{it} + \texttt{dlc}_{it} - \texttt{act}_{it})$.

Secondly, I will use three additional indicators that have been linked more directly to financial frictions: leverage, liquidity, and dividend payments with the idea being that firms with lower leverage, higher liquidity, and positive dividends are less financially constrained (Whited and Wu, 2006; Ottonello and Winberry, 2020). I follow Ottonello and Winberry (2020) in defining (net-)leverage as total debt minus short-run net assets divided by total assets:

$$\texttt{Net-leverage}_{it} \equiv \frac{1}{5}\sum_{w=0}^{4} \frac{(\texttt{dlc}_{it} + \texttt{dltt}_{it}) - (\texttt{act}_{it} - \texttt{lct}_{it})}{\texttt{at}_{it}}. \tag{B.6}$$

Following Whited and Wu (2006), liquidity is measured as cash holdings over assets:

$$\texttt{Liquidity}_{it} \equiv \frac{1}{5}\sum_{w=0}^{4} \frac{\texttt{ch}_{it}}{\texttt{at}_{it}}. \tag{B.7}$$

My dividend measure is a simple dummy whether the firm pays positive dividends.

Finally, I also investigate whether R&D returns differ across the firm lifecycle, which has been linked to financial frictions in Midrigan and Xu (2014).

For each of my indicators I estimate a simple OLS regressions of the form:

$$\texttt{ARDP}_{it} = \alpha_{j(i)\times t} + \beta \texttt{X}_{it} + \varepsilon_{it}, \tag{B.8}$$

where $\texttt{X}_{it}$ is one of the financial friction indicators.

The evidence in Table B.5 suggest that financial frictions are not a strong driver of R&D returns. The only variable that is significant at conventional levels is investment Q. Otherwise, coefficients have the right sign for all but firm age, but are highly imprecise. Furthermore, none of the R^2 exceeds 1% except for Firm Q, which is at 5%.

Table B.5: ARPDs and Financial Frictions

	(1)	(2)	(3)	(4)	(5)	(6)
			`Realized ARPD`$_{it}$			
Financial Friction Indicator	0.028	0.173***	0.839	-1.315	1.338	0.017
	(0.052)	(0.051)	(0.834)	(0.906)	(0.836)	(0.060)
Indicator	ARPK (log)	Firm Q (log)	Net-leverage	Liquidity	Dividend	Firm Age
Within R2	0.000	0.051	0.003	0.001	0.000	0.000
Observations	11,803	10,182	11,193	10,600	11,804	11,804

Note: This table reports coefficient estimates for . See text and Appendix for variable descriptions. Firm Age is an indicator for whether the firm exists for at least 20 year in Compustat. The outcome variable is in logs.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

Valuation-Impact Wedge. Finally, dispersion in ARDPs might not be harmful in the first place if it is driven by companies creating a lot of public per private value. As an economy, we might be better off if firms with large positive externalities of their R&D do too much of it from a private value maximization perspective. To investigate this possibility, I define the impact-valuation wedge as the ratio of patent impact, measured either via citations or using the text-based measure from Kelly et al. (2021), to patent valuation:[15]

$$\texttt{Impact-Valuation Wedge}_{it} = \frac{\sum_{w=0}^{4} \texttt{Patent Impact}_{it+w}}{\sum_{w=0}^{4} \texttt{Patent Valuations}_{it+w}}. \tag{B.9}$$

The idea behind the measure it that it captures inventions that have a high impact, which is not fully captured by the valuation associated with them. A large impact-valuation wedge might be due to spillovers or because firms find it harder to profit from particularly productive inventions.[16]

I test whether the variation in ARDP is associated with the impact-valuation wedge in a simple OLS framework:

$$\texttt{ARDP}_{it} = \alpha + \beta \texttt{Impact-Valuation Wedge}_{it} + \delta X_{it} + \varepsilon_{it} \tag{B.10}$$

[15]For citations, I follow Kogan et al. (2017) in normalizing all citations by the average citations within an application year. I only consider citations within the first five years since the patent was granted. I take the text-based impact measure directly from Kelly et al. (2021). Their data stops in 2009 and I thus lose some observations towards the end of the sample.

