The Diffusion of New General Purpose Technologies

Stephen Ayerst*

September 16, 2023

Abstract

What firms drive the diffusion of new general purpose technologies (GPT) and how does this diffusion affect productivity growth? I develop a model of innovative firms that adopt and diffuse information-communication technologies (ICT), a recent GPT. The central mechanism is that research productivity depends on ICT diffusion, leading to ICT firms investing more as ICT is diffused. I construct a measure of ICT-related patents capturing the application of ICT to new products. Empirically, firm-level citations-per-patent, patenting frequency, and productivity growth increase following ICT adoption. Quantifying the model shows that around two-thirds of ICT diffusion is driven by ICT firms, rather than new adopters, and growth falls by as much as 28% following the introduction of ICT, despite growth eventually recovering.

Keywords: Innovation, Endogenous Growth, Productivity, Information-Communication Technology (ICT), Technology Diffusion.
JEL classification: O11, O14, O33, O41, O43.

*International Monetary Fund, 700 19th Street NW, Washington DC, 20431. Email: stephen.b.ayerst@gmail.com. A previous version of the paper circulated as "The Adoption and Diffusion of New Technologies." I thank Diego Restuccia, Xiaodong Zhu and Murat Celik for their advice and guidance. This paper has benefited from comments from Ufuk Akcigit, Michelle Alexopoulos, Ariel Burstein, April Franco, Tomy Lee, Gaelan MacKenzie, Virgiliu Midrigan, Matt Mitchell, Swapnika Rachapalli, Baxter Robinson, and Marc-Antoine Schmidt as well as seminar participants at York University, Laurier University, the Bank of Canada, the Macro Brown Bag and Public Discussion Group at the University of Toronto, the 2018 CEA Meetings, the 2018 Asian Meeting of the Econometric Society, the 2019 North American Meeting of the Econometric Society, and the 2020 Econometric Society World Congress. The views expressed herein are those of the author and should not be attributed to the IMF, its Executive Board, or its management.

1 Introduction

General purpose technologies (GPTs), such as electricity or computers, transform the economy and fuel economic growth through their scope for improvement and broad applications throughout the economy (Bresnahan and Trajtenberg, 1995). Despite their importance, the diffusion of new GPTs often lasts several decades and coincides with prolonged growth slowdowns (Jovanovic and Rousseau, 2005). This leads to two questions: (1) how are new GPTs diffused through the economy? and (2) how does the diffusion of new GPTs affect growth?

To answer these questions, I develop a quantitative model of GPT diffusion by innovative firms and apply this model to the case of information-communication technologies (ICT). While being widely studied, the consequences of ICT diffusion remain unclear and debated.[1] The existing literature primarily focuses on equipment-based measures of ICT diffusion. In contrast, I focus on a novel dimension of ICT as a method of invention, capturing that part of ICT's value is that is allows for innovations that would be otherwise impossible. Consider, as an example, a vehicle's electronic control unit, which is an application of ICT to the automotive industry. The innovation builds directly on ICT and allows for future improvements to vehicles that would be otherwise impossible. However, equipment-based measures would only capture the value of the ICT equipment used in the vehicle, understating the innovations value as an input to future innovations. I develop a patent-based measure of ICT innovations that captures applications to new innovations and use this measure, along with the quantitative model, to answer the initial questions. First, slow ICT diffusion was due to high adoption costs and the difficulty expanding ICT into new markets through innovations. That is, the electronic control unit, and similar innovations, are both costly ideas to invent and difficult to leverage in new markets. Second, ICT's introduction contributed to a prolonged growth slowdown, with growth falling by as much as 28%, despite increased long-run growth potential.

I develop an endogenous growth model that bridges the innovation and technology adoption literatures. The model extends Klette and Kortum (2004) and Acemoglu et al. (2016a) to allow for adoption R&D as well as competing innovations between ICT and non-ICT firms. Firms invest in innovation and adoption R&D. On the innovation side, firms invest to create new, higher quality varieties of existing goods allowing the innovating firm to displace the good's incumbent producer. On the adoption side, firms invest to apply ICT to their goods and innovation process. In this regard, adoption captures the automotive firm learning to apply ICT to their cars (inventing the electronic control unit) rather than buying computer

[1] See, for example, Solow (1987); Brynjolfsson and Hitt (2000); Gordon (2000); Oliner and Sichel (2000); Acemoglu et al. (2014).

equipment for their offices.

ICT affects innovation dynamics through three channels. First, ICT has a direct value to consumers allowing adopters to charge higher markups. This could capture, for example, the novelty or extra value added relative to non-ICT goods. Second, ICT and non-ICT firms differ in the relative costs and returns to innovation R&D. Third, knowledge and competition spillovers differ between ICT and non-ICT firms. ICT (non-ICT) firms build more on the products of (knowledge spillover) and are more likely to be displaced by (competition spillover) other ICT (non-ICT) firms. This captures the inherent difficulty of applying new technologies to goods. The consequence of ICT-dependent spillovers is that ICT innovations embodied in goods are relatively more complementary to future ICT innovations leading ICT firms to invest more in R&D when ICT is more common, all else equal.

The model outcomes are ICT diffusion, which is measured as the ICT product share, and the aggregate productivity growth rate. ICT diffusion depends on both the adoption of ICT by non-ICT firms (*adoption channel*), such as the initial invention of the electronic control unit, and the expansion of ICT firms into new product lines (*expansion channel*), such as expanding the electronic control unit into adjacent markets (e.g., motorbikes). Growth depends on the relative research productivity of and allocation of resources across different types of R&D (e.g., adoption by non-ICT firms, innovation by ICT firms). Consequently, growth is tightly linked to ICT diffusion which affects both research productivity, through knowledge spillovers, and the R&D allocation, through incentives to innovate.

I analyze ICT diffusion by innovative US firms using administrative micro-level data on US patents. The main empirical challenge is to identify patents that apply ICT knowledge but are not in the ICT product market as these patents tend not to be ICT innovations under traditional classifications (e.g., USPTO technology classes). I address this challenge by using patent citation networks to construct a set of ICT-related patents that captures the application of ICT to new innovations, a key feature of general purpose technologies (Griliches, 1957; Jovanovic and Rousseau, 2005).

The ICT-related patent share increases from 21% in 1980 to 70% in 2000 compared with an increase from 15% to 25% using direct measures of ICT use. I use this measure to shed light on differences between ICT and non-ICT firm research productivity and the returns to ICT adoption in the model. This also serves to validate the measure of ICT-related patents by showings that firms characteristics change around the time of adoption. First, ICT firms produce higher quality patents and patent more frequently after adoption. Patent quality of adopting firms, measured by citations, increases by 12% relative to the firm's previous patents, controlling for trends in technology class and year. Adopting firms also patent, on average, 23% more than before adoption, controlling for firm heterogeneity and time trends.

Second, using linked Compustat data, firm productivity growth increases to around double its initial value in the adoption year and proceeding two years. The results are robust to alternative outcomes (e.g., employment growth), alternative controls (e.g., firm-level fixed effects) and do not hold for placebo measures of adoption based on other technology classes (e.g., chemical patents).

I calibrate the model's transition path to match ICT diffusion in the data. The empirical diffusion curve helps discipline the relative cost of adoption R&D. Empirical differences between ICT and non-ICT firm innovation characteristics (i.e., frequency, quality) and firm productivity growth around the time of adoption discipline the relative research productivity of ICT firms and the relative value of ICT products to consumers. The relative frequency of ICT patents citing non-ICT patents disciplines the strength of non-ICT to ICT knowledge spillovers. The calibrated model replicates the share of ICT diffusion driven by new adopters and the ICT product share, which are not targeted in the calibration.

I use the quantitative model to examine the drivers of ICT diffusion and the consequences for productivity growth. The calibration only directly targets the growth rate on the pre-ICT balanced growth path equilibrium. Long-run and transition-path growth are determined by the model using the relative innovative characteristics of ICT and non-ICT firms as well as the empirical path of ICT diffusion.[2] In the long-run, productivity growth increases by 37% of its initial value because of higher R&D spending by ICT firms. However, the introduction of ICT leads to a prolonged growth slowdown, in which growth falls by as much as 28% of its initial value, and growth does not recover to its initial value until around 2025. The growth slowdown is much longer lived than other mechanisms (e.g., Hornstein and Krusell, 1996; Yorukoglu, 1998) implying that ICT diffusion across innovations is more related to long-term secular decline in aggregate productivity growth.[3] The growth slowdown is driven by a decline in non-ICT innovation as ICT becomes more prevalent but is unable to fully accommodate the decline in non-ICT innovation.

ICT diffusion is primarily driven by the expansion channel in which ICT firms expand into new product lines, accounting for around two-thirds of cumulative diffusion. The expansion channel becomes stronger as ICT diffuses and ICT firms increase innovation. The adoption channel accounts for the remaining one-third of cumulative diffusion. The adoption channel is strongest early in the transition period when the returns to adopting are highest due to

[2] The calibration does not directly target the decline in research productivity (Bloom et al., 2020) or aggregate productivity growth over this period. Additionally, the innovative characteristics of ICT and non-ICT firms are set to target within sector and year moments such that these moments do not capture these trends. In this regard, the results should be thought of as isolating the impact of ICT adoption and diffusion on innovations without assuming a stance on how ICT adoption and diffusion applies to these trends.

[3] The longer growth slowdown is driven by the slower diffusion of ICT innovations in the data (see Section 3). For example, the share of ICT-related products reaches only around 50% by 2000.

ICT firms facing less competition.

Finally, I use the quantitative model to examine subsidies and taxes to non-ICT firm adoption and ICT firm innovation to provide insight for the role of policy and to highlight the model mechanisms. Subsidizing adoption R&D has a relatively small effect on ICT diffusion, growth, and welfare over the transition path. This is due to a negative feedback channel in which the initial increase in ICT diffusion lowers future benefits of adoption due to increased competition. Additionally, if the policymaker's time horizon is short enough, they find it beneficial to tax adoption R&D to flatten the drop in productivity growth. In contrast to the adoption R&D subsidy, subsidizing R&D by ICT firms encourages diffusion throughout the transition period and has a larger impact as ICT becomes more prevalent. This is driven by a reinforcing mechanism in which higher ICT innovation increases knowledge spillovers making future ICT R&D more productive. Taken together, the results suggest that policy should be designed over the entire transition path and should focus on users of new technologies, rather than non-users.

Related literature. My paper relates to the endogenous growth literature (Romer, 1990; Aghion and Howitt, 1992; Grossman and Helpman, 1991). Following Klette and Kortum (2004), the literature focuses on connecting endogenous growth models with micro-level data to identify key parameters (Lentz and Mortensen, 2008; Acemoglu et al., 2015; Garcia-Macia et al., 2016). Acemoglu et al. (2016a) extend Klette and Kortum (2004) to examine innovative interactions between environmentally clean and dirty innovators. I adapt this framework to examine interactions between non-ICT and ICT firms and extend the model to incorporate adoption R&D, allowing for firms to switch types. This allows me to examine endogenous ICT diffusion through both new adopters (through adoption R&D) and previous adopters (through innovation R&D). Others examine a closely related channel in which growth is driven by firms imitating the production techniques of frontier firms (e.g., Jovanovic and MacDonald, 1994; Perla and Tonetti, 2014; Lucas and Moll, 2014; Benhabib et al., 2018). ICT adoption similarly allows firms to improve productivity but differs by affecting the innovation dynamics of adopting firms through the effects on research productivity and knowledge and competition spillovers.

My paper also relates to the literature studying the determinants of technology adoption (Caselli and Coleman, 2001; Comin and Hobijn, 2004, 2010; Comin and Mestieri, 2018; Chen, 2018; Ayerst, 2020a). Due to data limitations, this literature generally focuses on either firm-level surveys that cover a small subset of the economy or on aggregate measures of technology stock (e.g., number of computers). In contrast, I develop a novel measure of ICT diffusion, unrelated to ICT equipment, that captures the application of ICT to new innovations using administrative micro-data on US patents and citation networks. The benefit of this measure is that it allows me to capture the entirety of ICT diffusion while maintaining the micro-level nature of the data, allowing me to examine firm-level innovation and market outcomes.

Finally, my paper relates to the literature examining the consequences of ICT on growth (e.g., Brynjolfsson and Hitt, 2000; Oliner and Sichel, 2000; Gordon, 2000).[4] Hornstein and Krusell (1996), Greenwood and Yorukoglu (1997), and Yorukoglu (1998) find that the growth slowdown following the introduction of ICT is driven by workers and firms slowly learning how to use new ICT equipment.[5] I provide an alternative explanation that is driven by declining investment by non-ICT firms. This channel leads to a longer-lived growth slowdown because of slow ICT diffusion across innovations in the data, and ties the paper more to the longer term slowdown in aggregate productivity growth. In a closely related paper, Baslandze (2018) uses patent data to examine the consequences of ICT diffusion on sectoral knowledge spillovers and innovation. In contrast, I focus on diffusion as an endogenous outcome depending on firm-level actions. While I similarly find that ICT improves long-run growth, this only occurs after a costly transition period in which productivity growth slows.

In closely related papers, De Ridder (2020) and Aghion et al. (2021) explain the secular decline in productivity growth as the consequence of the rise in intangible capital associated with the digitalization of the economy. Mechanically, intangible capital provides a competitive advantage to some firms, leading disadvantaged firms to reduce R&D investment. Using different data, I find a similar pattern in which productivity growth slows because non-ICT firms (the disadvantaged type) invest less in R&D as ICT firms (the advantaged type) become more common. My departure is to allow for endogenous firm types, which implies that the growth slowdown is temporary since all firms eventually become ICT firms, and use data to discipline the introduction of ICT. Despite the eventual recovery, I find that the growth rate does not recover for several decades indicating that this channel remains an important source of the productivity growth slowdown observed in the data.

Outline The remainder of this paper is organized as follows. Section 2 presents the model. Section 3 presents an overview of the data used as well as the main empirical results. Section 4 presents the calibration strategy and the estimated parameters. Section 5 presents the quantitative results. Section 6 concludes.

[4]Another closely related consequence of ICT diffusion is that it increases the scale of firms (e.g. Bloom et al., 2014; DeStefano et al., 2018; Lashkari et al., 2021). In my model, adopters (ICT firms) tend to have a larger production scale as ICT firms invest more in R&D and expand into more product lines, which increases sales and employment.

[5]Brynjolfsson et al. (2021) provide another explanation resulting from the poor measurement of complementary, intangible investment during early stages of new GPTs.

2 Model

I develop a model of innovative firms that adopt and diffuse a newly introduced general purpose technology, which is taken to be ICT (information communication technology), to be consistent with the empirical context. ICT is a firm type that dictates: (1) a persistent component of product quality, capturing the novelty of ICT products to consumers; (2) the relative costs and returns of innovations; and (3) the innovative interactions with other firms through knowledge and competition spillovers. In this regard, ICT firms capture a broader notion of ICT adoption (as opposed to using ICT equipment) in that ICT become embedded into the firm's production and innovation process. ICT is diffused through either direct adoption by non-ICT firms or innovation into new product lines by ICT firms.

2.1 Preferences and Production

Time is continuous and indexed by $t \in [0, \infty)$. The economy consists of a representative household and an endogenous mass of heterogeneous firms. Households supply labour to the production of goods. Firms produce goods and invest in R&D to innovate or adopt ICT. ICT is introduced to the economy in period 0 and all firms are initially non-ICT firms.

Households. The representative household is populated by a unit mass of workers with preferences described by:

$$U([C(t)]_{t=0}^{\infty}) = \int_0^{\infty} \exp\{-\rho t\} \log(C(t)) dt \tag{1}$$

where $C(t)$ is the consumption of the final good in period t and ρ is the household's discount rate. Members supply labor inelastically in the production of goods.

Production. The consumption good is produced by a representative producer using a continuum of differentiated intermediate goods $j \in [0, 1]$ produced by firms f. The production technology of the representative producer is:

$$\ln Y(t) = \int_0^1 \ln \left[\sum_{f \in \mathcal{F}_j} \eta_{j_f}(t) q_{j_f}(t) k_{j_f}(t) \right] dj, \tag{2}$$

where $k_{j_f}(t)$ the quantity of intermediate good j purchased from firm f; $q_{j_f}(t)$ is a quality component of good j that depends on the history of innovations; η_{j_f} is a quality component associated with whether the good is an ICT or non-ICT good; and $\mathcal{F}_j$ is the set of potential producers of good j. The ICT component of quality is given by $\eta_{j_f} = \eta^{\kappa_f}$ where $\kappa_f = ict$ if the firm uses ICT and $\kappa_f = non$ if the firm does not. The value η^{non} is normalized to 1 such that $\eta^{ict} \geq 1$ captures the relative preference of consumers for ICT products. For example, η^{ict} could be thought of as the novelty of computerized products to consumers.

Firms differ in whether they produce ICT goods $\kappa_f \in \{ict, non\}$, the number of goods that the firm produces n_f, the goods that the firm can produce $j_f \in \mathcal{J}_f$, and the associated quality of these goods $q_{j_f} \in \mathcal{Q}_f$. The production of intermediate good j_f by firm f is linear in labour ℓ and equal to $k_{j_f}(t) = \ell_{j_f}(t)$. ICT use directly affects production through the fixed quality component η and indirectly through the evolution of qualities q through innovations.

2.2 Research and Development

Incumbent firms invest in exploration and adoption R&D. Exploration R&D expands the set of goods that firms produce, such that $\mathcal{J}'_f = \mathcal{J}_f \cup \{j\}$. Adoption R&D switches non-ICT firms to ICT firms, such that $\kappa'_f = ict$. New firms are created through entry R&D.[6]

Exploration R&D. Exploration R&D follows Klette and Kortum (2004) in which successful innovators create a new, higher quality variety of a randomly drawn good $j \in [0, 1]$, which allows them to displace the incumbent producer. My point of departure is to allow for non-ICT and ICT firms to differ in terms of cost and returns of exploration R&D as well as in their innovative interactions with other firms.

A firm that invests S_x in R&D innovates at rate $A^\kappa(t) Z_x = A^\kappa(t) \left(S_x / \psi_x^\kappa Y(t)\right)^{1/\zeta_x} n_f^{1-1/\zeta_x}$, where $\psi_x^\kappa > 0$ and $\zeta_x > 1$ parameterize the level and curvature of the cost function; n_f is a firm-specific spillover term that depends on the number of goods produced by firm f, as in Klette and Kortum (2004); $A^\kappa(t)$ is a technology-specific term that determines knowledge and competition spillovers for ICT and non-ICT firms; and $Y(t)$ is a scaling term for the size of the economy.[7] The scaled cost of exploration R&D, $s_x = S_x/n_f$, can be written in terms of the scaled innovation intensity, $z_x = Z_x/n_f$, as:

$$s_x^{\kappa_f}(z_x, t) = \psi_x^{\kappa_f} z_x^{\zeta_x} Y(t).$$

[6] The literature also highlights the importance of innovator heterogeneity (as in Acemoglu et al., 2015; Ayerst, 2020b) and different forms of R&D (as in Akcigit and Kerr, 2018; Garcia-Macia et al., 2016). Earlier versions of this paper incorporated these features and found broadly similar quantitative results for growth and diffusion dynamics. Adding innovator heterogeneity reveals that innovation and adoption R&D are skewed towards a small subset of highly innovative (superstar) firms. Allowing firms to improve existing product lines (internal or exploitation R&D) has a quantitatively small effect because this type of innovation accounts for a small fraction of total innovation, as documented by Akcigit and Kerr (2018).

[7] Scaling R&D costs is necessary to avoid explosive growth. The scaling could also be thought of as R&D requiring some scarce resource as an input (e.g., skilled labor).

At rate $A^\kappa(t) z_x n_f$ exploration R&D draws an idea that improves the quality of a good $j \in [0,1]$ to $q_{j_f} = \lambda^{\kappa_f} q_{j_{f'}}$ for $\lambda^{\kappa_f} > 1$, where f is the innovating firm and f' is the current highest-quality producer of good j. The innovated good j is drawn randomly but is more likely to improve goods with the same technology κ as the innovating firm. The technology-specific spillover is set to

$$A^\kappa(t) = \mathcal{D}^\kappa(t) + \alpha^\kappa(1 - \mathcal{D}^\kappa(t)) \tag{3}$$

where $\mathcal{D}^\kappa(t)$ is the share of κ goods and $\alpha^\kappa \in [0,1]$ is a parameter that describes the relative likelihood of κ innovations drawing κ goods. A firm with innovation intensity z_x will draw goods with the same κ at rate $\mathcal{D}^\kappa(t) z_x n_f$ and goods with the other κ at rate $\alpha^\kappa(1-\mathcal{D}^\kappa(t)) z_x n_f$. Within each κ, the innovated good is drawn randomly from a uniform distribution.