[16]The latter naturally arises in Schumpeterian growth model with limit pricing where the profits associated with an invention are concave in the productivity improvement.

Note that within the framework introduced above, a large ARDP implies that a firm does not do enough R&D, while a low ARDP implies too much R&D effort. Thus, a negative β in this context implies that firms with high impact-valuation wedge conduct more R&D than warranted from a private perspective. In other words, a negative β might be a positive sign from a public perspective.

Indeed, this is what I find in column (1) of Table B.6. Both for the citation-based and text-based measures I find strong negative correlations. One concern with this finding might be that patent valuations are in the nominator of the outcome and denominator of the regressor, which might lead to a mechanical correlation, e.g. under measurement error. I address this concern by using the alternative ARDP measures developed above and confirm strong negative correlations for all, but changes in labor revenue productivity changes. Appendix Table B.7 repeats the exercise with firm fixed effects, which confirms the results in columns (1)-(4) and renders column (5) insignificant. Overall, this evidence suggest that at least some of the variation in ARDPs might be optimal.

Table B.6: ARDPs and Impact-Valuation Wedges

	(1)	(2)	(3)	(4)	(5)
A. Citation-based			**Realized ARPD**$_{it}$		
Impact-Valuation Wedge	-0.347***	-0.227***	-0.103***	-0.374***	0.732***
	(0.042)	(0.023)	(0.033)	(0.035)	(0.056)
Within R2	.15	.043	.0061	.12	.22
Observations	11,839	11,682	10,870	6,749	11,580
B. Text-based			**Realized ARPD**$_{it}$		
Impact-Valuation Wedge	-0.458***	-0.258***	-0.096**	-0.472***	0.654***
	(0.040)	(0.024)	(0.043)	(0.041)	(0.084)
Value Measure	Patent valuations	Revenue	Employment	Market valuation	Labor productivity
Within R2	.27	.056	.005	.19	.18
Observations	8,894	8,802	8,109	3,926	8,757

Note: All regressions control for industry× year fixed effects. Standard errors clustered at the 3-digit NAICS level.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

Table B.7: ARDPs and Impact-Valuation Wedges With Firm Fixed Effects

	(1)	(2)	(3)	(4)	(5)
A. Citation-based			**Realized** $\texttt{ARPD}_{it}$		
Impact-Valuation Wedge	-0.560***	-0.470***	-0.434***	-0.848***	-0.048
	(0.051)	(0.047)	(0.043)	(0.079)	(0.069)
Within R2	.2	.088	.044	.18	.00077
Observations	11,800	11,641	10,828	6,710	11,538
B. Text-based			**Realized** $\texttt{ARPD}_{it}$		
Impact-Valuation Wedge	-0.574***	-0.483***	-0.478***	-0.679***	-0.046
	(0.019)	(0.032)	(0.034)	(0.050)	(0.058)
Value Measure	Patent valuations	Revenue	Employment	Market valuation	Labor productivity
Within R2	.25	.11	.054	.15	.00093
Observations	8,860	8,769	8,073	3,892	8,720

Note: All regressions control for industry× year and firm fixed effects. Standard errors clustered at the 3-digit NAICS level.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

C Theoretical Results

C.1 Proof from Main Text

Proof of Proposition 1. Firstly, we can rearrange the growth rate expression as

$$
\begin{aligned}
g &= \int_0^1 \frac{z_{it} \times \mathbb{E}_t[\lambda_{it} - 1]}{z_{it}\mathbb{E}_t[\mathcal{V}_{it}]} \times \frac{z_{it}\mathbb{E}_t[\mathcal{V}_{it}]}{W_t \ell_{it}} \times W_t \ell_{it} di \\
&= \int_0^1 \zeta \times \frac{1+\phi}{\xi_{it}} \times W_t \ell_{it} di.
\end{aligned}
$$

Secondly, the first order conditions of R&D labor require

$$
\ell_{it} = \left(\frac{\gamma_{it}\xi_{it}}{(1+\phi)W_t} \right)^{\frac{1+\phi}{\phi}}
$$

Plugging labor demand into the aggregate labor supply, we can solve for W_t:

$$
W_t = \mathcal{L}^{-\frac{\phi}{1+\phi}} \times \frac{1}{1+\phi} \times \left(\int_0^1 (\gamma_{it}\xi_{it})^{\frac{1+\phi}{\phi}} di \right)^{\frac{\phi}{1+\phi}}.
$$

This expression also implies that equilibrium labor allocation is given by

$$
\ell_{it} = \mathcal{L} \times \frac{(\gamma_{it}\xi_{it})^{\frac{1+\phi}{\phi}}}{\int_0^1 (\gamma_{it}\xi_{it})^{\frac{1+\phi}{\phi}}}.
$$

Plugging in the expression for labor allocation and wage into the growth rate formula then yields the main expression. Note that the weights ω_{it} are the labor allocation share of a firm in absence of dispersion in ξ_{it}.

□

Proof of Corollary 1. It is straight-forward to show that $\Xi = 1$ if ξ_{it} is the same across all firms. Furthermore, one can show that the denominator is a convex transformation of the nominator, such that it is always larger in the presence of dispersion. □

Proof of Proposition 2. The proof proceeds along the same algebraic steps as the proof for Proposition 1. □

Proof of Corollary 3. The proof proceeds along the same algebraic steps as the proof for Corollary 1. □

C.2 Extensions

Entry. Entry of new firms has been an important force in Schumpeterian growth theory (Aghion et al., 2014). In the following I show how my result can be extended to the case of a flexible mass M_t of firms. While the formulas to not change significantly, the counterfactual now depends on how the mass of active firms responds to ARDP dispersion. As long as dispersion acts as a detractor to entry, the formula for Ξ developed in the main text is a lower bound for the gains from reducing ARDP dispersion. We could suspect this to be the case, for example, because R&D wages are increasing in ARDP dispersion, which increasing the cost of R&D and thus might lower the net-gain.

Proposition 4. *Consider the economy above, but with a mass M_t of innovative firms. We can express the economic growth rate in the model described above as*

$$g = \tilde{g}^{\max} \times \tilde{\Xi}, \tag{C.1}$$

where $\tilde{g}^{\max}$ is the growth-rate in absence of social ARDP dispersion via ξ_{it}/ζ_{it}, while holding M_t fixed, and allocative efficiency $\tilde{\Xi}$ is given by

$$\tilde{\Xi}^{\phi} \equiv \frac{\left(\int_0^{M_t} \tilde{\omega}_{it} \times (\xi_{it}/\zeta_{it})^{\frac{1}{\phi}} di\right)^{\phi}}{\left(\int_0^{M_t} \tilde{\omega}_{it} \times (\xi_{it}/\zeta_{it})^{\frac{1+\phi}{\phi}} di\right)^{\frac{\phi}{1+\phi}}}, \quad \text{with} \quad \tilde{\omega}_{it} \equiv \frac{\tilde{\gamma}_{it}^{\frac{1+\phi}{\phi}}}{\int_0^{M_t} \tilde{\gamma}_{it}^{\frac{1+\phi}{\phi}} di}. \tag{C.2}$$

The formula for $\tilde{g}^{\max}$ is given by

$$\tilde{g}^{\max} = \mathcal{L}^{\frac{1}{1+\phi}} \times M_t^{\frac{\phi}{1+\phi}} \times \left(\int_0^1 \tilde{\gamma}_{it}^{\frac{1+\phi}{\phi}} di\right)^{\frac{\phi}{1+\phi}}. \tag{C.3}$$

Proof of Proposition 4. The proof follows the same algebraic steps as the proof for Proposition 1. ☐

C.3 Welfare and Consumption Equivalents

In this section I briefly lay out how I calculate consumption equivalent measures. For this I will consider both a model with logarithmic and CRRA preferences over consumption.