I refer to α^κ as the cross-technology spillover because it dictates the strength of knowledge and competition spillovers between firms using different general purpose technologies. At the extremes, $\alpha^{non} = \alpha^{ict} = 1$ implies that knowledge is perfectly transferable and is nesting case for the standard Klette and Kortum (2004) innovation interactions. The case where $\alpha^{non} = \alpha^{ict} = 0$ implies that knowledge is perfectly isolated for each technology. For tractability, I assume that $\alpha^{non} = 0$ such that the non-ICT firms cannot innovate on goods produced by ICT firms. Appendix D.4 examines the sensitivity of the results to $\alpha^{non} > 0$.

The technology-specific spillover $A^\kappa(t)$ captures the unique interactions between general purpose technologies and innovation. Intuitively, the cross-technology spillover captures that firms tend to learn and compete more with firms using similar technologies. For example, an ICT-using car manufacturer may learn more from innovations by ICT firms (e.g., navigation systems) than firms in adjacent markets with similar products (e.g., bicycle manufacturers). More broadly, the cross-technology spillovers capture, in a reduced form, any factors that would lead to differences in spillovers across non-ICT and ICT firms. For example, cross-technology spillovers could capture the probability that ideas stemming from production—captured by n_f in the R&D cost function—arrive for goods using the same technology. Appendix A.4 provides a micro foundation for $A^\kappa(t)$ from costly conversion of non-ICT goods and differences in the relative benefit of converting goods (i.e., different values of η).

An alternative modeling assumption would be that ICT innovations become more beneficial over time through a larger step size λ^{ict}. In Appendix B.4, I show that diffusion through ICT innovations is mainly through the number of innovations and that the relative average quality of ICT and non-ICT innovations remains constant over time.

Adoption R&D. Adoption is modeled as a stochastic process in which firms invest to learn how to apply ICT to their goods. This reflects the idea that general purpose technologies, such as ICT, are a method of invention and integrated into the innovation process (e.g., inventing the electronic control unit), rather than as an input in production (e.g., buying computer equipment). Successful adoption switches the firm from a non-ICT firm ($\kappa = non$) to an ICT firm ($\kappa = ict$) and increases the quality component η from $\eta^{non} = 1$ to η^{ict} for all of the firm's goods. Like exploration R&D, firms choose a rate of adoption z_a by investing:

$$S_a(z_a, t) = s_a(z_a, t)n_f = \psi_a z_a^{\zeta_a} n_f Y(t),$$

where $\psi_a > 0$ and $\zeta_a > 1$ are cost function parameters. As with exploration R&D, the cost of adoption R&D scales with firm size through the number of products n_f.

Entry and exit. Entry occurs through successful research efforts to innovate on existing goods. Potential entrants choose the probability of entry z_e by investing $s_e(z_e) = \psi_e z_e Y(t)$. Successful entrants draw a good $j \in [0, 1]$ at random and enter with the same κ as the incumbent producer.[8] The quality step λ^κ is determined after the entrant's type κ.

Firms exit when they no longer produce any goods. Let $\bar{z}_x^\kappa(t)$ be the product-weighted average innovation intensity of κ firms, then a firm f loses products at rate:

$$\delta^{\kappa_f}(t) = z_e(t) + \left[\mathcal{D}^{\kappa_f}(t)\bar{z}_x^{\kappa_f}(t) + (1 - \mathcal{D}^{\kappa_f}(t))\alpha^{-\kappa_f}\bar{z}_x^{-\kappa_f}(t)\right], \quad (4)$$

where $-\kappa_f = \{ict, non\}/\kappa_f$. The expression shows how α^κ captures the competition spillovers across ICT and non-ICT firms. Larger $\alpha^{-\kappa_f}$ implies that the ICT (non-ICT) firms are more likely to be displaced by non-ICT (ICT) firm innovations.

2.3 Equilibrium

I now characterize the equilibrium of the economy. The analysis focuses on the diffusion of ICT throughout the economy. I assume that ICT is eventually diffused, which assumes that the ICT parameters $(\psi^{ict}, \lambda^{ict}, \psi_a, \eta^{ict})$ compare favorably with the non-ICT parameters $(\psi^{non}, \lambda^{non})$.[9] However, this assumption does not require that the long-run growth rate will

[8] Appendix B.8 shows that in the data entrants are slightly more likely, around 2.5%, to be ICT firms than the implied firm distribution. This channel is quantitatively small and so I abstract from it.

[9] This is the empirically relevant case for ICT. For $\alpha^{ict} > 0$ and $\lambda^{ict} > 1$, any equilibrium path that includes some diffusion of ICT implies that ICT will become fully diffused in the long run since ICT firms will always choose a strictly positive level of exploration R&D.

increase.[10] Appendix A.2 presents the balanced growth path equilibrium.

2.3.1 Production

Equilibrium production is standard and is briefly summarized here, with a full description provided in Appendix A.1. The final consumption good is taken to be the numeraire.

Firms compete in Bertrand competition, leading the highest quality producer to charge the limit price for good j with markup $\mu_j = \eta_{j_f} q_{j_f} / \eta_{j_{f'}} q_{j_{f'}}$ (where f' is the next lowest price-per-quality firm). The impact of ICT on production is then through the markups charged by ICT firms innovating on non-ICT goods, where η differs. The markup distribution $\Phi(\mu, t)$ is a key state variable for describing the distribution of production and prices as well as aggregate output in the economy. The markup distribution depends on the distribution of ICT goods and the previous producer of the good.

2.3.2 Research and Development

Dynamic problem. A good with markup μ earns profits $\pi(\mu, t) = (1 - \mu^{-1})Y(t)$. In what follows I drop firm f subscript and refer to firms based on type $\kappa \in \{ict, non\}$, the markups charged on actively-produced goods $[\mu_i]_{i=1}^{n}$, and the cardinality of this set n. The firm's dynamic problem is to set exploration and adoption R&D to maximize value:

$$\begin{aligned} rV_n^\kappa([\mu_i]) = \max_{z_x, z_a} & \sum_{i'=1}^{n} \left\{ (1 - \mu_{i'}^{-1})Y + \delta^\kappa \left[V_{n-1}^\kappa([\mu_i]/\{\mu_{i'}\}) - V_n^\kappa([\mu_i]) \right] \right\} \\ & + z_x n \left[\left[\mathcal{D}^\kappa(t)[V_{n+1}^\kappa([\mu_i] \cup \{\lambda^\kappa\}) - V_n^\kappa([\mu_i])] + \alpha^\kappa \mathcal{D}^{-\kappa}[V_{n+1}^\kappa([\mu_i] \cup \{\eta^\kappa \lambda^\kappa\}) - V_n^\kappa([\mu_i])] \right] \right] \\ & - n s_x^\kappa(z_x) Y + z_a^\kappa \left[V_n^{ict}([\eta^{ict} \mu_i]) - V_n^\kappa([\mu_i]) \right] - n s_a(z_a) Y + \dot{V}_n^\kappa([\mu_i]) \end{aligned} \tag{5}$$

where t is suppressed for brevity. I write the above expression under the restriction that adoption intensity $z_a^{ict} = 0$ for ICT firms such that the expression can be stated parsimoniously. Firm value depends on five components: (i) the current operating profits of the firm's products; (ii) the possibility of losing products to innovations by other firms; (iii) the net expected value of potentially of adding new products through exploration R&D; (iv) the net expected value of potentially adopting ICT; and (v) the change in firm value over time. Proposition 1 characterizes the value and policy functions.

[10] ICT may be privately preferable but not socially preferable or adoption may lead to large short-term gains that offset lower productivity growth in the long run. As a simple example, consider a parameterization where η^{ict} is very large and the ICT exploration R&D technology $(\lambda^{ict}, \psi_x^{ict})$ implies a lower BGP growth rate than the non-ICT BGP. Firms adopt to gain the immediate benefits of the high η^{ict} but transition to a lower growth BGP in the long run.

Proposition 1. *The value of a firm is described by:*

$$V_n^\kappa([\mu_i], t) = \sum_{i'=1}^{n} \left[B^\kappa(t) + v^\kappa(t)(1 - \mu_{i'}^{-1})\right] Y(t). \tag{6}$$

where $v^\kappa(t)$ and $B^\kappa(t)$ are defined in Appendix A. Exploration R&D is given by:

$$z_x^\kappa(t) = \left[\frac{\mathcal{D}^\kappa(t)[v^\kappa(t)(1 - (\lambda^\kappa)^{-1}) + B^\kappa(t)] + \alpha^\kappa \mathcal{D}^{-\kappa}(t)\left[v^\kappa(t)\left(1 - (\lambda^\kappa \eta^\kappa)^{-1}\right) + B^\kappa(t)\right]}{\zeta_x \psi_x^\kappa}\right]^{\frac{1}{\zeta_x - 1}}, \tag{7}$$

and adoption R&D is given by:

$$z_a(t) = \left[\underbrace{\frac{v^{ict}(t)\frac{1-(\eta^{ict})^{-1}}{\lambda^{non}}}{\psi_a \zeta_a}}_{Direct\ Benefit} + \underbrace{\frac{[v^{ict}(t)(1 - (\lambda^{non})^{-1}) + B^{ict}(t)] - [v^{non}(t)(1 - (\lambda^{non})^{-1}) + B^{non}(t)]}{\psi_a \zeta_a}}_{Indirect\ Benefit}\right]^{\frac{1}{\zeta_a - 1}}. \tag{8}$$

Firm value in (6) is composed of a portfolio value $v^\kappa(t)$ that captures the expected net present value of the firm's currently produced goods and an option value $B^\kappa(t)$ that captures the net value from potentially adding new products and adopting ICT.

Exploration R&D depends on the probability of a firm successfully innovating and the value of a new product line to the firm. All else equal, ICT firms invest more in exploration R&D when ICT is more common (higher $\mathcal{D}^{ict}(t)$) or if cross-technology spillovers α^{ict} are high. Adoption R&D depends on the direct and indirect benefits of adoption. The direct benefit captures that adopting ICT has a direct benefit on product quality through η^{ict} allowing firms to charge a higher markup. The indirect benefit captures the benefit from operating as a ICT firm relative to a non-ICT firm, which depends on the extent of knowledge and competition spillovers for each firm type. For example, early adopters of ICT face low competition from non-ICT firms, which increases the indirect benefits from adoption.

Potential entrants invest $\psi_e z_e$ to enter at rate z_e. The entry problem is:

$$V_e(t) = \max_{z_e} z_e \left[\left[\mathcal{D}^{ict}(t) V_1^{ict}(\lambda^{ict}, t) + \mathcal{D}^{non}(t) V_1^{non}(\lambda^{non}, t)\right] - V_e(t)\right] - \psi_e z_e Y(t), \tag{9}$$

where $V_e(t)$ is the value of a potential entrant at time t. Entry R&D is positive when the expected value of creating a new firm is at least as great as the marginal cost of entry. The linearity of the problem implies that the value of potential entrants equals zero in equilibrium when entry is positive $z_e > 0$, such that the expected value of a new firm equals the marginal cost of entry.

Aggregate dynamics. The aggregate dynamics of the economy are described by the share of ICT goods $\mathcal{D}^{ict}(t)$ and the growth rate of average good quality $g(t)$. The law of motion for the share of ICT goods $\mathcal{D}^{ict}(t)$ is given by:

$$\frac{d\mathcal{D}^{ict}(t)}{dt} = \underbrace{\alpha^{ict} z_x^{ict}(t)\mathcal{D}^{ict}(t)\mathcal{D}^{non}(t)}_{\text{Expansion Channel}} + \underbrace{\mathcal{D}^{non}(t) z_a(t)}_{\text{Adoption Channel}}, \tag{10}$$

where $\mathcal{D}^{non}(t) = 1 - \mathcal{D}^{ict}(t)$. The share of ICT goods $\mathcal{D}^{ict}(t)$ is the main measure of diffusion in the quantitative analysis because it is directly related to firm choices (see Proposition 1) and aggregate growth.

ICT diffusion in (10) is the first outcome of the model. The expansion channel depends on the total exploration R&D intensity of ICT firms $z_x^{ict}(t)$ multiplied by the number of ICT good $\mathcal{D}^{ict}(t)$ (since innovation scales with products) and the probability an ICT innovation draws and converts a non-ICT product $\alpha^{ict}\mathcal{D}^{non}(t)$. The adoption channel depends on adoption R&D intensity $z_a(t)$, which may increase or decrease in $\mathcal{D}^{ict}(t)$ depending on the relative strength of knowledge and competition spillovers. The strength of the two channels depends on the relative characteristics of non-ICT and ICT firms. All else equal, more innovative ICT firms (higher $z_x^{ict}(t)$ from lower costs ψ_x^{ict} or higher returns λ_x^{ict}) implies a stronger expansion channel while cheaper adoption (lower ψ_a) implies a stronger adoption channel.

Long-run output growth depends on growth in the average good quality $\bar{q}(t)$.[11] Average quality growth is given by

$$g(t) = \left[\sum_{\kappa} \left[\mathcal{D}^{\kappa}(t)\ln\lambda^{\kappa} + \alpha^{\kappa}\mathcal{D}^{-\kappa}(t)\ln\lambda^{\kappa}\eta^{\kappa}\right] z_x^{\kappa}(t)\right] + z_a(t)\mathcal{D}^{non}(t)\ln\eta^{ict} + z_e(t)\ln\lambda_e(t), \tag{11}$$

where $\ln\lambda_e(t) = \mathcal{D}^{non}(t)\ln\lambda^{non} + \mathcal{D}^{ict}(t)\ln\lambda^{ict}$ is the average step size of entrants.

Growth in (11) is the second outcome of the model. Growth depends on the relative use of ICT $\mathcal{D}^{ict}$, investment in each type of R&D (exploration, adoption, and entry), and the relative step size of each innovation activity. Growth is highly dependent on ICT diffusion, which determines both the research productivity of ICT and non-ICT firms (through knowledge spillovers) and the relative R&D allocations between ICT and non-ICT innovations.

[11] Output growth also depends on misallocation of factors of production (described in Appendix A.1).

2.3.3 Equilibrium Definition

Definition 1. *An equilibrium is a sequence of values:*

$$[r(t), w(t), C(t), p(\mu,t), k(\mu,t), \ell(\mu,t), z_x^{\kappa}(t), z_a(t), z_e(t), \mathcal{D}^{ict}(t), \delta^{\kappa}(t), \bar{q}(t), \Phi(\mu,t)],$$

for $\kappa \in \{ict, non\}$, $\mu \in \{\lambda^{non}, \lambda^{ict}, \eta^{ict}\lambda^{ict}, \eta^{ict}\lambda^{non}\}$ *and* $t \in [0,\infty)$ *such that: (i) given prices,* $C(t)$ *maximize household utility in (1); (ii) given prices,* $k_j(t)$ *maximizes the final good producer's profits; (iii) given* $w(t)$ *and demand,* $\ell(\mu,t)$ *and* $p(\mu,t)$ *maximize firm profits; (iv)* $z_a(t)$ *and* $z_x^{\kappa}(t)$ *maximize firm value in (5); (v) potential entrants maximize (9); (vi) the distribution of markups follows (17); (vii) growth of average quality* $\bar{q}(t)$ *follows (11); (viii) the creative destruction rate* $\delta^{\kappa}(t)$ *is given by (4); (ix) the ICT good share follows (10); (x) the labor, intermediate goods, and final good markets clear.*

2.4 Discussion

The key model outcomes describe ICT diffusion, in (10), and growth, in (11). The quantitative analysis uses the model structure to determine the drivers of ICT diffusion and the implications for aggregate productivity growth. Table 1 shows how the underlying model parameters relate to the main outcomes when future states of the economy are held fixed. This simplifying assumption is necessary since both diffusion and growth depend on highly interconnected and forward looking firm-level decisions.[12] The table ignores the feedback between channels to firm value, which are examined in the quantitative analysis.

Increasing the research productivity of adoption R&D is beneficial for growth, since adoption increases quality through η^{ict}, and diffusion, through the adoption channel. Increasing research productivity of ICT and non-ICT firms increases growth but increasing non-ICT research productivity can have an ambiguous effects on ICT diffusion. Increasing the non-ICT innovation step size λ^{non} increases the value of currently produced products ($v^{\kappa}(t)(1-(\lambda^{\kappa})^{-1})$ in Proposition 1), which can incentivize adoption if ICT firms face lower competition spillovers (lower δ^{κ}), but also decreases the marginal profitability gains from adoption (lowering adoption). Increasing the research productivity of ICT firms increases diffusion through more ICT innovations (expansion channel). Increasing cross-technology spillovers is qualitatively similar to increasing ICT research productivity since it results in more ICT innovations for a given R&D investment.

[12] Additionally, since growth may be higher or lower along the transition than on the balanced growth path, it is difficult to make direct predictions about the relationship between growth and ICT diffusion and the underlying model parameters. For example, say that growth temporarily declines over the transition path. Decreasing adoption costs could increase adoption initially resulting in lower growth over the transition path as well as a shorter transition path.

Table 1: Model Growth and Diffusion

		Growth	Diffusion		
		$g(t)$	$\frac{d\mathcal{D}^{ict}(t)}{dt}$	Expansion	Adoption
Adoption Productivity	$\uparrow \eta^{ict}, \downarrow \psi_a$	$\uparrow$	$\uparrow$	$-$	$\uparrow$
Non-ICT Research Productivity	$\uparrow \lambda^{non}, \downarrow \psi_x^{non}$	$\uparrow$	$\uparrow\downarrow$	$-$	$\uparrow\downarrow$
ICT Research Productivity	$\uparrow \lambda^{ict}, \downarrow \psi_x^{ict}$	$\uparrow$	$\uparrow$	$\uparrow$	$-$
Cross-Technology Spillovers	$\uparrow \alpha^{ict}$	$\uparrow$	$\uparrow$	$\uparrow$	$-$

Notes: Change in growth $g(t)$ and diffusion $d\mathcal{D}^{ict}(t)/dt$ for a change in the indicated parameter on a fixed path of $\mathcal{D}^{ict}(t)$ and $g(t)$. The results follow from (10) and (5) and Proposition 1 when the values $(v^\kappa(t), B^\kappa(t))$ are held fixed. $\uparrow$ and $\downarrow$ indicate a positive or negative direct impact, $-$ indicates no direct impact, and $\uparrow\downarrow$ indicates an ambiguous impact.

3 Empirical Analysis

The model describes ICT diffusion through goods and innovative capability, as opposed to ICT equipment. In this regard, diffusion of ICT equipment or capital stock does not capture the same diffusion as in the model. Given this, I start by constructing a new measure of ICT-related patents that captures the application of ICT to new inventions. ICT-related patents capture a broader measure of diffusion, compared with looking at ICT patents directly, and closely relates to the model concept of diffusion.

The final part of the section examines differences in ICT and non-ICT innovators. This serves two goals. First, the analysis validates the construction of ICT-related patents by showing that ICT-related firms and innovations differ from non-ICT firms and innovations. Second, the analysis disciplines key model parameters related to the relative research productivity (through λ^κ, ψ_x^κ, and η^{ict}) of ICT and non-ICT firms.

3.1 Data Sources

NBER USPTO Utility Patent Grant Database. The dataset contains information on all patents granted by the United States Patent and Trademark Office from 1976 to 2006. Citations are adjusted using Hall et al. (2001) weights to correct for sectoral and temporal differences in the likelihood of being cited. Additionally, I focus on within year and sector measures when comparing ICT and non-ICT firms to reduce the influence of aggregate trends, such as declining research productivity (Bloom et al., 2020). The date of a patent is taken to be the application date to avoid the influence of bureaucratic delays. For the analysis, the sample is restricted to patents issued to US non-governmental organizations and in the sub-period from 1980 to 2000. I trim the early periods because my measure of ICT-related patents is based on backward citations creating bias in early periods. Similarly, I trim the late periods because forward citations become truncated and to avoid the inclusion of the Dot-Com bubble. Variables are constructed using the full length of data and all patents in the dataset (i.e., not just those to non-governmental US organizations).