Consider a representative household discounting the future at rate ρ with CRRA preferences over consumption stream C_t. Welfare is given by

$$\mathcal{W} = \int_0 \infty e^{-\rho t} u(C_t) dt \quad \text{where} \quad u(C) = \begin{cases} \ln C & \text{if} \quad \sigma = 1 \\ \frac{C_t^{1-\sigma}-1}{1-\sigma} & \text{if} \quad \sigma \in [0, \infty) \setminus \{1\} \end{cases} \tag{C.4}$$

Assuming that the consumption grows at constant rate g, we can solve for welfare as

$$\mathcal{W} = \frac{1}{\rho + (\sigma-1)g} \left(u(C) + \frac{g}{\rho} \right). \tag{C.5}$$

For a given growth rate increase from g to $(1+\Delta_g)g$, the consumption equivalent is the permanent increase in consumption from C to $C(1+\Delta_C)$ such that welfare is equivalent in both scenarios. Δ_C thus solves

$$\frac{1}{\rho + (\sigma-1)g} \left(u(C(1+\Delta_C)) + \frac{g}{\rho} \right) = \frac{1}{\rho + (\sigma-1)g(1+\Delta_g)} \left(u(C) + \frac{g(1+\Delta_g)}{\rho} \right) \tag{C.6}$$

We can solve this explicitly as

$$\Delta_C = \begin{cases} \exp\left(\frac{g}{\rho} \times \Delta_g \right) - 1 & \text{if} \quad \sigma = 1 \\ \left(1 + \Delta_g \frac{(\sigma-1)g}{\rho+(\sigma-1)g} \right)^{\frac{1}{\sigma-1}} - 1 & \text{otherwise} \end{cases}. \tag{C.7}$$

D A Quantitative Schumpeterian Growth Model

This section provides derivations for the quantitative model introduced in Section 5.

D.1 General setup

Time is discrete and indexed by t.

Household. The representative household has log preferences over consumption C_t and a discount factor of β. For a given consumption sequence $\{C_t\}_{t=0}^{\infty}$, utility is given by

$$U\left(\{C_t\}_{t=0}^{\infty}\right) = \sum_{t=0}^{\infty} \beta^t \log C_t. \tag{D.1}$$

The household owns all the assets in the economy and earns income from wages and firm profits. The household has access to a riskless bond earning gross interest rate R_{t+1}. The first order conditions for the bond give rise to the Euler equation:

$$\frac{C_{t+1}}{C_t} = \beta R_{t+1}. \tag{D.2}$$

The consumption good itself is Cobb-Douglas composite over a unit mass of product lines:

$$\ln C_t = \int_0^1 \ln C_{it} di. \tag{D.3}$$

The household takes prices as given when making consumption decisions. First order conditions for optimal consumption allocation require

$$C_t = P_{it} C_{it}, \tag{D.4}$$

where I normalize the price of the aggregate consumption good to 1 and denote the price of individual goods by P_{it}.

Static production. Each product i can be produced by a set $\mathbb{F}_{it}$ of firms with linear production function in labor L_{ijt}:

$$Y_{ijt} = A_{ijt} L_{ijt} \quad \text{for} \quad j \in \mathbb{F}_{it}. \tag{D.5}$$

I will assume that there are no firms with the same productivity at any point in time. The microfoundation of productivity improvement introduced below will ensure that this will always hold in equilibrium. Market clearing requires

$$C_{it} = \sum_{j \in \mathbb{F}_{it}} Y_{ijt}. \tag{D.6}$$

Firms operate in Bertrand competition and hire labor at wage W_t^P. I will denote the most productive technology by $A_{it} =\equiv \max_{j \in \mathbb{F}_{it}}\{A_{ijt}\}$ and the second best technology by $\underline{A}_{it} \equiv \max_{j \in \mathbb{F}_{it}}\{A_{ijt}\} \setminus \{A_{it}\}$. The nature of the competition together with the demand structure imply that there is a unique Bertrand equilibrium where the first-best firm takes over the entire market and sets the price equal to the unit cost of the second best firm such that

$$P_{it} = \frac{W_t^P}{\underline{A}_{it}}. \tag{D.7}$$

The intuition is simple: At this price only the first and second best firm can supply the good without losses. Due to constant unit cost, both are willing to supply any quantity. Furthermore, the second best firm is exactly indifferent between supplying the good or not. Finally, positive supply by the second best firm cannot be an equilibrium as the first-best firm could decrease its price by a small amount, take over the entire market and make larger profits.