Compustat Database. I use the Compustat database to connect the patent measures to firm-level financial outcomes. The main financial variables of interest are 4-digit SIC classification, sales, number of employees, and R&D expenditure. To construct the final sample, I drop all non-US corporations; all firms in the financial and utilities sectors[13]; firms that experience major mergers or acquisitions; firms that report negative sales; and firms that do not patent over the sample. Finally, I winsorize all variables at the top and bottom 1% level using annual breakpoints to reduce the influence of outliers.

3.2 Construction of ICT-Related Variables

The challenge with mapping the data to the model involves constructing a measure of ICT that captures the application of ICT ideas throughout the economy. Patents in the *Computers & Communication* technology class capture innovations directly related to the ICT products but do not capture innovations in other markets that build on ICT innovations. For example, a car's diagnostic system may build on a microprocessor even though the end product is not directly classified in the *Computers & Communication* product market.[14]

To address this challenge, I use citations links between patents to construct a set of ICT-related patents that captures the application of ICT innovations to other markets. I then use this measure to construct empirical counterparts to the model concepts of innovations, firms, and products. Table 2 summarizes the constructed empirical variables.

At the center of these variables is *ICT distance* that measures the numbers of citations links between each patent and the *Computers & Communications* technology class. This provides a simple measure of how closely related a patent is to core ICT innovations. *ICT-related patents* are defined as patents with distance less than or equal to three. The cutoff distance is based on the distance that maximizes differences between adopters and non-adopters, discussed in Appendix B.2. *ICT-related firms* are firms in years after *ICT-related patents* exceed a 10% threshold of the firm's patent portfolio for the first time.[15] Finally, *ICT-related products* are technology subclasses with a significant share of *ICT-related patents*.[16]

[13]Specifically, SIC codes from 4900 to 5000 (utilities firms) and from 6900 to 7000 (financial firms).

[14]This example is from Patent 5,633,458 issued by Ford Motor Company. This patent has distance 2 to the ICT technology class using my measure.

[15]Appendix B.6 shows the robustness of the main results to the choice of cutoff.

[16]The measure of ICT-related products is used as a check on the fit of the model rather than being directly

Table 2: ICT Innovations, Firms, and Products

	Definition
ICT distance	Recursively defined. A patent is distance d if at least 10% of its citations are on distance $d-1$ or lower patents. *Computers & Communications* patents are distance $d = 0$.
ICT-related patents	Patent observations with *ICT distance* three or less.
ICT-related firms	Firm-year observations after ICT patents exceed 10% of patents in the firm's patent portfolio for the first year.
ICT-related products	Patent class (three-digit USPTO classification) by year observations where at least 25% of the depreciated patent stock are ICT-related patents.

Table 3: Patent Summary Statistics

	ICT-related	Non-ICT	All
Number of Observations	396,438	595,098	991,536
1980-1990	85,212	268,915	354,127
1990-2000	311,226	326,183	637,409
Average Citations	24.2	13.2	17.6

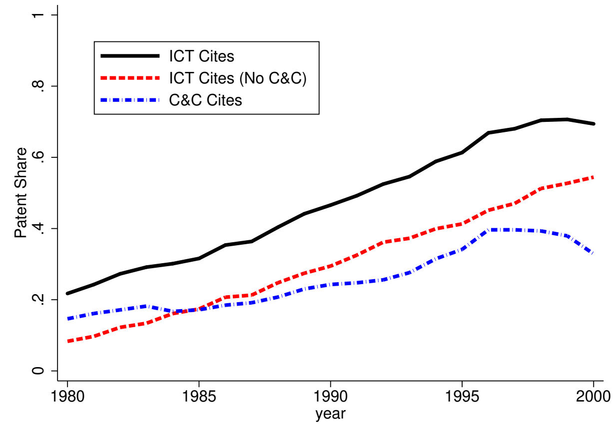
Table 3 reports summary statistics for ICT-related and non-ICT related patents. The fraction of ICT-related patents increases substantially in the latter half of the sample. Additionally, ICT-related patents are, on average, cited more frequently than non-ICT patents. However, this may mechanically be driven by ICT-related patents appearing more frequently in later periods.

Figure 1 reports different measures of ICT diffusion over time. *Computers & Communication* patents increase from just under 20% of total patents in 1980 to around 25% in 2000 while ICT-related patents increase from just over 20% in 1980 to around 70% by 2000. The difference between the two trends is that the share of ICT-related patents outside of the *Computers & Communication* technology class increases from around 10% to 45% over the period.

Figure 1 also shows that the ICT diffusion has been comparatively slower for innovations than for ICT equipment, which took around 30 years (Jovanovic and Rousseau, 2005). The

used in any of the analysis. I adapt the classification from Acemoglu et al. (2016a) who define products at the three-digit SIC level. To maintain the full patent sample, I define products at the three-digit USPTO classification, which I refer to as the product class. I define a product class as ICT-related once 25% of its depreciated patent stock is ICT-related patents. Appendix B.3 provides details and shows that the measure is robust to using a 50% cutoff.

Figure 1: Diffusion of ICT in the United States



Notes: The figure reports fraction of total patents for different sequences. The solid line plots the fraction of citation weighted ICT-related patents. The dashed line plots the fraction of citations weighted ICT-related patents excluding *Computers & Communication* (C&C) patents. The dotted line plots the fraction of citation weighted *Computers & Communications* (C&C) patents.

slow diffusion of ICT through innovations is consistent with anecdotal evidence about the spread of ICT applications to new products. For example, the application of ICT to some adjacent products (e.g., smart watches, self-checkout) has occurred only recently whereas other products (e.g., smart glasses) have been unable to breakthrough into broad markets. In contrast, the adoption of ICT equipment into the workplace and homes was relatively quick. Consequently, ICT diffusion through innovations should be expected to have a longer-lived impact on growth, compared with ICT diffusion through equipment.

3.3 How do ICT and Non-ICT Firms Differ?

ICT and non-ICT firms differ in the model through their relative research productivity, which depends on the relative cost and returns to R&D ($\psi_x^\kappa, \lambda^\kappa$), and the relative quality of ICT and non-ICT goods to consumers η^κ. I use differences in innovation outcomes to discipline the R&D functions and firm-level growth around the time of adoption to discipline the relative quality of goods. Additionally, the analysis validates the measure of ICT-related patents by showing differences in innovation and market outcomes of adopters and non-adopters.

Innovation outcomes. I begin by examining the differences in firm innovation outcomes using only the patent data. Table 3 suggests that ICT-related patents receive on average around double the citations of non-ICT-related patents. However, this comparison may be overstated since ICT-related patents appear in later periods and different sectors.

Firm-level citations and patenting frequency differences between adopters and non-adopters are estimated as:

$$\text{Outcome}_{h,t} = \exp\left\{\theta \times \text{Adoption}_{h,t} + \Gamma_f + \Gamma_{s,t} + \varepsilon_{h,t}\right\} \tag{12}$$

where $\text{Outcome}_{h,t}$ is a measure of either innovation quality or frequency at either the patent $h = j$ (Columns 1 and 2) or firm $h = f$ (Columns 3 and 4) level for period t. The estimation also includes firm-level fixed effects Γ_f and $\Gamma_{s,t}$ is a technology class-by-year fixed effect for the patent-level regressions and a year fixed effect for the firm-level regressions. The firm-level regressions control for only year, rather than year by technology class, because some firms patent in multiple technology classes. Firm-level fixed effects are included in all four specifications to account for firm heterogeneity, such as better innovators being more likely to adopt. Table 4 reports the results.

Table 4: Innovative Outcomes and Adoption

	Citations-per-Patent		Patents-per-Period	
	(1)	(2)	(3)	(4)
ICT-Rel Patent	0.255***			
	(0.00453)			
Adopter		0.111***	0.203***	0.260***
		(0.00808)	(0.0297)	(0.0353)
Firm FE	Yes	Yes	Yes	Yes
Year FE	No	No	Yes	Yes
Tech. class × Year FE	Yes	Yes	No	No
Observations	945723	945723	169219	168646

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results are estimated using Poisson pseudo-maximum likelihood (PPML). Robust standard errors clustered at the firm level are reported in parentheses. The unit of observation in columns (1) and (2) is at the patent-level. The unit of observation in columns (3) and (4) is at the firm-level. Patents in column (3) are unweighted and patents in column (4) are weighted by citations. Technology classes are defined for *Chemical*, *Computers & Communications*, *Drug*, *Electronics*, *Mechanics*, and *Other* patents.

Column (1) confirms that ICT-related patents receive, on average, 29% ($\approx \exp(0.255)$) more citations compared to non-ICT patents issued in the same year and technology class, even after controlling for firm-specific differences. Column (2) shows that, on average, a firm's patent quality increases by 12% after adopting, compared with other firms in the same technology class and year. Both values are smaller than in Table 3, as expected. Appendix C.2 shows that the results are robust to using Kogan et al. (2017) patent values as quality measures, although the interpretation is different.

The coefficient in column (3) shows that, on average, adopters apply for 23% more patents than prior to adopting, relative to other firms in the same year. Column (4) shows that weighting by citations increases the coefficient estimate because adopters tend to apply for more highly cited patents. Together, the results indicate that adopters tend to be more frequent innovators following adoption.

The implication is that ICT firms tend to have higher returns to innovating, captured by λ_x^κ, and innovate more frequently, which could be a results of higher returns λ_x^κ, lower costs ψ_x^κ, or higher knowledge spillovers α^{ict}. The model is necessary to disentangle these factors.

Firm-level productivity. The limitation of the patent database is that it restricts to patent outcomes, which may not capture changes in economic activity. I use the linked Compustat-patent data to examine changes in firm-level labour productivity (sales-per-employee) following adoption.[17] Specifically, I estimate the relationship as:

$$g_{f,t} = \beta_0 + \sum_{t'=-4}^{4} \beta_{A,t'} \text{Adopt}_{f,t+t'} + \mathbf{X}_{f,t}\beta_X + \Gamma_s + \Gamma_t + \varepsilon_{f,t} \tag{13}$$

where $g_{f,t}$ is the labour productivity growth of firm f in year t; $\text{Adopt}_{f,t}$ takes value one if firm f adopts ICT in year t, and zero otherwise; $\mathbf{X}_{f,t}$ is a vector of controls; Γ_s is a sector (four-digit SIC) fixed effect; and Γ_t is a year fixed effect.[18] Controls used in the baseline specification are: the log of one plus the number of (forward) citations received by the patents applied for by firm f in period t; log of firm size (measured by employment); the log of one plus the firms R&D expenditure over sales; firm age in period t; as well as dummy variables to control for whether firm f patents in period t, whether firm f issues its first patent in period t, whether firm f issues a *Computers & Communications* patent in period

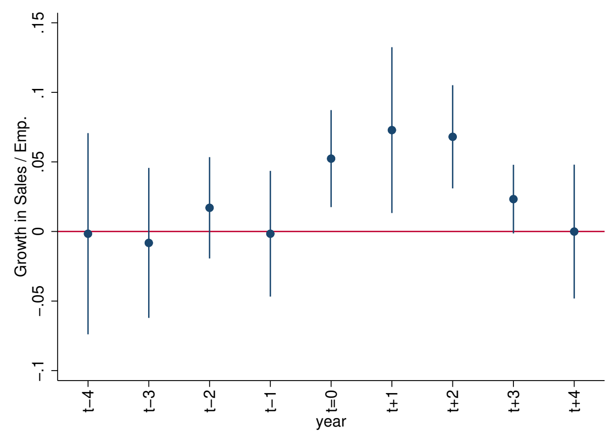
[17] A concern is that the relationship is being driven by either the endogeneity of the relationship (e.g., firms are more likely to adopt in high growth periods) or an omitted variables (e.g., new management leads to both adoption and high growth). In both cases, the timing of the growth increase provides some reassurance that adoption is driving growth. Specifically, given the difficulty of adopting ICT, alternative explanations should lead to growth increasing prior to adoption. Additionally, the robustness of the results to firm-level fixed effects and other firm-level controls (Appendix B.5) suggests that these alternative channels would need to be relatively short lived. Another concern is that adoption is mismeasured because either the timing is incorrect (i.e., adoption occurs earlier or later than the recorded year) or some non-adopters are being recorded as adopters. This would create downward bias on the coefficient estimates since these non-adopter firm-years should have average growth.

[18] Growth is $g_{f,t} = (Sales/Emp_{f,t} - Sales/Emp_{f,t-1})/(0.5(Sales/Emp_{f,t} + Sales/Emp_{f,t-1}))$. In the regressions, I do not deflate sales since changes in prices are captured by the year fixed effect. For the comparison with average firm growth, I deflate sales using GDP deflators from the BEA national income and product account tables.

t.[19] Standard errors are clustered at the two-digit SIC sector level.

The coefficient $\beta_{A,t'}$ is interpreted as the average increase in firm growth t' periods before or after adoption. Figure 2 reports the coefficient estimates with 95% confidence bands.

Figure 2: Labour Productivity Growth Before and After Adoption



Notes: The figure plots the coefficients $\beta_{A,t'}$ from the regression in (13). The bands report the 95% confidence interval based on robust standard errors clustered at the 2 digit SIC sector level. Year $t-1$ is the year prior to the adoption of the new technology. Average growth in sales-per-employee is 5.84%. Controls include the log of one plus the number of (forward) citations received by the patents applied for by firm f in period t; log of firm size (measured by employment); the log of one plus the firms R&D expenditure over sales; firm age in period t; as well as dummy variables to control for whether firm f patents in period t, whether firm f issues their first patent in period t and whether firm f issues a *Computers & Communications* patent in period t. The regression includes year and 4 digit SIC sector fixed effects.

Figure 2 reports that labour productivity growth doubles in the year of and two years following adoption. In the third year, labour productivity growth is elevated and statistically significant at the 10% level. The total increase in growth corresponds to a 3.7-fold increase in the average one-year growth rate of labour productivity. While labour productivity growth increases immediately following adoption, there does not appear to be a long-term effect on the growth rate. However, the temporary increase in the growth rate does imply a persistent increase in labor productivity between adopters and non-adopters. The growth pattern is consistent with the model in which adopting firms get a one-time increase in sales per employee from the increase in the ICT-specific component of product quality η^{ict}. The lack of a long-lasting increase in growth is also consistent with the model since exploration innovation increases the innovating firm's number of products but not markups.

[19] In addition, I check that the results are robust to a variety of alternative specifications: different time horizons; controls for patenting; different measures of size including employment and book value; research stock; R&D intensity; fixed effects measured at the 2 and 3 digit SIC code; firm-level fixed effects; including non-patenting firms.

Appendix B.5 shows that the results are robust to alternative controls (e.g., sector definition, firm controls), outcomes (e.g., employment growth), and estimation techniques. I also find that placebo measures of adoption based on other technology classes (e.g., chemicals) do not produce similar increases in growth.

4 Calibration and Measurement

I calibrate the model transition path to match ICT diffusion in the data. Empirical differences between adopter and non-adopter innovation characteristics discipline ICT and non-ICT innovation technologies and the aggregate diffusion curve (Figure 1) disciplines adoption R&D costs. A key aspect of the calibration is that moments related to growth over the transition period are not targeted since a goal of the calibration is to examine the model implications for growth. The next section examines the model-implied growth path and the drivers of ICT diffusion using the quantitative model.

4.1 Calibration Strategy

Calibration preliminaries. The discount rate ρ is set to 0.02, which corresponds to an annual discount rate of 4% on the initial balanced growth path (BGP). The curvature of the exploration R&D cost function is set to $\zeta_x = 2$, consistent with Acemoglu et al. (2015).[20]

Directly calibrated moments. The relative innovation step size is set to $\ln \lambda^{ict} / \ln \lambda^{non} = \exp(0.255)$ to match the relative patent quality of ICT-related and non-ICT-related patents from Column (1) of Table 4. The data moment follows the common assumption that more highly cited patents tend to be more valuable to the innovating firm (see Hall et al., 2005; Kogan et al., 2017; Abrams et al., 2018, for empirical support).[21]

Jointly calibrated moments. The parameters $\{\lambda^{non}, \psi_x^{non}, \lambda^{ict}, \psi_x^{ict}, \eta^{ict}, \psi_a, \zeta_a, \alpha^{ict}\}$ are jointly chosen to minimize the mean squared error between the model and data moments in Table 5. Unless otherwise stated, the moments are calculated as the mean value over the 1980 to 2000 period in both the data and model.

Model moments are calculated over the transition path that follows the introduction of the ICT. For a given set of parameter values, the transition path and moments are calculated as follows. To start, the BGPs in which no firms use ICT and all firms use ICT are solved (see

[20] I confirm the estimate by regressing $\ln(pat/sales_{f,t})$ on $\ln(R\&D/sales_{f,t})$ in levels and in differences with a firm fixed effect. The estimated coefficients, corresponding to $1/\zeta$, are 0.51 and 0.53.

[21] In the data firms with higher citations also tend to have higher growth (from Table 14).

Table 5: Model and Data Moments

Description of Moment	Data	Model
Initial BGP Growth (%)	2.00	1.99
Entry Share (%)	17.9	20.1
Profit Margin (%)	12.2	12.4
Relative Innovation Rate	0.203	0.203
Post-Adoption Growth	3.7	3.8
ICT Patent Share in 1999 (%)	70.7	71.9
MSE ICT Patent Share	0	0.04
Prob ICT Cites Non-ICT (%)	28.9	28.4

Notes: Model and data moments are calculated as mean values over the period 1980 to 2000 with the exception of the Initial BGP Growth and the ICT Patent Share in 1999.

Proposition 2 in Appendix A). The discretized model is then solved over the transition path that follows the introduction of ICT, with $\mathcal{D}^{ict}(0) = 0$, and ends on the full ICT diffusion BGP (see Appendix D.1 for details). Model periods are mapped into data years using the ICT patent share in 1980. The period that reaches 21.8% (the data share) of ICT patents in the model transition path is set to be 1980. This allows me to use the model to infer the date that ICT is introduced. Finally, moments are calculated in the model periods corresponding to 1980 to 2000.

4.2 Calibration Moments

I discuss the construction of the calibration moments and closely related parameters for intuition below. Additional details on the sensitivity of the model moments to parameters are reported in Appendix D.2.

Initial BGP Growth. The initial BGP growth rate is set to 2.0% which roughly corresponds to the post-war growth rate in the US until the start of the data in the late 1970s. The initial BGP growth rate is related to the parameters dictating the research productivity of non-ICT firms, $(\lambda_x^{non}, \psi_x^{non})$. I include the initial BGP growth rate rather than directly targeting growth over the transition since a goal of the estimation is to use the model to assess growth over the transition path and in the long run.

Entry Share. The data moment is the share of citation-weighted patents applied for by firms in their first year of observation in the dataset. For the calculation of this moment, I also top code the total patent count at 1,000 patents (affecting 159 firms or 0.1% of the sample) to reduce the influence of outliers. The model moment is calculated as average entry rate $z_e(t)$ divided by the average innovation rate $z_e(t) + A^{non}(t)z_x^{non}(t)\mathcal{D}^{non}(t) + A^{ict}(t)z_x^{ict}(t)\mathcal{D}^{ict}(t)$. The entry rate $z_e(t)$ is closely related to the cost of entry ψ_e but is also related to the costs and benefits of exploration R&D.

Profit Margin. The data moment is the median value of operating income before depreciation divided by sales.[22] The model moment is calculated as the average value of $1 - \mu^{-1}$ across products where the markup μ depends on the step sizes $(\lambda^{non}, \lambda^{ict}, \eta^{ict})$.

Relative Innovation Rate. The data moment is from Column (3) of Table 4. The model moment is calculated as $\ln A^{ict}(t)z_x^{ict}(t) - \ln A^{non}(t)z_x^{non}(t)$. The moment is closely related with the R&D technologies, $(\psi_x^\kappa, \lambda_x^\kappa)$.

Post-Adoption Growth. The data moment is the increase in labor productivity growth, divided by average firm-level growth, in the the periods following adoption (see Figure 2). The model moment is calculated as the relative growth of a firm at the time of adoption $\eta^{ict}/g(t)$ and is closely related to the ICT quality component η^{ict}. This loads the entirety of the productivity increase following adoption on a single point in time. Following adoption, the firm's labour productivity growth returns to its pre-adoption value, as in the data.[23]

ICT Patent Share in 1999. Diffusion in the data is measured as the ICT-related patent share. This does not directly correspond to the ICT product share since ICT firms may innovate more or less frequently than non-ICT firms. The model citation-weighted ICT-related patent share is:

$$\text{ICT Patent Share in } t = \frac{[\mathcal{D}^{ict}(t)z_x^{ict}(t)(\mathcal{D}^{ict}(t) + \alpha^{ict}\mathcal{D}^{non}(t))) + z_e(t)\mathcal{D}^{ict}(t)]\ln\lambda^{ict}}{\sum_\kappa [\mathcal{D}^\kappa z_x^\kappa(t)(\mathcal{D}^\kappa(t) + \alpha^\kappa(1 - \mathcal{D}^\kappa(t))) + z_e(t)\mathcal{D}^\kappa(t)]\ln\lambda^\kappa}. \quad (14)$$

The numerator measures total quality-weighted ICT innovations and the denominator measure the total quality-weighted innovations. I use (14) to construct two moments related to the length and shape of the diffusion. The first moment is the share ICT-related patents in

[22]Additionally, I calculate the statistic on a sub-sample of innovative firms that only includes firms that report profit margins above a 25% loss to avoid the influence of outliers. The target does not change much for alternative cutoffs—e.g., the target would be 13.5% if the loss cutoff is 10% and 10.5% if the loss cutoff is 50%—but is sensitive to including the bottom tail of the distribution where low sales or highly negative operating incomes skew the distribution.

[23]In the model, exploration innovation and creative destruction primarily affect the innovating firm's scale through the number of product lines implying that growth in sales-per-employee is close to output growth for both adopters and non-adopters. Exploration innovation and creative destruction may add or drop higher or lower markup products, but this tends to be negligible.

1999, since this is the highest recorded ICT-related patent share in the data. Diffusion in the model is closely related to the adoption R&D cost function parameters (ψ_a, ζ_a). Figure 5a shows the graphical comparison of the data and model.

MSE ICT Patents. The second moment targeted using (14) is the mean-squared error of ICT-related patents between the model and data over the full transition period (1980 to 2000). This moment is distinct from the previous moment in that it incorporates information on ICT diffusion from all years of the data, rather than just one. This helps to discipline the curvature of the adoption cost function ζ_a that determines the elasticity of adoption R&D investment to changes in the relative value of ICT.

Prob ICT Cites Non-ICT. The probability that an ICT patent cites a non-ICT patent is included as a measure of the knowledge spillovers from non-ICT to ICT innovations. The moment follows the idea of modeling ICT as a method of innovation, rather than a production input, and is supported by empirical evidence showing the importance of knowledge flows along citation networks (e.g., Acemoglu et al., 2016b).[24] The data moment is constructed as the average of the share of non-ICT patents cited by ICT patents, excluding citations to patents produced by the same firm. The model moment is constructed as

$$\text{Prob ICT Cites Non-ICT}_t = \frac{(z_x^{ict}(t)\mathcal{D}^{ict}(t))\alpha^{ict}\mathcal{D}^{non}(t)}{z_t^{ict}(t)\mathcal{D}^{ict}(t)\left[\mathcal{D}^{ict}(t)+\alpha^{ict}\mathcal{D}^{non}(t)\right]} = \frac{\alpha^{ict}(1-\mathcal{D}^{ict}(t))}{\mathcal{D}^{ict}(t)+\alpha^{ict}(1-\mathcal{D}^{ict}(t))}.$$

If the ICT product share $\mathcal{D}^{ict}(t)$ was directly observed in the data, the above expression could be rearranged to calculate α^{ict} directly.

4.3 Parameter Values

The parameter values from the calibration are listed in Table 6. Panel B reports the moments chosen to match moments in Table 5.

Table 6 shows that ICT firms have a relatively higher innovation step size ($\lambda^{ict} > \lambda^{non}$) but also face higher costs of exploration R&D ($\psi_x^{ict} > \psi_x^{non}$). The higher step size of ICT firms leads to increasing markups and profits over time, consistent with evidence from De Loecker et al. (2018). The parameters also imply that measured research productivity of non-ICT firms is higher than ICT firms such that long-run research productivity is declining, consistent

[24] See also evidence in Berkes et al. (2022), Liu and Ma (2022), and Ayerst et al. (2022) showing that productivity, firm value, and innovation are affected by shocks to upstream sector knowledge flows, as measured by citations links.

Table 6: Calibrated Parameters of the Model

A. Externally Calibrated Parameters		
Parameter		Value
Discount Rate	ρ	0.02
Crv Exploration R&D	ζ_x	2

B. Parameters Calibrated to Internal Targets		
Parameter		Value
Cost of Entry	ψ_e	1.41
Cost of Non-ICT Exploration R&D	ψ_x^{non}	2.80
Cost of ICT Exploration R&D	ψ_x^{ict}	3.20
Non-ICT Step Size	λ^{non}	1.119
ICT Step Size	λ^{ict}	1.157
Level Cost of Adoption R&D	ψ_a	10.66
Rel Quality of ICT Products	η^{ict}	1.055
Cross-technology Spillovers	α^{ict}	0.259
Crv Adoption R&D	ζ_a	1.67

with Bloom et al. (2020).[25] However, the impact on growth is ambiguous since the higher profitability of ICT innovations ($\lambda^{ict} > \lambda^{non}$) causes more resources to be allocated to R&D.

Adoption R&D compares unfavorably in both costs ($\psi_a > \psi_x^\kappa$) and returns ($\eta^{ict} < \lambda_x^\kappa$) to exploration R&D. The low relative return on adoption η^{ict} is seemingly in contradiction with the empirical results showing that adopting firms experience a large increase in productivity growth relative to average firm growth. The contradiction is resolved by noting that the comparison is between the ex-post adoption growth, which depends on only η^{ict}, and the ex-ante average growth, which depends on both λ^κ and the low value of z_x^κ.

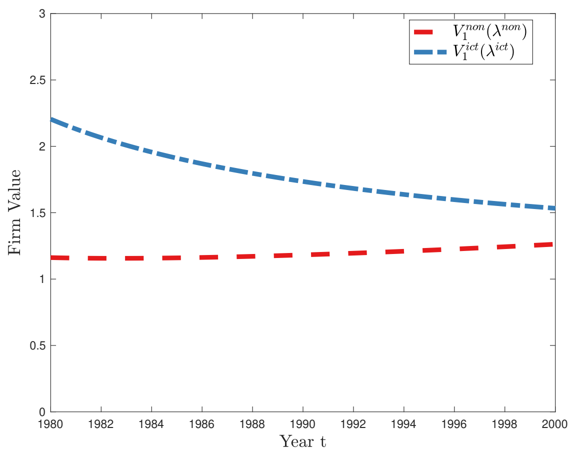
The cross-technology spillover term α^{ict} lies in the middle of the no spillover ($\alpha^{ict} = 0$) and perfect spillover ($\alpha^{ict} = 1$) parameterizations. Being greater than zero implies that ICT firms build on and compete with non-ICT firms while being less than one implies that ICT firms are relatively biased towards innovating on other ICT goods. On its own, the cross-technology spillover value implies that growth falls over the transition period as knowledge spillovers are lower when goods are divided between ICT and non-ICT firms.

[25] Following Bloom et al. (2020), measured research productivity is calculated as average growth in sales, $g(t) + (z_x^\kappa(t) - \delta^\kappa)$, divided by scaled R&D input, $\psi_x^\kappa (z_x^\kappa)^{\zeta_x}$. An alternative measure, in the spirit of TFP, would be to measure research productivity as growth generated with a fixed input of R&D, i.e., $\ln(\lambda^\kappa)(s_x/\psi_x^\kappa)^{1/\zeta_x}$. This would indicate that ICT firms have higher research productivity.

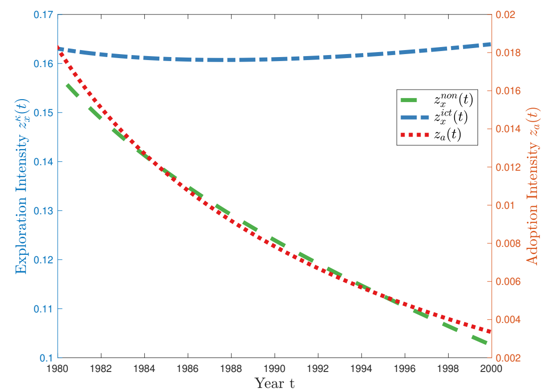
4.4 Value and Policy Functions

Figure 3 plots the value and policy functions over the transition period. To provide insight into the changes in the value and policy functions, Figure 4 plots the technology-specific spillover and the probability of losing products δ^{κ} for ICT and non-ICT firms.

Figure 3: Value and Policy Functions



(a) Value Functions



(b) Policy Functions

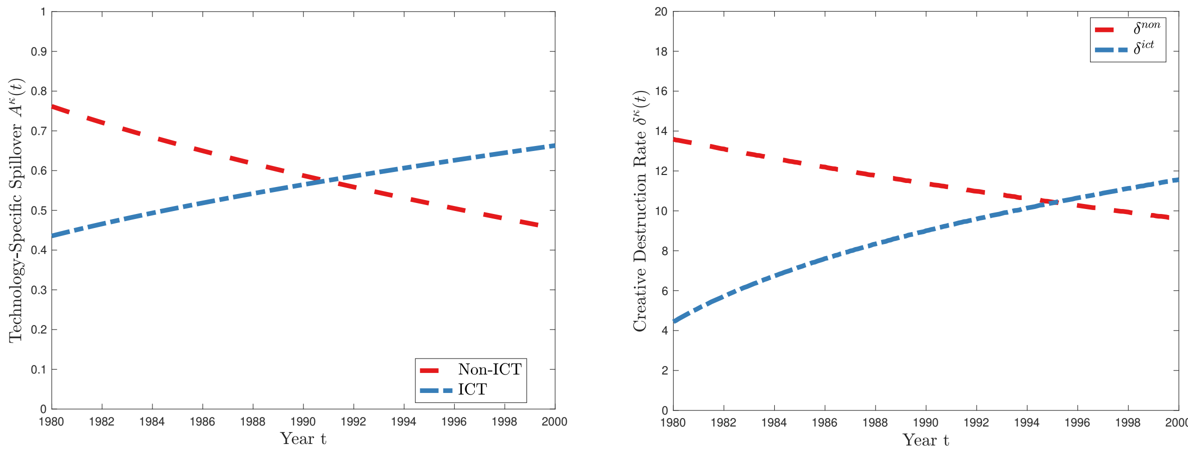
Notes: Figure (a) reports the value of non-ICT and ICT firms with a single product line and average markups over the transition period. Figure (b) reports the policy functions for exploration and adoption R&D over the transition period. Both sets of functions are described in Proposition 1.

The value of ICT firms declines steadily over the transition path whereas the value of non-ICT firms increases slightly. The change in firm values is driven by the relative strength of knowledge and competition spillovers over the transition. Knowledge spillovers increase for ICT firms over the transition period because ICT becomes more widely used making ICT firm innovation more productive (Figure 4a). Similarly, competition spillovers also increase for ICT firms over the transition period because increasing ICT innovation leads to ICT firms being more likely to lose products to competitors (Figure 4b). On net, the increase in competition spillovers dominates knowledge spillovers leading to ICT firm value falling over the transition, while the reverse holds for non-ICT firm value. The decline in the value of ICT firms is large enough to decrease ICT innovation intensity $z_x^{ict}(t)$ over the transition despite a contemporaneous increase in the ICT-specific spillover (Figure 4a).[26]

Adoption R&D also declines over the transition period. From (8), adoption R&D depends on both the direct and indirect benefits from adoption, which both decline over the transition.

[26] The relative values of z_x^{κ} and the technology-specific spillovers imply that ICT firms are larger in terms of employment and sales on average compared with non-ICT firms. While ICT adoption in the model does not necessarily correlate with ICT equipment use, the result is similar to empirical evidence showing that ICT adopters tend to have larger scale (e.g. Bloom et al., 2014; DeStefano et al., 2018; Lashkari et al., 2021).

Figure 4: Innovation Success and Creative Destruction



(a) Technology-Specific Spillovers

(b) Creative Destruction

Notes: Figure (a) reports the value of $A^\kappa(t) = \mathcal{D}^\kappa(t) + \alpha^\kappa(1 - \mathcal{D}^\kappa)$ that scales the innovation rate. Figure (b) reports the probability of losing a product $\delta^\kappa(t)$, given in (4), for both ICT and non-ICT firms.

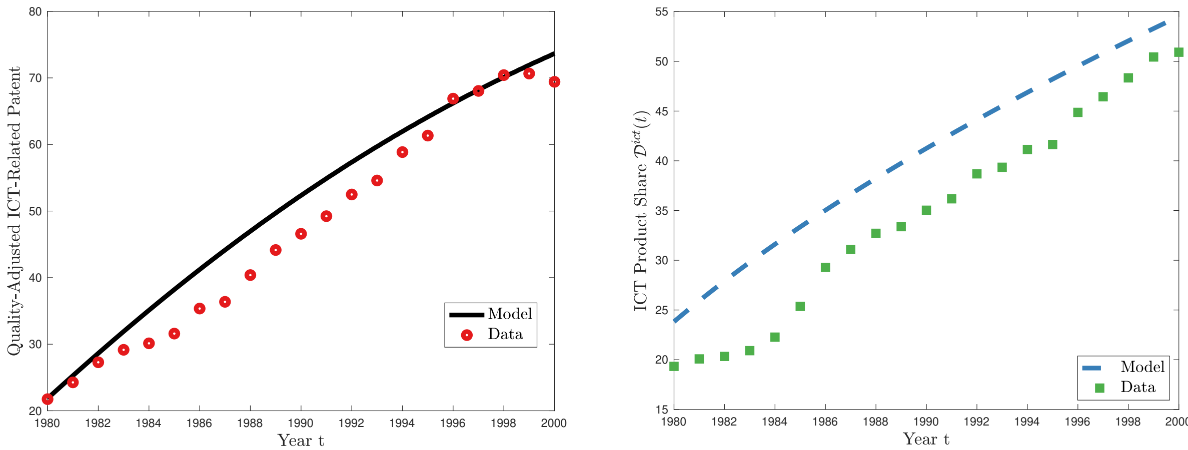
The direct benefits depend on the increase in product quality, which is constant, and ICT firm value, which declines. The indirect benefits depend on the relative value of ICT and non-ICT firms, which is positive but declines over the transition path. The decomposition of adoption intensity $z_a(t)$ into the direct and indirect benefits is reported in Appendix D.3.

4.5 Goodness of Fit and Other Model Moments

Figure 5a plots the (citation adjusted) ICT-related patent share in the data and model, which is directly targeted in the calibration. The patent share depends both on the product share and the innovation intensity of the two firm types. This implies that the citation share tends to overstate the product share of the firm type that innovates more frequently. Over the data period, the ICT product share is lower than the ICT citation share because ICT firms innovate relatively more frequently than non-ICT firms.

Figure 5b compares the ICT-related product share in the model with data, which is not directly targeted or mechanically implied by the calibration. the model replicates the overall level and trend of the data. Following (14), the gap between the ICT product share and ICT patent share in the model depends on the relative innovation intensity of ICT and non-ICT firms. It follows that the moment provides a check on parameters related to the relative research intensity of ICT and non-ICT firms (ψ_x^κ, λ^κ, and α^{ict}). The moment also validates the structure of innovations and adoption the model. For example, an alternative innovation structure with single product firms that only innovate on their own product would flatten the

Figure 5: ICT Diffusion in the Data and Model



(a) ICT Diffusion (Targeted)

(b) ICT Product Share (Not Targeted)

Notes: Figure (a) reports the citation-adjusted ICT patent share in the model and data. The model measure is the share of quality-weighted ICT innovations from (14). The data measure is the share of citation-weighted ICT-related patents from Section 3. Figure (b) reports the ICT product share for the model and data. The product share in the model is $\mathcal{D}^{ict}(t)$. The product share in the data is calculated as the patent-weighted share of sectors (US three-digit class) with a depreciated patent stock of ICT-related patents above 25% of the patent stock (details in Appendix B.3).

ICT-product share over time since there would be no scope for ICT innovations to convert non-ICT products.

Table 7: Other Model and Data Moments

Description of Moment	Data	Model
R&D to Sales (%)	5.7	6.0
New ICT Patenter Share (%)	14.7	12.1
Final BGP Growth (%)	-	2.72

Notes: Model and data moments are calculated as mean values over the period 1980 to 2000 with the exception of the Initial BGP Growth and the ICT Patent Share in 1999. The model moment for New ICT Patenter Share is calculated using simulation data.

Table 7 reports moments that are not directly targeted in the calibration. The first moment reported is R&D to sales. The data moment is constructed as the median value for firms that report R&D, excluding firms that higher R&D than sales. The model moment is close to the data and consistent with values reported in the literature, such as Acemoglu et al. (2015) who report values between 5.9% and 8.6%. The second moment shows that the model replicates the share of ICT patents that are issued by firms that have not issued ICT innovations in the past.[27] ICT innovations can be divided into innovations by new ICT innovators (recent adopters) or by previous ICT innovators. The moment can be viewed as a check on the model's ability to replicate the overall number of adopters, which is something not directly targeted in the calibration.

The final moment reports that the predicted growth rate for the long-run BGP, in which ICT becomes fully diffused, is 37% higher than the initial growth rate. The model does not account for the secular decline in research productivity found by Bloom et al. (2020), which would dominate the predicted increase in growth.

The quantitative model makes several additional predictions about ICT and non-ICT firms over the transition path that can be tested in the data. Appendix C tests the prediction in Figures 3 and 4. Using patent values from Kogan et al. (2017), the relative value of ICT patents was highest at the start of the sample and declined over the period (Appendix C.2), consistent with the convergence in values $V_1^\kappa(\lambda^\kappa)$ in Figure 3a. The innovation frequency gap between ICT and non-ICT firms has also increased substantially over time (Appendix C.3), consistent with the widening gap in innovation rates z_x^κ in Figures 3b and 4a. Finally, ICT firms are less likely than non-ICT to exit but these differences decline over time (Appendix C.4), consistent with the creative destruction rates δ^κ in Figure 4b.

5 Quantitative Results

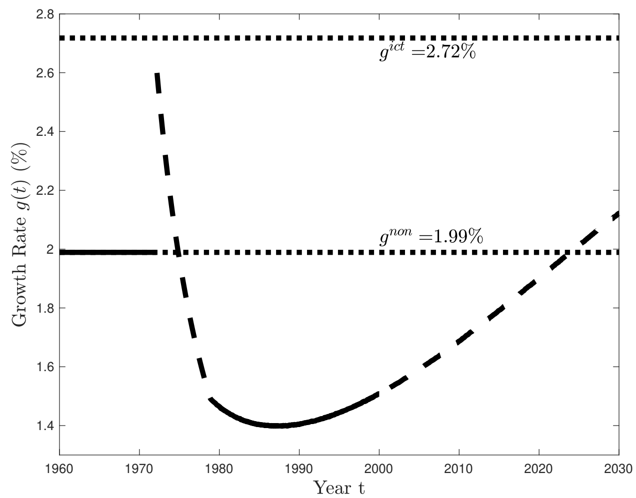
I use the calibrated model to examine the consequences of ICT diffusion for aggregate productivity growth. The calibrated model predicts how aggregate productivity growth is affected by ICT diffusion through changes in firm-level R&D allocations and knowledge spillovers over the transition path. I then examine the drivers of ICT diffusion to help shed light on the relatively slow diffusion. In the final part of the section, I implement two simple policy experiments that illustrate the dynamic interactions of firms over the transition path.