Combing demand, equilibrium price and production function, we can calculate the equilibrium profits by the best firm, which I will call incumbent from now on, as

$$\pi_t(\lambda_{it}) = (1 - \lambda_{it}^{-1})C_t \qquad \text{where} \quad \lambda_{it} \equiv \frac{A_{it}}{\underline{A}_{it}}. \tag{D.8}$$

The productivity advantage λ_{it} is thus a sufficient statistic for firm profits conditional on aggregate production C_t.

Aggregate output. Having solved the static firm problem, we can calculate aggregates. I will assume that there is a fixed mass of production workers L. Using the aggregate production function, optimal firm prices, and the labor market clearing for production workers we can then express aggregate output as the product of current technology A_t, production labor, and a markup dispersion term Λ_t:

$$C_t = A_t L \Lambda_t, \quad \text{where} \quad \ln A_t \equiv \int_0^1 \ln A_{it} di \quad \text{and} \quad \Lambda_t \equiv \frac{\exp\left(\int_0^1 \ln(1/\lambda_{it}) di\right)}{\int_0^1 (1/\lambda_{it}) di}. \tag{D.9}$$

The markup dispersion term is decreasing in the dispersion in λ_{it} and captures static misallocation due to heterogeneous markups (Acemoglu and Autor, 2011; Peters, 2020; de Ridder, 2021). Note, however, that this term is second order from a welfare perspective in practice (Acemoglu and Autor, 2011; Peters, 2020).

Innovation. There is a unit mass of firms with productivity portfolios $\{A_{ijt}\}_{i\in[0,1]}$ across product lines. As discussed above, the factor determining whether a firm is making profits is its incumbent status and the associated productivity gap. I will, thus, instead denote the relevant state-variable from the firm perspective by $\Lambda_{jt} \equiv \{\lambda_{it}\}_{i\in\mathbb{I}_{jt}}$, where $\mathbb{I}_{jt}$ is the set of product lines where firm j is the incumbent at time t. Note that this set could be empty.

Firms have R&D efficiency φ_{it} and hire researchers ℓ_{jt} to innovate with probability z_{jt}

$$z_{jt} = \varphi_{jt} \ell_{jt}^{\frac{1}{1+\phi}}. \tag{D.10}$$

An innovation conducted at time t is realized in $t+1$. Inventions improve upon the leading technology in a random product lines by factor λ_{jt}. Successful invention thus adds $\{\lambda_{jt}\}$ to Λ_{jt}. On the other hand, innovation by rival firms displaces incumbents at the aggregate innovation probability $z_t \equiv \int_0^1 z_{jt} di$ for each incumbent product line.

To add some ex-ante uncertainty around the valuation of R&D output, I assume that $\lambda_{jt} = \lambda^{\Delta_{jt}}$, where Δ_{jt} is an i.i.d. random variable when the innovation investments are made and distributed according to

$$P(\Delta) = \frac{(1-p)^{\Delta-1} p}{1-(1-p)^{\bar{\Delta}}} \qquad \text{for} \quad \Delta = 1, ..., \bar{\Delta}. \tag{D.11}$$

Note that this formulation implies that there is a time-invariant distribution over λ such that Λ_t becomes a constant.