5.1 Productivity Growth

Figure 6 reports the path of output growth over the transition path. In the calibration, only the initial BGP growth rate ($g^{non} = 1.99\%$) is targeted whereas the long-run and transition values are implied by model parameters. In the long run, the growth rate ($g^{ict} = 2.72\%$) rises by around 37% because of higher returns to ICT innovations ($\lambda^{ict} > \lambda^{non}$). However, the increase only follows after a prolonged growth slowdown over the transition period.

[27]The model moment is constructed using simulation data for an initial set of 10,000 firms, where the overall number of firms varies over the simulation. For ease of calculation, the new ICT patenters are taken to be ICT firms in their first year of adoption.

Figure 6: Growth Over the Transition Path



Notes: The figure reports output growth over the transition path. g^{non} and g^{ict} are the growth rates on the non-ICT and ICT balanced growth paths. The solid line corresponds to the years 1980 to 2000.

The model generates a decline and recovery in productivity growth following the introduction of ICT, consistent with empirical evidence on general purpose technologies (e.g., Jovanovic and Rousseau, 2005). Aggregate productivity initially spikes because the introduction of ICT leas to more R&D investment in adoption and non-ICT exploration R&D, where the latter is driven by an increase in the option value of non-ICT firms. Aggregate productivity growth falls quickly in early periods and from 1980 to 2000, the average growth rate is 72% of its initial value on the non-ICT balanced growth path. Growth eventually recovers reaching its initial value in the 2020s and converging slowly to the ICT balanced growth rate in the long run.[28]

The growth slowdown is much longer lived than other channels that explain the growth decline following ICT's introduction (e.g., Hornstein and Krusell, 1996; Greenwood and Yorukoglu, 1997; Yorukoglu, 1998). The length of the growth slowdown stems from the relatively slow diffusion of ICT across innovations in the data (Figure 1). This implies that the model is more relevant for the secular slowdown in aggregate productivity growth (e.g., Aghion et al., 2019), rather than only explaining the isolated decline in aggregate productivity growth in the 1970s and 1980s.[29] The results also point to a recovery in productivity

[28]ICT diffusion has a very long tail in the calibrated model, where the ICT product share reaches 90% by 2035, 99% by 2072, and 99.9% after 2100.

[29]The combination of the results in Figure 6 and other channels explaining the initial drop and recovery in growth—for example, learning-by-doing with ICT equipment as in Greenwood and Yorukoglu (1997)—would lead to growth dynamics similar to those documented in the data. The growth dynamics can be superimposed since the model is unrelated to ICT equipment. The addition of learning-by-doing would imply a sharper decline following the introduction of ICT followed by a growth acceleration in the early 1990s as firms and

growth as ICT becomes fully diffused. However, the recovery in Figure 6 does not overtake the initial growth rate until the 2020s and neglects other channels weighing on growth and research productivity (e.g., Akcigit and Ates, 2019; Peters and Walsh, 2021; Ayerst, 2020b).

The growth slowdown following the introduction of ICT is closely related to mechanisms proposed by De Ridder (2020) and Aghion et al. (2021). In both papers and in Figure 6, aggregate productivity growth initially jumps as some advantaged firms enter the market and then slows as the disadvantaged firms reduce R&D investment. There are two main features in my model that are not present in De Ridder (2020) or Aghion et al. (2021). First, firms endogenously transition from non-ICT to ICT firms over time. This leads to the growth slowdown in Figure 6 being temporary as eventually there are no disadvantaged non-ICT firms. That said, the recovery is gradual and so the growth slowdown remains quantitatively relevant. Second, I use the model structure and the data to inform the introduction of ICT in the data.[30] This leads to the timing of the spike in growth in Figure 6 occurring before what is found by De Ridder (2020) and Aghion et al. (2021).

To understand the decline in growth over the transition path, I decompose quality growth $g(t)$ into three channels based on the associated R&D activity:

$$g(t) = \underbrace{\mathcal{D}^{non}(t) z_x^{non}(t) \left[\mathcal{D}^{non}(t) \ln \lambda^{non}\right] + \mathcal{D}^{ict}(t) z_x^{ict}(t) \left[\mathcal{D}^{ict}(t) \ln \lambda^{ict} + \alpha^{ict} \mathcal{D}^{non}(t) \ln \lambda^{ict} \eta^{ict}\right]}_{\text{Exploration R\&D (=79.1\%)}}$$
$$+ \underbrace{\mathcal{D}^{non}(t) z_a(t) \ln \eta^{ict}}_{\text{Adoption R\&D (=2.1\%)}} + \underbrace{z_e(t) \ln \left(\mathcal{D}^{non}(t)\lambda^{non} + \mathcal{D}^{ict}(t)\lambda^{ict}\right)}_{\text{Entry R\&D (=18.9\%)}}, \tag{15}$$

where the numbers indicate the contribution of each type of R&D to productivity growth over the 1980-2000 period.

Exploration R&D is the main driver of aggregate growth dynamics over the transition path. Exploration R&D follows a U-shaped pattern driven by a combination of declining non-ICT exploration R&D, rising ICT exploration R&D, and an increasing share of ICT products. The bottom of the curve is the point at which the ICT product share is enough to discourage non-ICT exploration R&D but not enough that ICT exploration R&D can fully compensate for the decline. Before (after) this point, increasing the share of ICT products leads to a larger (smaller) decline in non-ICT exploration R&D than the rise in ICT exploration R&D.

workers learn to use ICT equipment and then a second decline as the gains from learning-by-doing spillovers fade. This dual slowdown matches growth dynamics in the data (see, for example, Aghion et al., 2019) that could not be explained by learning-by-doing alone.

[30] This is not to say that the dates used by De Ridder (2020) or Aghion et al. (2021) are incorrect. While the overarching mechanisms are similar, the exact mechanisms are different.

The other two channels account for a small share of growth. Adoption R&D accounts for a negligible share of growth. This is informed by the slow ICT diffusion in the data, which the model interprets as adoption being costly relative to its benefits. Entry R&D makes up the remainder of growth accounting for around one-sixth of total growth over the period, similar to the value found by Akcigit and Kerr (2018). Entry R&D increases early in the transition as the value of ICT firms is initially high and the option value from adopting ICT increases firm value of non-ICT firms but eventually declines as firm values begin to decline as ICT becomes diffused.[31]

5.2 Drivers of Diffusion

The decline in growth depends on the share of ICT products $\mathcal{D}^{ict}(t)$ over the transition path. Understanding what drives diffusion is then key to understanding the growth slowdown and how policy can address it. Diffusion is composed of two channels:

$$\frac{d\mathcal{D}^{ict}(t)}{dt} = \underbrace{\alpha^{ict} z_x^{ict}(t)\mathcal{D}^{ict}(t)\mathcal{D}^{non}(t)}_{\text{Expansion} \quad (=67.3\%)} + \underbrace{\mathcal{D}^{non}(t) z_a(t)}_{\text{Adoption} \quad (=32.8\%)}, \tag{16}$$

where the numbers listed indicate the channel's contribution over the entire transition path.

The expansion channel measures ICT diffusion through ICT firm exploration R&D and accounts for two-thirds of cumulative ICT diffusion. The importance of the expansion channel increases over the transition path driven by the combination of rising ICT exploration R&D (Figure 3b) and the rising ICT product share (Figure 5a). The expansion channel is self-reinforcing in that ICT exploration R&D increases the share of ICT products leading to higher knowledge spillovers (see Figure 4a) and more profitable exploration R&D. In the data, the expansion channel is disciplined by the relative innovation rate of adopters (Table 4) and the relative frequency with which ICT patents cite non-ICT patents.[32]

The adoption channel measures ICT diffusion through non-ICT firms adopting ICT and accounts for around one-third of cumulative ICT diffusion. In contrast with the expansion

[31] Qualitatively, the rise and fall in entry R&D is also consistent with two stylized facts. First, the creation of new general purpose technologies are associated with a boom in entry activity (e.g. Jovanovic and Rousseau, 2005) as the high market value of new technologies encourages entry. Second, the decline in entry is consistent with the decline in business dynamism found in the US (see Decker et al., 2014).

[32] While it is clear that if $\alpha^{ict} = 0$ then the expansion channel is zero, changes in α^{ict} would have an offsetting effect on the expansion channel if the model is recalibrated. Specifically, lower values of α^{ict} increases the value of non-ICT firms because of lower competition spillovers, which increases non-ICT exploration innovation $z_x^{non}(t)$. The calibration then requires a compensating decrease in the cost of ICT exploration innovation ψ_x^{ict} in order to match the targeted relative innovation rate. The intuition is similar to the model extension adding non-ICT spillovers α^{non} in Appendix D.4.

channel, the adoption channel is self-cannibalizing in that adoption discourages future adoption since it increases competition spillovers (see $\delta^{ict}(t)$ Figure 4b).[33] The relatively lackluster contribution of the adoption channel is determined by the slow ICT diffusion in the data.

The cumulative contributions are potentially misleading since the adoption channel accounts for most of the early diffusion. In early periods, ICT diffusion can only be driven by the adoption channel since there are no ICT firms to invest in exploration R&D. While the expansion channel overtakes the adoption channel relatively quickly, this is only possible because the economy reaches a critical mass of ICT products to support ICT exploration R&D. As a result, delaying adoption R&D delays ICT diffusion more than is suggested by the cumulative share of the adoption channel. In the extreme case, eliminating adoption R&D outright would stop diffusion. Consequently, it is also important to account for the interactions between the two channels, which is explored in the next part of the section.

The majority of the existing literature focuses on adoption as the sole channel of diffusion (e.g. Parente and Prescott, 1994; Yorukoglu, 1998; Comin and Hobijn, 2010; Ayerst, 2020a). I find that this channel accounts for a relatively small share of cumulative diffusion. This can be especially misleading for policy because models with only adoption lead to policy responses that target non-adopters whereas the results in (16) suggests that policy should also consider previous adopters. This is explored next.

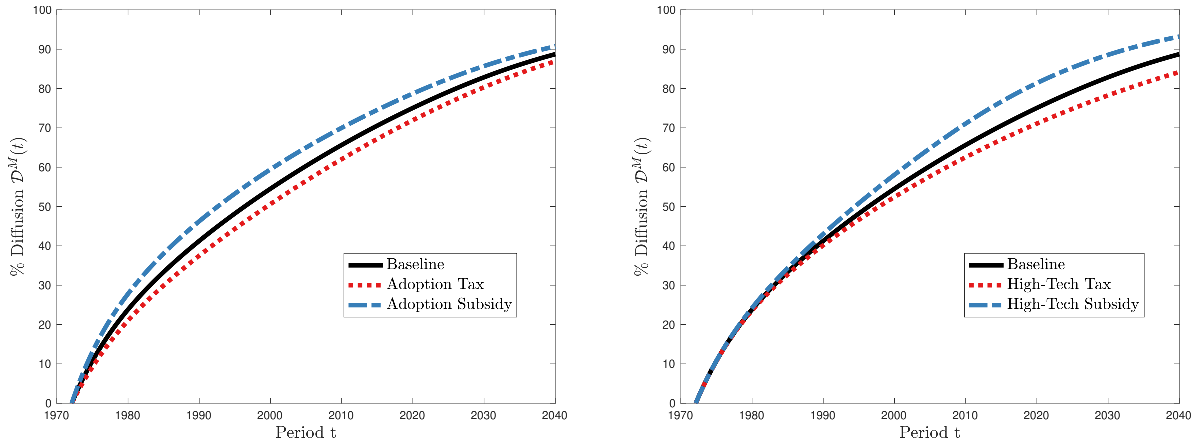
5.3 Policy Experiments

Adoption R&D creates spillovers to other firms through its impact on the distribution of ICT-related products. Along with the knowledge and competition spillovers from exploration R&D, there is a role for policy to reduce inefficiencies and improve welfare. Supporting diffusion can shorten the low-growth transition period and bring the economy to the high-growth BGP earlier. However, policy is inherently complex because of the forward-looking dynamics and interconnectedness of the environment. Given this, I explore two simple policy experiments to provide insights into policy design and highlight the model mechanisms.

The first set of policy experiments implements a simple subsidy or tax on adoption R&D, such that the net adoption R&D cost becomes $(1-\tau_a)w_s(t)ns_a(z_a)$ for a non-ICT firm with n product lines. The second set of policy experiments implements a high-tech subsidy on exploration R&D by ICT firms, such that the net exploration R&D cost becomes $(1-\tau^{ict})w_s(t)ns_x(z_x)$ for an ICT firm with n product lines. I focus on a 25% subsidy and tax for each type of policy. Figure 7 reports diffusion in the policy experiments and Table 8 summarizes the welfare and growth results from the policy experiments.

[33] Appendix D.3 decomposes adoption R&D into the direct and indirect benefits and reports the contribution of the adoption and expansion channels to ICT diffusion by period.

Figure 7: ICT Diffusion under Policy Experiments



(a) Adoption Policy (b) High-Tech Policy

Notes: The figure reports diffusion $\mathcal{D}^{ict}(t)$ for a 25% tax or subsidy to adoption R&D costs, Panel (a), or exploration R&D costs of ICT firms, Panel (b).

Both the adoption and high-tech subsidies have the intended effects of increasing diffusion and shortening the period of slow growth. The adoption subsidy increases early diffusion but has relatively no effect on later diffusion. The adoption subsidy creates a negative feedback in which the rise in early diffusion leads to higher competition for ICT firms lowering the incentives for firms to invest in adoption. In contrast, the high-tech subsidy has a fanning out effect on diffusion. The high-tech subsidy is self-reinforcing in that higher ICT exploration R&D increases the share of ICT products making future exploration R&D more profitable.

Table 8 summarizes the change in welfare from the two subsidies as well as the corresponding taxes. Welfare is reported at a short horizon from the introduction of ICT t_0 until 2010 and an infinitely forward looking (long) horizon.[34] At the short horizon, taxing adoption R&D is beneficial since it delays the drop in aggregate productivity growth, raising consumption. At the long horizon, subsidizing diffusion becomes more beneficial because it leads to a shorter low growth period. Welfare becomes relatively more elastic to the high-tech policy at longer horizons since the high-tech subsidy impacts diffusion in later periods. In contrast, the magnitude of the adoption policy on welfare is relatively consistent at both horizons since it has the largest effect early periods.

The final row of the table reports the year that the transition growth rate overtakes the initial balanced growth path rate, relative to the baseline calibration. In the baseline

[34]The choice of 2010 is arbitrary for the comparison of welfare. I choose 2010 since it is a mid-point between the lowest growth rate on the transition path (around 1990) and the point where growth overtakes its initial value (around 2030). The main intuition of the welfare comparison holds for different constructions of the time horizons.

Table 8: Summary of Policy Experiments

	Adoption Policy		High-Tech Policy	
	Tax	Subsidy	Tax	Subsidy
Welfare Short Horizon $[t_0, 2010]$ (%)	0.15	-0.14	-0.19	0.20
Welfare Long Horizon $[t_0, \infty)$ (%)	-1.10	1.58	-6.60	11.94
Year $g(t) \geq g^{non}$ (Change in Years)	3	-5	20	-13

Notes: Welfare values report the change in welfare relative to the baseline economy for a 25% tax or subsidy to adoption R&D costs or exploration R&D costs of ICT firms. Welfare is reported as a consumption equivalent (defined in Appendix A.2). The short horizon calculated the change in welfare from the introduction of ICT (t_0) until the year corresponding to 2010. The long horizon calculates the change in welfare over the entire transition period. Year $g(t) \geq g^{non}$ reports the first year (after the initial drop) that growth surpasses the non-ICT balanced growth path growth rate relative to the benchmark economy. For example, -5 years indicates the the adoption subsidy reaches a growth rate greater than g^{non} five years earlier than in the benchmark economy.

calibration, growth overtakes its initial value in the 2020s (Figure 6). The year changes only slightly for both the adoption tax and subsidy reflecting that diffusion in Figure 7 is relatively unchanged. Since the high-tech policies have a large impact on diffusion, it follows that the high-tech policies also have a large impact on the year that growth overtakes its initial value. For the subsidy, this occurs 13 years earlier than in the baseline economy.

The experiments have two broad insights for policy design related to new general purpose technologies (GPTs). First, policies that focus on shorter horizons may incorrectly assess welfare. Intuitively, shorter horizons place a greater weight on the growth slowdown that follows the introduction of new GPTs motivating delayed diffusion. At longer horizons, this is damaging because it causes the economy to remain in the low growth transition for longer. Second, policies should focus on current users of new GPTs rather than non-adopters. Focusing on non-adopters can be effective in early periods of the transition but creates a negative feedback that weakens the policy impact over time. Additionally, the window of opportunity for policymakers is relatively short and requires that policymakers can correctly identify new GPTs. Focusing on previous GPT adopters tends to be self-reinforcing and is more effective in later periods when policymakers are more likely to know the value of new GPTs.

6 Conclusion

In this paper, I examine the diffusion of new general purpose technologies (GPTs) and the consequences for innovation and growth. This paper contributes along three dimensions.

First, I extend the standard Klette and Kortum (2004) framework to incorporate the adoption and diffusion of ICT. The model highlights that the R&D allocation and intensity of firms changes over the transition path as ICT diffusion dictates the relative knowledge and competition spillovers firms face. Second, I construct a measure of ICT-related innovation using administrative data on patents and citation networks, capturing the idea that the applications of ICT spurs innovations in other fields. I find that ICT adopter are relatively better innovators in both patent quality and frequency and that adopters experience a sharp increase in productivity growth in the years immediately following adoption. Third, I use the calibrated model to examine the consequences and drivers of ICT diffusion. I find that most of the diffusion is driven by the expansion of current users rather than new adopters. Additionally, I provide a new explanation of growth slowdowns following the introduction of GPTs in which non-users lower R&D investment as the new GPT becomes more prevalent.

The structure of the model is flexible and could be applied to new general purpose technologies (e.g., artificial intelligence) or other technologies (e.g., green technologies). An interesting application would be to use the model to examine the introduction of artificial intelligence. Another path for future work would be to extend the framework to examine cross-country differences in ICT diffusion and technology gaps documented by Comin and Hobijn (2010) and Comin and Mestieri (2018).

References

Abrams, D. S., Akcigit, U., and Greenan, J. (2018). Patent value and citations: Creative destruction or strategic disruption. Technical report.

Acemoglu, D., Akcigit, U., Alp, H., Bloom, N., and Kerr, W. R. (2015). Innovation, reallocation and growth. *American Economic Review*.

Acemoglu, D., Akcigit, U., Hanley, D., and Kerr, W. R. (2016a). Transition to clean technology. *Journal of Political Economy*, pages 52–104.

Acemoglu, D., Akcigit, U., and Kerr, W. (2016b). Innovation network. *Proceedings of the National Academy of Sciences*, 113(41):11483–11488.

Acemoglu, D., Autor, D., Dorn, D., Hanson, G. H., and Price, B. (2014). Return of the solow paradox? it, productivity, and employment in us manufacturing. *American Economic Review*, 104(5):394–399.

Aghion, P., Bergeaud, A., Boppart, T., Klenow, P. J., and Li, H. (2019). Missing growth from creative destruction. *American Economic Review*, 109:2795–2822.

Aghion, P., Bergeaud, A., Boppart, T., Klenow, P. J., and Li, H. (2021). A theory of falling growth and rising rents. Technical report.

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

Akcigit, U. and Ates, S. T. (2019). What happened to us business dynamism. Technical report.

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

Ayerst, S. (2020a). Distorted technology adoption. Technical report.

Ayerst, S. (2020b). Innovator heterogeneity, r&d misallocation and the productivity growth slowdown. Technical report.

Ayerst, S., Ibrahim, F., MacKenzie, G., and Rachapalli, S. (2022). Trade and diffusion of embodied technology: An empirical analysis. Technical report.

Baslandze, S. (2018). The role of the it revolution in knowledge diffusion, innovation and reallocation. Technical report, Working Paper.

Benhabib, J., Perla, J., and Tonetti, C. (2018). Reconciling models of diffusion and innovation: A theoryof the productivity distribution and technology frontier. Technical report, Working Paper.

Berkes, E., Manysheva, K., and Mestieri, M. (2022). Global innovation spillovers and productivity evidence from 100 years of world patent data. Technical report.

Bloom, N., Garicano, L., Sadun, R., and Van Reenen, J. (2014). The distinct effects of information technology and commnuication technology on firm organization. *Management Science*, 60(12):2859–2885.