Firm R&D productivity evolves according to a simple AR(1) process:

$$\varphi_{it} = (1-\rho)\mu + \rho\varphi_{it} + \varepsilon_{it} \qquad \varepsilon_{it} \overset{i.i.d.}{\sim} N(0,\sigma). \tag{D.12}$$

The firm is subject to two frictions in its R&D labor input choice: Adjustment cost and labor market power. The cost of R&D are given by

$$C(\ell_{jt}, \ell_{jt-1}) = W_t \ell_{jt} \left(1 + \xi^{Adj.} \left(\frac{\ell_{jt} - (1-\delta)\ell_{jt-1}}{\ell_{jt}}\right)^2\right) \left(1 + \frac{\ell_{jt}}{\underline{\ell}}\right)^{\xi^{LMP}}. \tag{D.13}$$

Here, $\xi^{Adj.}$ regulates the adjustment cost, while ξ^{LMP} determines the degree of labor market power. The wage factor W_t determines a baseline wage level. Importantly, the elasticity of the final term with respect to R&D workers is increasing in them such that firms with high R&D labor demand effectively face inelastic supply. This formulation allows me to match the patterns documented in Seegmiller (2021) qualitatively. The depreciation term δ captures the natural turnover in the labor market and is useful for quantitative purposes. Adjustment cost for R&D workers can be motivated by congestion externalities and more generally with decreasing returns to scale in the hiring process.

The firms then solve a simple value function maximization problem:

$$V_t(\Lambda_{jt}, \ell_{jt-1}, \varphi_{it}) = \max_{\ell_{jt}} \left\{ \sum_{\lambda \in \Lambda_{jt}} \pi_t(\lambda) - C(\ell_{jt}, \ell_{jt-1}) + \frac{1}{R_{t+1}} \mathbb{E}_t \left[V_{t+1}(\Lambda_{jt+1}, \ell_{jt}, \varphi_{it+1}) \right] \right\},$$

where expectations are taken with respect to the evolution of φ_{it} as well as Λ_{jt}.

It is straight-forward to show via a guess and verify approach that the value function for this problem can be simplified as

$$V_t(\Lambda_{jt}, \ell_{jt-1}, \varphi_{it}) = \tilde{V}_t(\ell_{jt-1}, \varphi_{it}) + \sum_{\lambda \in \Lambda_{jt}} \mathcal{V}_t(\lambda).$$

The value function is composed of a profit part and a R&D part. The profit part simply values the expected current and future profits from an innovation:

$$\mathcal{V}_t(\lambda) = \pi_t(\lambda) + \left(\frac{1 - z_t}{R_{t+1}}\right) \mathcal{V}_{t+1}(\lambda). \tag{D.14}$$

On the other hand, the R&D part captures the value of profitable conducting R&D and generating new inventions:

$$\tilde{V}_t(\ell_{jt-1}, \varphi_{it}) = \max_{\ell_{jt}} \left\{ -C(\ell_{jt}, \ell_{jt-1}) + \frac{1}{R_{t+1}} \left(z_{jt} \mathbb{E}_t[\mathcal{V}_{t+1}(\lambda_{jt})] + \mathbb{E}_t[\tilde{V}_t(\ell_{jt}, \varphi_{it+1})] \right) \right\}. \tag{D.15}$$

This formulation is much more tractable and computationally manageable. Furthermore, it clarifies that the relevant state for the firm R&D choice is $(\ell_{it-1}, \varphi_{it})$. The associated distribution evolves according to

$$f_{t+1}(\ell', \varphi') = \int \int f_t(\ell, \varphi)\mathbb{I}\{\ell' = \ell_t^*(\ell, \varphi)\}d\ell f(\varphi'|\varphi)di, \tag{D.16}$$

where $f(\varphi'|\varphi)$ is the conditional density of being a φ' firm conditional on being a φ firm today, which is time-invariant due to the law of motion for φ_{it}. Furthermore, $\ell_t^*(\ell, \varphi)\}$ is the time t policy function evaluated at R&D workers ℓ and productivity level φ.