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

Bresnahan, T. F. and Trajtenberg, M. (1995). General purpose technologies 'engines of growth'? *Journal of Econometrics*, 65(1):83–108.

Brynjolfsson, E. and Hitt, L. M. (2000). Beyond computation: Information technology, organizational transformation and business performance. *Journal of Economic Perspectives*, 14(4):23–48.

Brynjolfsson, E., Rock, D., and Syverson, C. (2021). The productivity j-curve: How intangible complement general purpose technologies. *American Economic Journal: Macroeconomics*, 13(1):333–372.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Caselli, F. and Coleman, W. J. (2001). Cross-country technology diffusion: The case of computers. *American Economic Review*, 91(2):328–335.

Chen, C. (2018). Technology adoption, capital deepening, and international productivity differences. Technical report, Working Paper.

Comin, D. and Hobijn, B. (2004). Cross-country technological adoption: Making the theories face the facts. *Journal of Monetary Economics*, 51:39–83.

Comin, D. and Hobijn, B. (2010). An exploration of technology diffusion. *American Economic Review*, 100(5):2031–2059.

Comin, D. and Mestieri, M. (2018). If technology has arrived everywhere, why has income diverged? *American Economic Journal: Macroeconomics*, 10(3):137–178.

De Loecker, J., Eeckhout, J., and Unger, G. (2018). The rise of market power and the macroeconomic implications. *Quarterly Journal of Economics*, 135(2):561–644.

De Ridder, M. (2020). Market power and innovation in the intangible economy. Technical report, Working Paper.

Decker, R., Haltiwanger, J., Jarmin, R. S., and Miranda, J. (2014). The secular decline of business dynamisms in the united states. Technical report, Working Paper.

DeStefano, T., Kneller, R., and Timmis, J. (2018). Broadband infrastructure, ict use and firm performance: Evidence for uk firms. *Journal of Economic Behavior and Organization*, 155(November):110–139.

Garcia-Macia, D., Hsieh, C.-T., and Klenow, P. J. (2016). How destructive is innovation? Technical report, NBER Working Paper.

Gordon, R. J. (2000). Does the "new economy" measure up to the great inventions of the past? *Journal of Economic Perspectives*, 14(4):49–74.

Greenwood, J. and Yorukoglu, M. (1997). 1974. *Carnegie-Rochester Conference Series on Public Policy*, 46:49–95.

Griliches, Z. (1957). Hybrid corn: An exploration in the economies of technological change. *Econometrica*, 25(4):501–522.

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

Hall, B., Jaffe, A., and Trajtenberg, M. (2001). The nber patent citation data file: Lessons insight and methodological tool. Technical report.

Hall, B., Jaffe, A., and Trajtenberg, M. (2005). Market value and patent citations. *RAND Journal of Economics*.

Hornstein, A. and Krusell, P. (1996). Can technology improvements cause productivity slowdowns? *NBER Macroeconomics Annual 1996*, 11.

Jovanovic, B. and MacDonald, G. M. (1994). Competitive diffusion. *Journal of Political Economy*, 102(1):24–52.

Jovanovic, B. and Rousseau, P. L. (2005). General purpose technologies. *Handbook of Economic Growth*, pages 1181–1223.

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

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

Lashkari, D., Bauer, A., and Boussard, J. (2021). Information technology and returns to scale. Technical report, Working Paper.

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

Liu, E. and Ma, S. (2022). Innovation networks and r&d allocation. Technical report.

Lucas, R. E. J. and Moll, B. (2014). Knowledge growth and the allocation of time. *Journal of Political Economy*, 122(1):1–51.

Oliner, S. D. and Sichel, D. E. (2000). The resurgence of growth in the late 1990s: Is information technology the story? *Journal of Economic Perspectives*, 14(4):3–22.

Parente, S. L. and Prescott, E. C. (1994). Barriers to technology adoption and development. *Journal of Political Economy*, 102(2):298–321.

Perla, J. and Tonetti, C. (2014). Equilibrium imitation and growth. *Journal of Political Economy*, 122(1):52–76.

Peters, M. (2016). Heterogeneous markups, growth and endogenous misallocation. Technical report.

Peters, M. and Walsh, C. (2021). Population growth and firm dynamics. Technical report.

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

Solow, R. M. (1987). Review of *Manufacturing Matters: The Myth of the Post-Industrial Economy.*

Yorukoglu, M. (1998). The information technology productivity paradox. *Review of Economic Dynamics*, 1(2):551–592.

Online Appendix (Not for Publication)

A Model Appendix

A.1 Description of Equilibrium Production

The final consumption good is taken to be the numeraire. The household's problem implies the Euler equation $\dot{C}(t)/C(t) = r(t) - \rho$.

The final good producer demands $Y(t)$ units of each good j. Because varieties of good j are quality-adjusted perfect substitutes, the final good producer only purchases good j from the firm that offers the lowest price-per-quality $p_{j_f}/\eta_{j_f} q_{j_f}$. This implies that the quality leader $f = L$ sets the limit price given the quality of the follower $f = F$ such that $p_j(t) = \mu_j w(t)$ where the markup is $\mu_j = \eta_{j_L} q_{j_L}/\eta_{j_F} q_{j_F}$ and $w(t)$ is the wage rate. The labor and output allocated to good j are $k_j(t) = \ell_j(t) = Y(t)/(\mu_j w(t))$.

Markup distribution. Prices depend on the quality leadership of the current producer. Markups take one of four values in equilibrium $\{\lambda^{non}, \lambda^{ict}, \eta^{ict}\lambda^{ict}, \eta^{ict}\lambda^{non}\}$. The first two values correspond non-ICT and ICT firms innovating on goods with the same technology. The third value occurs when ICT firms successfully innovate on goods produced by non-ICT firms. The fourth value occurs when non-ICT firms adopt ICT increasing the markup from λ^{non} to $\eta^{ict}\lambda^{non}$. The distribution $\Phi(\mu, t)$ of markups is characterized by the law of motion:

$$\dot{\Phi}(\mu, t) = \begin{cases} \left(z_x^{non}(t)\mathcal{D}^{non}(t) + z_e(t) - z_a(t)\right)\mathcal{D}^{non}(t) - \delta^{non}(t)\Phi(\lambda^{non}, t) & \text{for } \mu = \lambda^{non} \\ (z_x^{ict}(t)\mathcal{D}^{ict}(t) + z_e(t))\mathcal{D}^{ict}(t) - \delta^{ict}(t)\Phi(\lambda^{ict}, t) & \text{for } \mu = \lambda^{ict} \\ \alpha^{ict} z_x^{ict}(t)\mathcal{D}^{ict}(t)\mathcal{D}^{non}(t) - \delta^{ict}(t)\Phi(\eta^{ict}\lambda^{ict}, t) & \text{for } \mu = \eta^{ict}\lambda^{ict} \\ z_a(t)\mathcal{D}^{non}(t) - \delta^{ict}(t)\Phi(\eta^{ict}\lambda^{non}, t) & \text{for } \mu = \eta^{ict}\lambda^{non} \end{cases}. \tag{17}$$

In what follows, I refer to goods by the markup charged by the current producer, rather than the index j, since this describes the price and labor and output allocations of the good.

Allocations, prices, and output. Labor market clearing requires that $1 = \int_j \ell_j dj$. The equilibrium wage rate that clears the labour market is equal to:

$$w(t) = \bar{q}(t) \sum\nolimits_{\mu} \mu^{-1}\Phi(\mu, t), \tag{18}$$

where $\bar{q}(t) = \exp \int_0^1 \ln q_j dj$ is the average quality of goods.

Given the wage rate, the price, output, and labor allocated to good j can be written in terms of the markup as:

$$p(\mu,t)=\mu w(t), \qquad k(\mu,t)=\ell(\mu,t)=\frac{\mu^{-1}}{\sum_{\mu}\mu^{-1}\Phi(\mu,t)}. \tag{19}$$

Intuitively, Cobb-Douglas preferences imply that expenditure on each good is the same and so there is less production of goods with higher prices.

The goods market clearing requires that all intermediate inputs are used in the production of the final consumption good and that the final consumption good is used for consumption or R&D, $Y(t)=C(t)+\psi_e z_e(t)+\sum_{\kappa\in\{T,M\}}\mathcal{D}^{\kappa}(t)s_x(z_x,t)+\mathcal{D}^{non}(t)s_a(z_a,t)$. Aggregate output depends on both the distribution of markups across goods and the average quality of goods:

$$Y(t)=\chi(t)\bar{q}(t), \tag{20}$$

where $\chi(t)=\exp\left(\sum_{\iota}\ln\mu_{\iota}^{-1}\Phi(\mu_{\iota},t)\right)/\sum_{\iota}\mu_{\iota}\Phi(\mu_{\iota},t)\leq 1$ describes misallocation from dispersion in markups. This output loss from markup dispersion is examined by Peters (2016) while the loss in growth from persistent markup differences is examined by Ayerst (2020b).

A.2 Additional Model Results and Outcomes

Welfare Welfare over a period $[\underline{t},\bar{t}]$ is calculated as:

$$W=\int_{\underline{t}}^{\bar{t}}e^{-\rho\hat{t}}\log(C(\hat{t}))d\hat{t}.$$

The policy experiments compare the change in welfare using the consumption equivalent. Let $\tilde{C}(t)$ be the series of consumption under the proposed policy and $C(t)$ be the consumption under the baseline scenario. The consumption equivalent ξ solves:

$$\int_{\underline{t}}^{\bar{t}}e^{-\rho\hat{t}}\log((1+\xi)C(\hat{t}))d\hat{t}=\int_{\underline{t}}^{\bar{t}}e^{-\rho\hat{t}}\log(\tilde{C}(\hat{t}))d\hat{t}. \tag{21}$$

Balanced Growth Path Equilibrium Proposition 2 characterizes the balanced growth path equilibrium.

Proposition 2. *If $z_e>0$, the $\kappa\in\{non,ict\}$ technology balanced growth path (BGP$^{\kappa}$) is characterized by the innovation rates*

$$z_x^{\kappa}=\left(\frac{\psi_e}{\zeta_x\psi_x^{\kappa}}\right)^{\frac{1}{\zeta_x-1}} \qquad z_e=\frac{1}{\psi_e}\left[(1-(\lambda^{\kappa})^{-1})+\psi_x^{\kappa}(z_x^{\kappa})^{\zeta_x}\right]-\rho-z_x^{\kappa}$$

and if $z_e = 0$ then $z_x^\kappa = [(B^\kappa + (1-(\lambda^\kappa)^{-1})v^\kappa)/\zeta\psi_x^\kappa]^{1/(\zeta_x - 1)}$ is implicitly defined along with the value function below. The creative destruction and growth rates

$$\delta = z_e + z_x^\kappa \qquad\qquad g = \delta \ln \lambda^\kappa;$$

and the value function

$$V_n^\kappa = Y(t)\left[B^\kappa + v^\kappa(1-(\lambda^\kappa)^{-1})\right]n$$

where

$$v^\kappa = \frac{1}{\rho + z_x^\kappa + z_e} \qquad\qquad B^\kappa = \frac{\psi_x^\kappa(\zeta_x - 1)}{\rho + z_x^\kappa + z_e}(z_x^\kappa)^{\zeta_x}.$$

Proof The balanced growth path (BGP) equilibrium requires that only one technology κ can be used by firms. This follows from the law of motion for the ICT product share:

$$\frac{d\mathcal{D}^{ict}(t)}{dt} = z_a(t)\mathcal{D}^{non}(t) + \alpha^{ict} z_x^{ict}(t)\mathcal{D}^{non}(t)\mathcal{D}^{ict}(t)$$

which can only equal 0 under one of three cases: (1) $\alpha^{ict} z_x^{ict}(t) = z_a(t) = 0$ with $\mathcal{D}^{non}(t), \mathcal{D}^{ict}(t) > 0$; (2) $\mathcal{D}^{non}(t) = 0$; or (3) $z_a(t) = 0$ and $\mathcal{D}^{ict}(t) = 0$. The first case is ruled out by assumption that there is some adoption $z_a(t) > 0$ for some t and that $\alpha^{ict} > 0$. The exploration R&D cost function implies that ICT firms always choose positive $z_x^{ict}(t)$ at non-infinite R&D costs. The second and third cases represent the non-ICT and ICT BGPs.

To solve the BGP values, I use the guess and verify method with guess:

$$V_n^\kappa = Y\left[B^\kappa + v^\kappa(1-(\lambda^\kappa)^{-1})\right]n.$$

Plugging the guess into the value function and solving for z_x^κ implies that $z_x^\kappa = \left[V_1^\kappa/\zeta_x\psi_x^\kappa\right]^{1/\zeta_x - 1}$. If entry is positive then the entry condition implies that $V_1^\kappa = \psi_e Y$ and innovation intensities are:

$$z_x^\kappa = \left(\frac{\psi_e}{\zeta_x\psi_x^\kappa}\right)^{\frac{1}{\zeta_x - 1}} \qquad\qquad z_e = \frac{1}{\psi_e}\left[(1-(\lambda^\kappa)^{-1}) + \psi_x^\kappa(z_x^\kappa)^{\zeta_x}\right] - \rho - z_x^\kappa$$

If entry is zero, then the value function takes the same form and $z_x^\kappa = [(B^\kappa + (1-(\lambda^\kappa)^{-1})v^\kappa)/\zeta\psi_x^\kappa]^{1/(\zeta_x - 1)}$ follows from solving the firm's problem. The growth rate is $g = [z_x^\kappa + z_e]\ln\lambda^\kappa = \delta\ln\lambda^\kappa$ with the values of z_x^κ and z_e given above.

Finally, substituting the optimal value of R&D intensity z_x and the creative destruction rate δ into the value function, confirms the guess for values:

$$v^{\kappa} = \frac{1}{\rho + z_x^{\kappa} + z_e} \qquad\qquad B^{\kappa} = \frac{\psi_x^{\kappa}(\zeta_x - 1)}{\rho + z_x^{\kappa} + z_e}(z_x^{\kappa})^{\zeta_x}.$$

A.3 Transition Path Solution

Proof of Proposition 1: For brevity, I write the problem of a firm with technology κ to include adoption with the understanding that the benefit of adoption is equal to zero for ICT firms. The discrete approximation of the value function is

$$\begin{aligned} V_n^{\kappa}([\mu_i], t) = \max_{z_x^{\kappa}, z_a^{\kappa}} & \left[\sum_{i'=1}^{n}(1-\mu_{i'}^{-1})Y(t)\right]\Delta t - n\left[s_x(z_x) + s_a(z_a)\right]Y(t)\Delta t + o(\Delta t) \\ & + \exp\{-r(t)\Delta t\}\left[\begin{array}{c} \sum_{i'=1}^{n}\left[\left(\delta^{\kappa}(t)\Delta t\right)V_n^{\kappa}([\mu_i]/\{\mu_{i'}\}, t+\Delta t)\right] \\ + (z_x^{\kappa} n\Delta t)\left[\mathcal{D}^{\kappa}(t)V_{n+1}^{\kappa}([\mu_i]\cup\{\lambda^{\kappa}\}, t+\Delta t) + \alpha^{\kappa}\mathcal{D}^{-\kappa}(t)V_{n+1}^{\kappa}([\mu_i]\cup\{\eta^{\kappa}\lambda^{\kappa}\}, t+\Delta t)\right] \\ + (z_a^{\kappa} n\Delta t)\,V_n^{ict}([\eta^{ict}\mu_i], t+\Delta t) + (1 - (\delta^{\kappa}(t) + z_x^{\kappa} + z_a^{\kappa})\Delta t)V_n^{k}([\mu_i], t+\Delta t) \end{array}\right]. \end{aligned}$$

I solve the value function using the guess and verify method, with guess:

$$V_n^{\kappa}([\mu_i], t) = \sum_{i=1}^{n}\left[B^{\kappa}(t) + v^{\kappa}(t)(1-\mu_i^{-1})\right]Y(t).$$

Using the above guess, the implied policy functions are given by:

$$z_x^{\kappa}(t) = \left[e^{-(r(t)-g_Y(t))\Delta t}\frac{\mathcal{D}^{\kappa}(t)[v^{\kappa}(t)(1-(\lambda^{\kappa})^{-1}) + B^{\kappa}(t)] + \alpha^{\kappa}\mathcal{D}^{-\kappa}(t)\left[v(t)\left(1-(\eta^{\kappa}\lambda^{\kappa})^{-1}\right) + B^{\kappa}(t)\right]}{\zeta_x\psi_x^{\kappa}}\right]^{\frac{1}{\zeta_x - 1}}$$

$$z_a(t) = \Bigg[e^{-(r(t)-g_Y(t))\Delta t}\frac{v^{ict}(t)\frac{1-(\eta^{ict})^{-1}}{\lambda^{non}} + [v^{ict}(t)(1-(\lambda^{non})^{-1}) + B^{ict}(t)] - [v^{non}(t)(1-(\lambda^{non})^{-1}) + B^{non}(t)]}{\psi_a\zeta_a}$$

As $\Delta t \to 0$ the policy functions imply the functions given in Proposition 1. Substituting the policy functions into the firm's problem implies that the value function is described by the equations:

$$\begin{aligned} v^{\kappa}(t) - e^{(g_Y(t)-r(t))\Delta t}v^{\kappa}(t+\Delta t) &= \Delta t - e^{(g_Y(t)-r(t))\Delta t}(\delta^{\kappa}(t)\Delta t)v^{\kappa}(t+\Delta t) + o(\Delta t) \\ B^{\kappa}(t) - e^{(g_Y(t)-r(t))\Delta t}B^{\kappa}(t+\Delta t) &= -\left[\psi_x^{\kappa}(z_x^{\kappa}(t))^{\zeta_x} + \psi_a(z_a(t))^{\zeta_a}\right]\Delta t - e^{(g_Y(t)-r(t))\Delta t}(\delta^{\kappa}(t)\Delta t)B^{\kappa}(t+\Delta t) \\ &\quad + e^{(g_Y(t)-r(t))\Delta t}(z_x^{\kappa}(t)\Delta t)\left[\mathcal{D}^{\kappa}(t)[v^{\kappa}(t)(1-(\lambda^{\kappa})^{-1}) + B^{\kappa}(t)] + \alpha^{\kappa}\mathcal{D}^{-\kappa}(t)\left[v(t)\left(1-(\eta^{\kappa}\lambda^{\kappa})^{-1}\right) + B^{\kappa}(t)\right]\right. \\ &\quad + e^{(g_Y(t)-r(t))\Delta t}(z_a(t)\Delta t)\left[v^{ict}(t)(1-(\eta^{ict}\lambda^{non})^{-1}) + B^{ict}(t) - v^{non}(t)(1-(\lambda^{non})^{-1}) - B^{non}(t)\right] + o(\Delta t \end{aligned}$$

Setting $r(t) = \rho + g_C(t)$ (from the household's Euler equation) and taking the limit $\Delta t \to 0$

implies that the components of the value function are described by the differential equations:

$$
\begin{aligned}
(\rho + (g_C(t) - g_Y(t)) + \delta^\kappa(t))v^\kappa(t) - \dot{v}^\kappa(t) &= 1 \\
(\rho + (g_C(t) - g_Y(t)) + \delta^\kappa(t))B^\kappa(t) - \dot{B}^\kappa(t) &= \psi_x^\kappa(\zeta_x - 1)(z_x^\kappa(t))^{\zeta_x} + \psi_a(\zeta_a - 1)(z_a(t))^{\zeta_a}.
\end{aligned}
$$

A.4 Technology-Specific Spillovers

In this Appendix Section, I discuss a simple extension of the model that endogenizes the cross-technology spillovers. The extended model assumes that innovations do not build perfectly on cross-technology products and fail when the innovating firm is unable to charge a lower price-per-quality than the incumbent producer. I use the extended model to show that the model has the same structural parameters as in the baseline model.