Inventor market clearing. Finally, I close the model by assuming that there is a fixed mass of inventors $\mathcal{L}$ such that

$$\mathcal{L} = \int_0^1 \ell_{jt}dj. \tag{D.17}$$

Economic growth. The innovation process implies that productivity evolves according to the simple law of motion

$$g \equiv \ln A_{t+1} - \ln A_t = \int_0^1 z_{jt} \ln \lambda_{jt}dj. \tag{D.18}$$

Definition 2. *A competitive equilibrium is a sequence of prices $\{W_t, R_t\}$, aggregates $\{C_t, A_t\}$, firm-level variables $\{\varphi_{it}, \ell_{it}\}$, value functions $\tilde{V}_t(\cdot), \mathcal{V}(\cdot)$ and a joint distributions $f_t(\ell_{it-1}, \varphi_{it})$ such that the value functions solve the firm optimization problem with policy function ℓ_{jt}, consumption and interest rate respect the Euler equation, productivity evolves according to its law of motion, and the distributions evolve according to their law of motion.*

D.2 Balanced growth path

Definition 3. *A Balanced Growth Path (BGP) is a competitive equilibrium such that aggregates and wages grow at the same, constant rate.*

By definition of a Balanced growth path, we have that aggregate consumption grows by factor $1 + g$, which implies that we have a constant interest rate R via the Euler equation. I will normalize denote values normalized by C_t in lower case or by dropping time indices for the value functions.

It is straight-forward to verify that the normalized profit value is simply given by

$$\mathcal{V}(\lambda) = \frac{\pi(\lambda)}{1 - \beta(1 - z)}, \tag{D.19}$$

where I have used the Euler equation to substitute out the interest rate. Note that z is a constant since g is a constant and we can an iid distribution over innovation step sizes λ:

$$g = \int_0^1 z_{ij} \ln \lambda_{jt} dij = \int_0^1 z_{jt} dj \times \mathbb{E}[\ln \lambda] = z \times \mathbb{E}[\ln \lambda].$$

We can similarly normalize the value function for R&D such that

$$\tilde{V}(\ell_{jt-1}, \varphi_{it}) = \max_{\ell_{jt}} \left\{ -c(\ell_{jt}, \ell_{jt-1}) + \beta \left(z_{jt} \mathbb{E}[\mathcal{V}(\lambda)] + \mathbb{E}[\tilde{V}(\ell_{jt}, \varphi_{it+1})] \right) \right\}, \tag{D.20}$$

where the cost function is given by

$$c(\ell_{jt}, \ell_{jt-1}) = w\ell_{jt} \left(1 + \xi^{Adj.} \left(\frac{\ell_{jt} - (1-\delta)\ell_{jt-1}}{\ell_{jt}} \right)^2 \right) \left(1 + \frac{\ell_{jt}}{\underline{\ell}} \right)^{\xi^{LMP}}. \tag{D.21}$$

Expressing the value function in recursive form, we have

$$\tilde{V}(\ell, \varphi) = \max_{\ell'} \left\{ -c(\ell', \ell) + \beta \left(z(\ell', \varphi) \frac{\mathbb{E}[(1-\lambda^{-1})]}{1 - \beta(1-z)} + \mathbb{E}[\tilde{V}(\ell', \varphi')] \right) \right\}, \tag{D.22}$$

where substituted in the value of $\mathcal{V}(\lambda)$. I will denote the associated $\arg\max$ by $\ell^*(\ell, \varphi)$. The ergodic distribution over the state-space is then the solution to

$$f(\ell', \varphi') = \int \int f(\ell, \varphi) \mathbb{I}\{\ell' = \ell^*(\ell, \varphi)\} d\ell f(\varphi' | \varphi) di. \tag{D.23}$$

Furthermore, labor market clearing requires

$$\mathcal{L} = \int_0^1 \ell^*(\ell, \varphi) dF(\ell, \varphi). \tag{D.24}$$

Proposition 5. *A value function $\tilde{V}(\ell, \varphi)$ with policy function $\ell^*(\ell, \varphi)$, distribution $f(\ell', \varphi')$ and normalized wage w that solve equations* (D.22)*,* (D.23)*, and* (D.24) *subject to the productivity law of motion constitute a Balanced Growth Path, where the aggregate innovation rate and growth rate are then simply given by*

$$z = \int z(\ell^*(\ell, \varphi), \varphi) dF(\ell, \varphi) \quad \text{and} \quad g = z \times \mathbb{E}[\ln \lambda]. \tag{D.25}$$

E Quantitative Results

E.1 Additional Robustness

Table E.1: Specification Robustness for Aggregate Results

Specification	Baseline			Unweighted		
	Private	Public		Private	Public	
	Adjusted	Citations	Impact	Adjusted	Citations	Impact
A. Allocative Efficiency Ξ						
$\phi = 1/2$	59.2%	47.7%	53.0%	43.2%	43.8%	48.1%
$\phi = 1$	76.1%	68.6%	72.8%	60.6%	59.9%	67.1%
$\phi = 2$	86.7%	82.6%	85.4%	77.7%	76.6%	82.0%
B. Growth Cost						
$\phi = 1/2$	1.03 p.p.	1.65 p.p.	1.33 p.p.	1.97 p.p.	1.93 p.p.	1.62 p.p.
$\phi = 1$	0.47 p.p.	0.69 p.p.	0.56 p.p.	0.97 p.p.	1.00 p.p.	0.74 p.p.
$\phi = 2$	0.23 p.p.	0.32 p.p.	0.26 p.p.	0.43 p.p.	0.46 p.p.	0.33 p.p.
C. Consumption Equivalent						
$\phi = 1/2$	180.9%	419.3%	278.8%	617.3%	587.5%	404.3%
$\phi = 1$	60.3%	99.0%	75.0%	164.9%	172.9%	108.6%
$\phi = 2$	25.8%	37.3%	29.1%	53.9%	58.1%	39.0%
D. Consumption Equivalent with $\sigma = 2$						
$\phi = 1/2$	41.3%	65.9%	53.3%	78.8%	77.1%	64.7%
$\phi = 1$	18.9%	27.5%	22.4%	39.0%	40.2%	29.4%
$\phi = 2$	9.2%	12.7%	10.2%	17.3%	18.3%	13.2%

Note: Baseline results uses the standard formula, while unweighted sets ω_{it} to a uniform weight over active firms. "Growth Cost" report $g^{\max} - g$ assuming a baseline growth rate g of 1.5% per annum. "Consumption Equilvalent" translates this number into a consumption equivalent value for welfare on a Balanced Growth Path with logarithmic preferences over consumption using the formula $\Delta_C = \exp\left(\frac{g}{\rho}\Delta_g\right) - 1$. "Consumption Equivalent with $\sigma = 2$" instead uses a CRRA utility with intertemporal elasticity of substitution σ giving rise to consumption equivalent $\Delta_C = \left(1 + \Delta_g \frac{(\sigma-1)g}{\rho+(\sigma-1)g}\right)^{\frac{1}{\sigma-1}} - 1$. See Appendix C.3 for details on the welfare calculations.

Table E.2: Contribution of Declining Allocative Efficiency to Growth Slowdown — Robustness

Specification	Private	Public	
	Adjusted	Citations	Impact
A. Allocative Efficiency 1974-94			
$\phi = 1/2$	63.3%	53.6%	55.6%
$\phi = 1$	79.9%	73.4%	74.9%
$\phi = 2$	89.4%	85.7%	86.8%
B. Allocative Efficiency 2004-14			
$\phi = 1/2$	53.9%	39.7%	40.4%
$\phi = 1$	71.8%	61.5%	64.5%
$\phi = 2$	84.0%	78.2%	80.1%
C. Contribution to Declining Growth			
$\phi = 1/2$	81.7%	164.9%	177.6%
$\phi = 1$	52.8%	90.7%	75.5%
$\phi = 2$	30.3%	45.2%	39.4%

Note: This tables reports the contribution of changing allocative efficiency to the growth slowdown. The values in Panel C are calculated as $\frac{\Xi_{74-94}/\Xi_{04-14}-1}{1.54/1.27-1}$.