As in the baseline model, the technology-specific quality component for non-ICT firms is normalized $\eta^{non} = 1$ while the ICT quality $\eta^{ict} = \eta$ is drawn from a Pareto distribution with CDF $1 - \eta^{-\gamma}$. When a firm using the κ technology draws a product produced with the other $\kappa' \neq \kappa$ technology they pay a productivity cost described by $b \leq 1$ such that the overall step size becomes λb. In a reduced form, the cost b captures many of the difficulties of moving into new technology markets. The cost b can be thought of as capturing some incompatibilities between producing goods with a different general purpose technologies or applying innovations stemming from the use of ICT to new markets. For example, consider the difficulty of an ICT-using car manufacturer moving into an adjacent non-ICT market (e.g., motorcycles). The application of the ICT-related technology (e.g., electronic control unit) may require additional R&D to be applied to the new good or the production process used by the ICT-using car manufacturer may not be suitable for the newly innovated good, requiring additional costs such as changing machinery or intermediate input suppliers. Brynjolfsson and Hitt (2000) provides several examples of the difficulties faced by specific firms transitioning to ICT. Alternatively, if the innovation is developed from the insights in the production of ICT-using cars (what Klette and Kortum (2004) refer to as *knowledge capital*) then the innovation itself may be incompatible with the targeted good and require additional investment to bridge this gap. The cost could also be thought of as capturing the difficulties of moving to cross-technology markets. For example, a smartphone manufacturer may have the technical capabilities to expand into adjacent non-ICT markets (e.g., watches) but the costs may be prohibitively high or the consumer base too small to justify the benefits. The explanation is more similar to the modeling of $A^\kappa(t)$ in the baseline model where α^κ captures a bias towards same-technology innovations.

The probability that a ICT firm innovates on a non-ICT product is then given by

$$\alpha^{ict} = \Pr[\lambda^{ict}\eta b > 1] = \min\left\{(\lambda^{ict}b)^{\gamma}, 1\right\}.$$

The probability that a non-ICT firm innovates on an ICT product is given by

$$\alpha^{non} = \Pr\left[\lambda^{non}b > \eta\lambda^{ict}\right] = \max\left\{0, 1 - \left(\frac{\lambda^{non}b}{\lambda^{ict}}\right)^{-\gamma}\right\},$$

where the baseline assumption that non-ICT firms do not innovate on ICT products requires parameters to be set such that $b < \lambda^{ict}/\lambda^{non}$. The values of α^{non} and α^{ict} are then implied by the parameters b and γ. In this regard, the cross-technology spillovers α^{κ} in the baseline model nests a structure in which firms require an additional cost to build on cross-technology products.

B Empirical Appendix

B.1 Additional Summary Statistics

Table 9 reports summary statistics on the final set of innovative Compustat firms.

Table 9: Firm Summary Statistics

Statistic	Adopters	Non-Adopters	All
Observations	7,357	15,324	22,681
Initial Type	316	1,284	2,600
Final Type	1,369	1,231	2,600

The final dataset contains information on 22,681 firm-year observations from 2,600 unique firms. The initial and final types report the firm's type for the first and last observations in the dataset. While this does not directly correspond to the beginning and end of the sample, the comparison shows that a large fraction of firms adopt ICT over the sample.

B.2 Distance Cutoff

To estimate the appropriate distance to cutoff ICT-related patents, I perform a variation of the exercise in Section 3 in which I examine firm-level responses to adoption based on different measures of distance. The set of ICT patents is defined as:

$$\text{ICT-Related Patents} = \cup_{d=0}^{\bar{d}} \mathcal{P}_d.$$

The cutoff $\bar{d}$ is the maximum distance from the *Computers & Communication* technology class included in the set of ICT-related patents, which is taken to be $\bar{d} = 3$ in the baseline analysis. Table 10 summarizes the number of new firm observations in each distance category.

Table 10: Observations by Patent Distance

Distance	Observations
0	624
1	165
2	186
3	129
4	85

The table shows that higher distances include relatively fewer observations. A caveat with the values in the table above is that the minimum distance is time independent. That is, increasing the cutoff may lead to some firms adopting earlier because they issue higher distance patents before lower distance patents. This leads to a downward bias of the estimated response at lower cutoffs to the extent that the "true" adoption date is missed.

The decision of the cutoff weighs the benefits of including a broader set of potentially ICT-related patents against the cost of mislabeling patents as being ICT-related. Given this, I set the cutoff $\bar{d}$ based on the response of firms following ICT adoption. The main idea being that the response captures a change at the firm (ICT adoption) implying that smaller response indicates a worse measure of adoption. This favors classifying ICT-related patents as non-ICT patents if at higher distances smaller responses are likely to be observed. I base the cutoff on is the empirical response of labour productivity to adoption estimated as:

$$g_{f,t} = \beta_0 + \sum_{s=0,-1,-2} \beta_{A,s} \text{Adopt}_{f,t+s} + \mathbf{X}_{f,t}\beta_X + \Gamma_s + \Gamma_t + \varepsilon_{f,t}$$

where $\text{Adopt}_{f,t}$ is the year in which firm f's patent portfolio exceeds 10% patents with distance $d \leq \bar{d}$. The same controls and fixed effects as in the main text are included. Table 11 reports the results for different cutoffs $\bar{d}$. The three citation link cutoff performs comparably better than the other specifications and is consequently used as the baseline.

Table 11: Distance Cutoff

	Sales / Emp. Growth				
	(1)	(2)	(3)	(4)	(5)
Adopt t	0.0118	0.0232	0.0460**	0.0500***	0.0336
	(0.0349)	(0.0334)	(0.0225)	(0.0175)	(0.0225)
Adopt t-1	0.0691**	0.0493*	0.0638**	0.0702**	0.0587***
	(0.0327)	(0.0289)	(0.0240)	(0.0283)	(0.0217)
Adopt t-2	0.0554***	0.0385*	0.0502***	0.0652***	0.0624***
	(0.0130)	(0.0220)	(0.0141)	(0.0166)	(0.0188)
Cutoff Distance	$\bar{d} = 0$	$\bar{d} = 1$	$\bar{d} = 2$	$\bar{d} = 3$	$\bar{d} = 4$
R^2	0.0421	0.0418	0.0426	0.0433	0.0430
Observations	16378	16378	16378	16378	16378

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors, clustered at the 2 digit SIC level are in parenthesis. Year level and 4 digit SIC fixed effects are included in all columns. Baseline controls include: the log of one plus the number of (forward) citations received by the patents applied for by firm f in period t; log of firm size (measured by employment); the log of one plus the firms R&D expenditure over sales; firm age in period t; as well as dummy variables to control for whether firm f patents in period t, whether firm f issues their first patent in period t and whether firm f issues a *Computers & Communications* patent in period t. Unreported R&D is taken to be equal to 0.

B.3 ICT-related Products

In this appendix section, I discuss the construction of the ICT-related product measure and discuss the implications for the model structure. ICT-related products provide a useful benchmark to compare the model predictions and structure with data.

I adapt the basic classification of products from Acemoglu et al. (2016a) who examine within product competition between clean and dirty products. I define products j at the US three-digit classification (product class) level to maintain the full sample of the patent data. I construct the total stock of κ-related patents $K^{\kappa}_{j,t}$ in period t for product j as:

$$K^{\kappa}_{j,t} = 0.9K^{\kappa}_{j,t} + N^{\kappa}_{j,t},$$

where $N^{\kappa}_{j,t}$ is the number of κ-related product j patents issued in period t. The set of ICT-related products is then constructed as:

$$\mathcal{J}^{ict}_t = \left\{ j \in \mathcal{J} | K^{ict}_{j,t} > 0.25(K^{ict}_{j,t} + K^{non}_{j,t}) \text{ or } j \in \mathcal{J}^{ict}_{t-1} \right\}.$$

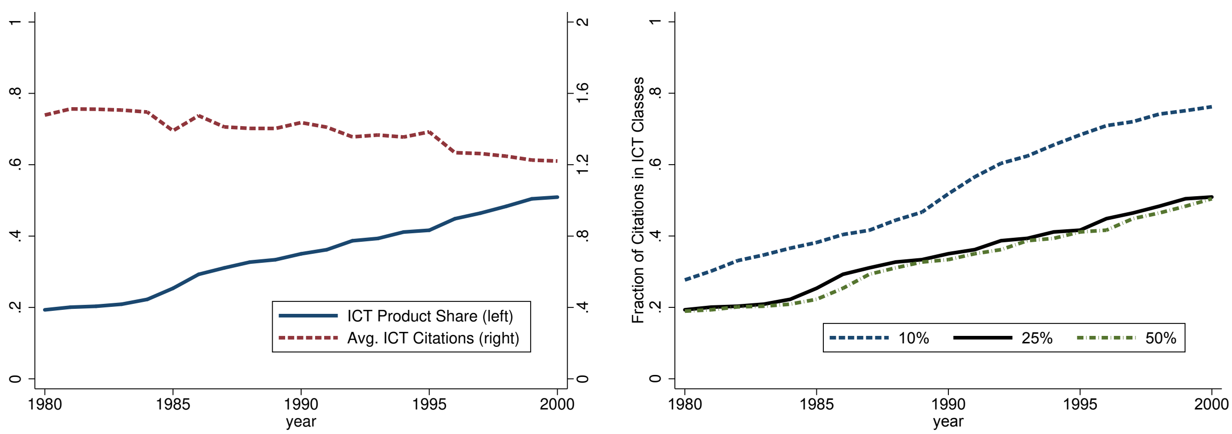
ICT-related products are those that hit a certain threshold (25%) or have hit that threshold in the past. I set the weights at 25% in the baseline analysis to capture the idea that the products are trending towards ICT-related patents, but Figure 8 shows that the results are similar if 50% is used as a threshold instead. The share of ICT-related products is given by

$$\hat{\mathcal{D}}_t^{ict} = \sum_{j \in \mathcal{J}_t^{ict}} W_j$$

where W_j is the weight of product j and is set equal to total share of product j patents over the entire period. To construct the above measures, I use the full sample of data and then trim to the same sample (1980-2000) used in the main analysis.

I also construct a measure of relative quality of ICT-related products as citation share of ICT-related patents divided by the patent share. Relative quality increases if ICT-related products receive more citations on average than other products. Figure 8 plots the implied ICT-related product share $\hat{\mathcal{D}}_t^{ict}$ and the relative quality of ICT-related products.

Figure 8: Diffusion by ICT-Related Products



(a) ICT Product Share

(b) Alternative Cutoffs

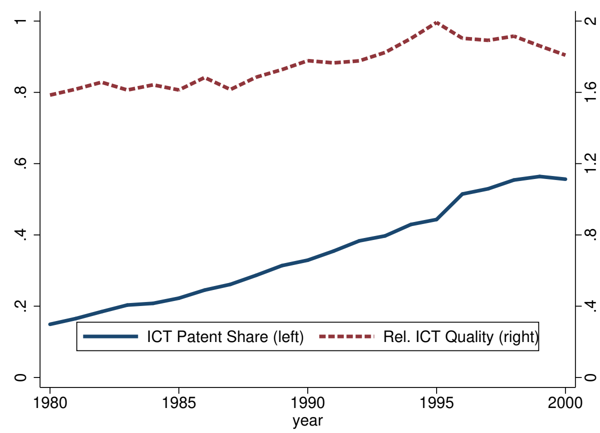
Notes: ICT-related products are defined as technology classes with at least 25% ICT-related patents. ICT Product Share measures the total number of patents associated with ICT products divided by total patents. Avg. ICT Citations measures total citations on ICT-related products divided by total citations on all patents divided by the number of ICT-related products.

The figure highlights that the share of ICT-related products increases from around 20% in 1980 to around 50% in 2000, consistent with the ICT product share implied by the model. The relative quality measure provides an additional test on the model structure. The model assumes that the entirety of diffusion is driven by the increase in the number of ICT-related products and not an increase in the relative quality of ICT-related products. The stability of the relative quality of ICT products is consistent with this view.

B.4 Extensive and Intensive Margins of ICT Diffusion

The model structure assumes that the diffusion of ICT occurs through the application of ICT ideas to new goods. An alternative possibility is that the share of ICT goods remains relatively constant over time, but that the relative quality of ICT-related goods increases. Figure 9 tests this alternative by comparing the unweighted patent share and the relative average citations received on ICT-related patents over the sample, which act as stand-ins for the share of ICT products $\mathcal{D}^{ict}(t)$ and the relative quality of innovations $\ln \lambda^{ict} / \ln \lambda^{non}$ in the model. The figure shows that ICT diffusion is primarily through new patents, consistent with the model structure. While ICT-related patents experience an advantage in terms of citations, this advantage appears to be relatively constant throughout the sample.

Figure 9: Intensive and Extensive Margins of ICT Diffusion



Notes: The figure reports the extensive and intensive margins of the citation-adjusted ICT-related patent share. ICT Patent Share measures the unweighted share of ICT-related patents. Rel. ICT Quality measures the average citations received by ICT-related patents divided by the average citations received by non-ICT-related patents.

B.5 Robustness of Main Results

The central result in the empirical analysis is the relationship between firm-level sales-per-employee growth and adoption. This serves to validate the measure and is also used as a target in the calibration. In this appendix section, I consider the robustness of the results to alternative specifications and controls. The empirical specification is given by:

$$g_{f,t}^{X} = \beta_0 + \sum_{s=0,-1,-2} \beta_{A,s} \text{Adopt}_{f,t+s} + \mathbf{X}_{f,t}\beta_X + \Gamma_s + \Gamma_t + \varepsilon_{f,t} \tag{22}$$

where $g_{f,t}^X$ is the growth in variable X in period t and other variables are defined as in the main text. I focus on the period of and two periods following adoption to simplify reporting.

Alternative Controls Table 12 presents robustness of the main firm-level regression under alternative controls. Column (1) presents the set of baseline controls, which serves as the baseline specification for the robustness analysis. Columns (2) and (3) alter the measure of R&D intensity to exclude missing values or measure a firm's cumulative R&D stock. Both alternatives do not have a substantial effect on the results. Column (4) uses an alternative measure of firm size (assets) in place of the baseline measure (employment). Column (5) adds the adopter variable which takes value one in all periods including and following the firm's adoption of ICT. The estimated coefficient is statistically insignificant and small indicating that adoption does not lead to a long-term increase in the adopting firm's labor productivity growth rate, consistent with the model.

Table 12: Alternative Controls

	Sales / Emp. Growth				
	(1)	(2)	(3)	(4)	(5)
Adopt t	0.0500***	0.0532***	0.0525***	0.0523***	0.0513***
	(0.0175)	(0.0180)	(0.0178)	(0.0179)	(0.0183)
Adopt t-1	0.0702**	0.0709**	0.0665**	0.0705**	0.0715***
	(0.0283)	(0.0308)	(0.0267)	(0.0282)	(0.0259)
Adopt t-2	0.0652***	0.0763***	0.0608***	0.0665***	0.0664***
	(0.0166)	(0.0176)	(0.0163)	(0.0166)	(0.0151)
Adopter					-0.00240
					(0.00923)
Alt. Controls	Baseline	Drop Missing R&D	R&D Stock	Assets	Adopter
R^2	0.0433	0.0448	0.0427	0.0390	0.0433
Observations	16378	11927	16378	16378	16378

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors, clustered at the 2 digit SIC level are in parentheses. Year and 4 digit SIC fixed effects are included in all columns. Baseline controls include: the log of one plus the number of (forward) citations received by the patents applied for by firm f in period t; log of firm size (measured by employment); the log of one plus the firms R&D expenditure over sales; firm age in period t; as well as dummy variables to control for whether firm f patents in period t, whether firm f issues their first patent in period t and whether firm f issues a *Computers & Communications* patent in period t. Unreported R&D is set to 0. Column (2) removes unreported R&D. Column (3) adds log of one plus the R&D stock and remove R&D from the controls. R&D stock is calculated as R&D $\text{Stock}_t = \text{R\&D}_t + 0.90\text{R\&D Stock}_{t-1}$. Column (4) adds log of total assets as a control and removes log employment. Column (5) add an adopter control for whether the firm has adopted the technology prior to or in period t.

Alternative Sector Definitions Table 13 reports the results using different sectoral definitions. The baseline specification defines sectors at the 4 digit SIC level. The results in columns (2) and (3), which change the sector definitions to SIC 3 and SIC2, are quantitatively similar to the baseline specification in column (1). Both sector definitions are broader than considered in the baseline specification. Column (4) examines firm-level controls in place of sector-level controls. This is the narrowest control and captures additional heterogeneity across firms, such as innovator heterogeneity, that could potentially drive the estimates. The overall impact of adoption is similar to the baseline regression, but the coefficients indicate that the gains to productivity growth are closer to the adoption date.

Table 13: Alternative Sector Definition

	Sales / Emp. Growth			
	(1)	(2)	(3)	(4)
Adopt t	0.0500***	0.0515***	0.0508***	0.0765***
	(0.0175)	(0.0171)	(0.0167)	(0.0175)
Adopt t-1	0.0702**	0.0735**	0.0719**	0.0571*
	(0.0283)	(0.0319)	(0.0309)	(0.0290)
Adopt t-2	0.0652***	0.0686***	0.0687***	0.0319*
	(0.0166)	(0.0160)	(0.0151)	(0.0180)
Sector Definition	Baseline	SIC 3	SIC 2	Firm-Level
R^2	0.0433	0.0359	0.0276	0.237
Observations	16378	16382	16387	16101

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors, clustered at the 2 digit SIC level are in parentheses. Year and sector level fixed effects are included in all columns. Sectors are defined as indicated. Baseline controls include: the log of one plus the number of (forward) citations received by the patents applied for by firm f in period t; log of firm size (measured by employment); the log of one plus the firms R&D expenditure over sales; firm age in period t; as well as dummy variables to control for whether firm f patents in period t, whether firm f issues their first patent in period t and whether firm f issues a *Computers & Communications* patent in period t. Unreported R&D is set to 0. Age is dropped as a control when the Sector is defined at the firm-level.

Alternative Outcomes Table 14 reports the results using alternative measures of growth. The baseline results focus on sales-per-employee because this measure is the most relevant for the calibration since it directly relates to the parameter η^{ict} in the model. The results in table 14 show that qualitatively the results holds for sales, employment and capital. Together, the results show that firms become larger following ICT adoption.

Table 14: Alternative Outcomes

	(1) Sales/Emp. Growth	(2) Sales Growth	(3) Emp. Growth	(4) Asset Growth
Adopt t	0.0500***	0.0909***	0.0446**	0.0495
	(0.0175)	(0.0299)	(0.0200)	(0.0312)
Adopt t-1	0.0702**	0.102***	0.0463***	0.126***
	(0.0283)	(0.0232)	(0.0170)	(0.0269)
Adopt t-2	0.0652***	0.0929***	0.0342***	0.0108
	(0.0166)	(0.0161)	(0.00856)	(0.0124)
R^2	0.0433	0.0863	0.115	0.0898
Observations	16378	16378	16378	16376

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors, clustered at the 2 digit SIC level are in parentheses. Year level and 4 digit SIC fixed effects are included in all columns. Baseline controls include: the log of one plus the number of (forward) citations received by the patents applied for by firm f in period t; log of firm size (measured by employment); the log of one plus the firms R&D expenditure over sales; firm age in period t; as well as dummy variables to control for whether firm f patents in period t, whether firm f issues their first patent in period t and whether firm f issues a *Computers & Communications* patent in period t. Unreported R&D is taken to be equal to 0.

Alternative Estimator Callaway and Sant'Anna (2021) develop an alternative estimator to deal with two-way fixed effects. Table 15 reports the results from this estimator when the comparison group is taken to be either the not yet and never treated group. I report the consolidated results for the four periods ahead lead and lag and suppress the remaining results for readability. For the most part, the results outside of this range lacks statistical significance. The results are broadly consistent with the baseline results, albeit with weaker significance, in that labor productivity growth increases around the time of adoption and remains elevated for several periods after, before reverting to its-pre-adoption average.

B.6 Alternative Firm and Patent Cutoffs

In the baseline construction of variables, I use a 10% cutoff for both ICT-related patents and firms. The cutoff is chosen such that the typical patent would quality with a single patent since the median patent has close to 10 citations. The cutoff is included to prevent patents that cite more frequently and firms that issue more patents from being mechanically more likely to be labeled as ICT-related. In this appendix section, I examine the robustness of the main results to using 0% and 20% cutoffs in place of the baseline 10% cutoff for both definitions. I also examine an adjustment to a 10% depreciation per year of past patenting

Table 15: Callaway and Sant'Anna (2021) Estimation Results

	Not Yet		Never	
	Coefficient	SE	Coefficient	SE
Tm4	-0.0537	(0.0443)	-0.0523	(0.0440)
Tm3	-0.0176	(0.0399)	-0.0203	(0.0395)
Tm2	0.0301	(0.0356)	0.0292	(0.0350)
Tm1	-0.0337	(0.0332)	-0.0343	(0.0329)
Tp0	0.0552	(0.0320)	0.0550	(0.0315)
Tp1	0.0189	(0.0301)	0.0179	(0.0298)
Tp2	0.0339	(0.0391)	0.0311	(0.0391)
Tp3	0.0686	(0.0389)	0.0636	(0.0389)
Tp4	0.0597	(0.0435)	0.0545	(0.0437)

Notes: Outcome variable is growth in firm labor productivity $g_{f,t}$. Not Yet and Never refer to the estimator using the not yet treated and never treated groups as the comparison. TmX refers to the adoption period minus X periods. TpX refers to the adoption period plus X periods.

for the ICT-related firm definition. Table 16 summarizes the results

The table shows that the results are robust to alternative cutoff rules. Except for the 20% patent cutoff, all of the estimated coefficients are also statistically significant. The 20% patent cutoff is mechanically the most restrictive as patent citations are a more continuous measure and so the change in cutoff leads to a large change in the number of ICT-related patents. Overall, the similarity of the estimates provides reassurance that the cutoff rules are not driving the results.

B.7 Placebo Groups

A concern with the measure of ICT-related patents is that the links between patents are correlated with some other aspect of innovation that are not specific to ICT. For example, it could be the case that building on patents in other technology classes is inherently valuable. To address this concern, I consider several placebo groups and use them to construct alternative measures of adoption. I then perform the same analysis on the placebo groups as on the baseline measure of ICT diffusion. The hypothesis is that these groups should not be related to firm growth as they do not capture ICT adoption. On the other hand, if there is an underlying variable picked up by the construction of the ICT group, then these groups should also pick up this relationship. I consider five placebo groups corresponding to other technology classes (Chemicals, Drugs, Electronics, Mechanics, and Other). These groups are constructed following an identical methodology as with the baseline group. Table 17 presents

Table 16: Alternative Cutoff Assumptions

	Sales / Emp. Growth					
	(1)	(2)	(3)	(4)	(5)	(6)
Adopt t	0.0515***	0.0362	0.0481***	0.0415*	0.0400**	0.0484***
	(0.0171)	(0.0231)	(0.0178)	(0.0208)	(0.0160)	(0.0168)
Adopt t-1	0.0735**	0.0482	0.0761**	0.0742**	0.0626**	0.0704**
	(0.0319)	(0.0354)	(0.0293)	(0.0311)	(0.0254)	(0.0297)
Adopt t-2	0.0686***	0.0573***	0.0700***	0.0754***	0.0546***	0.0668***
	(0.0160)	(0.0144)	(0.0188)	(0.0168)	(0.0146)	(0.0141)
Alt. Cutoff	Baseline	20% Pat.	0% Pat.	20% Firm	0% Firm	10% Dep.
R^2	0.0359	0.0264	0.0280	0.0276	0.0271	0.0276
Observations	16382	16387	16387	16387	16387	16387

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors, clustered at the 2 digit SIC level are in parentheses. Year and 2 digit SIC sector-level fixed effects are included in all columns. All columns include the log of one plus the number of (forward) citations received by the patents applied for by firm f in period t; log of firm size (measured by employment); the log of one plus the firms R&D expenditure over sales; firm age in period t; as well as dummy variables to control for whether firm f patents in period t, whether firm f issues their first patent in period t and whether firm f issues a *Computers & Communications* patent in period t. Unreported R&D is taken to be equal to 0. Modification indicates the change in the definition of ICT-related patents and firms used to construct the adoption variables. X% patent changes the cutoff used for ICT-related patents to X% from 10% in the baseline analysis. X% firm changes the cutoff used for ICT-related firms to X% from 10% in the baseline analysis. 10% dep. changes the firm portfolio calculation to a 10% deprecation rate from 0% in the baseline analysis.

the estimates for adoption based on the other technology class.

The results in Table 17 show that only adoption based on the *Computers & Communications* technology classes leads to a positive and statistically significant increase in firm productivity. The results for adoption based on the *Drug* and *Mechanics* technology classes lead to productivity falling after adoption. The results are consistent with the hypothesis that the application of ICT to other areas is capturing a fundamental change to firms.

B.8 Entrant Adoption

Entrants enter the market by innovating on an existing good to become the new leading-edge producer of that good. The baseline model assumes that entrants use ICT if the current producer of the good uses ICT. However, it may be the case that entrants have a comparative advantage with ICT because they are not restricted by previous practices. In this appendix section, I examine the role of entrant adoption.

Table 17: Labour Productivity Growth and Adoption

	Sales / Emp. Growth						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Adopt Chem	-0.00323						0.0130
	(0.0326)						(0.0324)
Adopt C&C		0.0382**					0.0471*
		(0.0185)					(0.0256)
Adopt Drug			-0.0480**				-0.0536***
			(0.0223)				(0.0174)
Adopt Elec				0.0180			0.0143
				(0.0282)			(0.0385)
Adopt Mech					-0.0382*		-0.0592**
					(0.0214)		(0.0232)
Adopt Other						0.00856	0.0230
						(0.0228)	(0.0229)
R^2	0.0413	0.0415	0.0415	0.0413	0.0415	0.0413	0.0423
Observations	16378	16378	16378	16378	16378	16378	16378

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors, clustered at the 2 digit SIC level are in parentheses. Year level and 4 digit SIC fixed effects are included in all columns. Controls include: the log of one plus the number of (forward) citations received by the patents applied for by firm f in period t; log of firm size (measured by employment); the log of one plus the firms R&D expenditure over sales; firm age in period t; as well as dummy variables to control for whether firm f issues their first patent in period t and whether firm f issues a *Computers & Communications* patent in period t. Unreported R&D is set to 0.

I define entrants as first-time assignees in the patent database. I use the model to construct a measure of entrant comparative advantage that can be analyzed in the data. Let $\alpha_e(t) \in [0, 1]$ be the probability that an entrant that draws a non-ICT good and enters as an ICT firm. In discrete time, the mass of ICT entrants in period t is $z_e(t)\left[\mathcal{D}^{ict}(t-1) + \alpha_e(t)\mathcal{D}^{non}(t-1)\right]$. For the empirical comparison, I approximate the share of ICT products $\mathcal{D}^{ict}(t)$ with the share of ICT patents.[35] The comparative advantage of entrants with ICT can be written as:

$$\mathrm{CA}_t = \underbrace{\frac{\text{ICT Entrants}_t}{\text{Entrants}_t}}_{\approx \mathcal{D}^{ict}(t-1)+\alpha_e \mathcal{D}^{non}(t-1)} - \underbrace{\frac{\text{ICT-Related Patents}_{t-1}}{\text{Total Patents}_{t-1}}}_{\approx \mathcal{D}^{ict}(t-1)}. \tag{23}$$

[35] In the baseline calibration the average ICT citation share is 49% while the average ICT product share is 37% over the period from 1980 to 2000. Together this would imply that the estimates value of α_e is around 30% larger than indicated by the constructed value of CA_t.

The term CA_t is the ICT entrants beyond what is implied by the prior stock of knowledge. The value of CA_t is positive if entrants have a comparative advantage ($\alpha_e > 0$).

Estimating the average value of CA_t controlling for differences across technology class implies an average value of $\alpha_e = 2.5\%$, which indicates that an entrant that draws a non-ICT good will become an ICT firm 2.5% of the time.

C Other Testable Model Predictions

In this appendix section, I compare the model predictions with data to provide additional support for the fit of the model.

C.1 ICT Innovations

The model assumes that firms after a firm adopts ICT it only produces ICT-related innovation. While this does not hold in a strict sense in the data, I use the data to test the extent that adopters are more likely to issue ICT-related patents. I estimate:

$$\text{ICT-Related Patent}_{f,t} = \exp\left\{ \underset{(0.0084)}{1.757} * \text{Adopter}_{f,t} + \Gamma_{c,t} \right\} + \varepsilon_{f,t} \tag{24}$$

where the number in parentheses is the robust standard error. The estimate shows that adopters are around six times more likely to apply for ICT-related patents than non-adopters.

C.2 Innovation Value

The model also makes predictions about firm value over the transition. I use the patent value measures developed by Kogan et al. (2017) to examine how innovation value changes over time. The advantage of using innovation value, as opposed to firm value, is that it does not rely on observing the number of products that a firm operates and is directly related to the innovative value of the firm. Figure 3a makes two predictions about firm value $V_1^\kappa(\mu)$: (i) ICT firm innovations are relatively more valuable than non-ICT firm innovations; (ii) ICT firm innovations become relatively less valuable over time. Table 18 summarizes the results.

The estimates are consistent with the model predictions. Column (1) shows that adopters have higher estimated patent values compared with non-adopters. The results also serve as a robustness check to Table 4 by showing that the higher quality of adopter innovations holds using an alternative measure. Column (2) shows that adopter patents are relatively more valuable in early periods and that the relative value compared with non-adopters diminishes over time, consistent with the second prediction.

Table 18: Patent Value by Firm Type

	log Patent Value	
	(1)	(2)
Adopter	0.100***	0.424***
	(0.00651)	(0.00820)
Adopter × t		-0.0398***
		(0.000590)
Observations	523173	523173

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors are reported in parentheses. All columns include firm-level fixed effects, technology class fixed effects, and year fixed effects. Period t is defined such that 1980 is $t = 0$.

C.3 Innovation Frequency

The model makes three predictions about the exploration innovation frequency $A^{\kappa}(t)z^{\kappa}(t)$ over the transition: (i) ICT firms innovate relatively less frequently in the early periods; (ii) ICT firms innovate relatively more frequently over time; and (iii) on average, ICT firms innovate relatively more frequently over the 1980 to 2000 period. The third prediction is directly targeted in the model calibration and so it holds by construction. The first and second predictions can be tested using the patent data. Table 19 summarizes the results.

Table 19: Innovation Frequency by Firm Type

	Unweighted		Citation Weighted	
	(1)	(2)	(3)	(4)
Adopter	0.888***	-0.237***	1.014***	-0.439***
	(0.0553)	(0.0471)	(0.0566)	(0.0520)
Adopter × t	0.0102**	0.0472***	0.0381***	0.0780***
	(0.00430)	(0.00314)	(0.00462)	(0.00341)
Firm FE	No	Yes	No	Yes
Observations	219232	169219	219232	168646

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The values are estimated using Poisson pseudo-maximum likelihood (PPML). Robust standard errors are reported in parentheses. Period t is defined such that 1980 is $t = 0$. All columns include year fixed effects.

The results in the table show that adopters tend to innovate more frequently over time, consistent with the predictions of the model. Additionally, once firm-level fixed effects are accounted for, as in columns (2) and (4), then ICT firms tend to innovate less frequently in early periods of the data, also consistent with the model. One reason that the firm-level fixed effects may be important for highlighting this feature is that more innovative firms tend to also be early adopters of ICT.

C.4 Competition Spillovers

The model predicts that competition $\delta^{\kappa}(t)$ faced by firms varies over the transition path. I use the probability that a firm exits the sample to test competition spillovers. The model makes two predictions about firm exit over the transition path: (i) ICT firms are less likely than non-ICT firm to exit; and (ii) ICT firms become relatively more likely to exit over time. Table 20 explores these trends empirically using the patent data. In the model, larger firms are less likely to exit because they accumulate more product lines that shield them from being pushed out of the market. To control for this, I include the firm's discounted patent stock, calculated as $\text{Pat Stock}_{f,t} = 0.9 * \text{Pat Stock}_{f,t-1} + N_{f,t}$ where $N_{f,t}$ is is the number of patents applied for by firm f in period t.

Table 20: Exit Probability by Firm Type

	Exit	
	(1)	(2)
Adopter	-0.135***	-0.204***
	(0.00882)	(0.0190)
Adopter × t		0.00527***
		(0.00122)
Observations	140870	140870

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The values are estimated using Poisson pseudo-maximum likelihood (PPML). Robust standard errors are reported in parentheses. Both columns include year fixed effects and a control for the firm's discounted stock of patents. Period t is defined such that 1980 is $t = 0$.

The relationships in Table 20 are consistent with the model predictions. Adopters are both relatively less likely to exit than other firms and become more likely to exit over the transition period.

D Quantitative Appendix

D.1 Algorithm to Calculate the Transition Path

Solving the transition path involves solving for the sequence of value functions $\{v^{\kappa}(t), B^{\kappa}(t)\}$ and aggregate states $\{\mathcal{D}^{\kappa}(t), \bar{q}(t), \mu(t)\}$ for each period in the transition path $t \in [0, \bar{t}]$. I use a modified version of the algorithm from Acemoglu et al. (2016a). I solve the discrete approximation to the transition path equilibrium for a small interval $\Delta t = 0.10$ for 1,100 periods (110 years). I initialize the economy at the non-ICT BGP and introduce ICT in period $t = 0$. Following its introduction, firms become aware of the properties of ICT as well as the evolution of the economy over the transition period. I assume that ICT becomes fully diffused by the final period, which roughly corresponds to the year 2080 in the data. I solve the transition path using the following algorithm:

1. Guess a sequence of value functions $\{v^{\kappa}(t), B^{\kappa}(t)\}$ over the transition path. I take the starting point to be BGP values for each technology in all periods.

2. Starting in period 0, solve the economy forward using the the guess of the value functions. The innovation rates are solved using the discrete time approximations from the Proof of Proposition A.3. The entry rate $z_e(t)$ solves the entry problem. Update the states, $(g(t), \mathcal{D}^{\kappa}(t), \Phi(\mu, t))$, using the discrete approximations for the laws of motions from the main text and the innovation intensities.

3. Solve backwards over the transition path for the value functions using the discrete time approximations from the Proof of Proposition A.3.

4. Repeat steps (2) and (3) until the value functions converge.

D.2 Sensitivity of Model Moments

Table 21 reports the change in model moments used in the joint calibration to a 10% shock to the parameter value. For the case of the step sizes, the shock is calculated on the component greater than one, e.g., $\lambda_x^{non\prime} = 1 + 1.1 \times (\lambda_x^{non} - 1)$. This highlights the interconnectedness of the calibration and how each model parameter relates to the moments targeted in the joint calibration.

The table shows that while some moments are more closely related to some parameters, as discussed in the baseline text, the determination of most moments depends on the joint choice of parameters. For example, the mean squared error for ICT citations in the model and data is closely related to the adoption cost function parameters $\{\psi_a, \zeta_a\}$, as is expected.

Table 21: Sensitivity of Model Moments to a 10% Change in Parameter Values

	Avg Entry	Rel Inn Rate	Profit	Adopt Growth	ICT 1999	g^{non}	MSE	Prob ICT
ζ_a	5.0	98.6	1.3	-3.3	12.3	0.0	600.7	-13.1
ψ_x^{non}	9.4	16.1	-0.3	2.1	-1.0	-5.1	-17.2	2.9
ψ_x^{ict}	12.0	-34.0	0.0	2.1	-2.9	0.0	-34.9	-1.0
λ^{non}	18.3	-19.9	5.4	-7.6	-3.9	14.6	-54.1	-1.9
ψ_e	-50.2	31.3	0.7	1.8	7.5	0.0	234.2	-4.1
η^{ict}	0.4	8.8	0.8	8.9	1.2	0.0	32.7	-0.9
ψ_a	-1.2	-15.4	-0.2	0.3	-2.1	0.0	-39.5	2.3
α^{ict}	-7.7	48.5	0.4	-2.0	5.2	0.0	125.6	5.7

Notes: The table reports the percent change in model moments from a 10% change in the indicated parameter, except for MSE where the value is calculated as 100 times the change in the MSE.

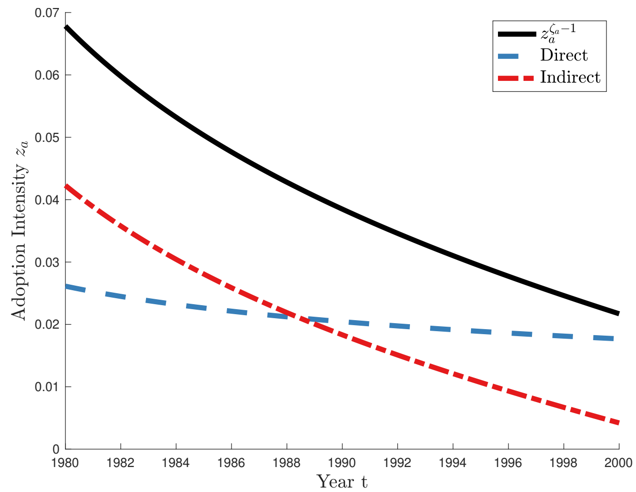
However, this moment is also dependent on the entry cost ψ_e, which pins down the average value of firms along the transition, and the cost of ICT exploration R&D since this is another channel of diffusion.

D.3 Drivers of Adoption

In this appendix section, I examine the drivers of adoption R&D over the transition path. Figure 10 reports the adoption intensity $z_a(t)$ of non-ICT firms over the transition path as well as the direct and indirect benefits from adoption, that are described in Proposition 1. As discussed in the main text, the adoption intensity is relatively small compared to the other R&D activities because of the relatively high cost of adoption ψ_a and low returns η^{ict}.

The figure shows that adoption intensity declines over the transition period because of a decline in the indirect benefits of adoption. The indirect benefits of adoption measures the relative increase in firm value of producing a set of products with ICT, as opposed to with non-ICT, holding the markups fixed. The indirect benefits depend on the extent of knowledge and competition spillovers in the economy, which are reported in Figure 4. On one hand, ICT firms face relatively lower competition spillovers in early periods when there are relatively few ICT firms, since $\alpha^{non} = 0$. On the other hand, ICT firms are relatively less likely to innovate in early periods because of lower knowledge spillovers, since $\alpha^{ict} < 1$. Quantitatively, the first channel dominates leading to positive and declining indirect benefits from adoption.

Figure 10: Adoption Decomposition



Notes: The figure reports adoption intensity $z_a(t)$ and the direct and indirect benefits from adoption. Expressions are given in Proposition 1.

D.4 Sensitivity to Non-ICT Spillovers

The baseline model assumes that non-ICT firms only innovate on non-ICT goods. In this section, I examine the sensitivity of the main results to setting $\alpha^{non} = 4.3\%$. Since ICT-related patents are defined using citation links, the moment used to calibrate α^{ict} in the baseline calibration is mechanically small. Instead, I construct the target for α^{non} using the probability that non *Computers & Communications* patents cite *Computers & Communications* patents. I then approximate $\mathcal{D}^{ict}$ using the share of ICT-related patents to directly infer the value of α^{non} from the data. For tractability, I assume that non-ICT firms that innovate on ICT products are able to build on both components of quality, such that a non-ICT firm f that innovates on good j produced by ICT firm f' has quality $q_{j_f} = \lambda^{non}(\eta^{ict} q_{j_{f'}})$.

I recalibrate the other model parameters to match the same moments as in the baseline calibration. Table 22 summarizes the model and data moments.

Table 23 summarizes the calibrated parameters in the extended model. Both the relative cost of ICT exploration R&D $\psi_x^{ict}/\psi_x^{non}$ and cost of adoption R&D ψ_a decline relative to the baseline calibration. The decline compensates for the increase in α^{non} creating a hinderance to diffusion since it leads to some ICT products being converted to non-ICT products. A consequence of the lower cost of ICT exploration R&D is that the long-run growth rate increases relative to the predicted long-run growth rate in the baseline economy.

Figure 11 shows that the path for ICT diffusion and growth are similar to the baseline experiment. The transition path growth rate is slightly higher than in the baseline calibration, which is explained by the relative decline in the costs ICT exploration and adoption

Table 22: Calibration Moments with Non-ICT Spillovers

Description of Moment	Data	Model
Initial BGP Growth (%)	2.00	1.99
Entry Share (%)	17.9	17.7
Profit Margin (%)	12.2	12.4
Relative Innovation Rate	0.203	0.177
Post-Adoption Growth	3.7	3.7
ICT Patent Share in 1999 (%)	70.7	68.4
MSE ICT Patent Share	0	0.02
Prob ICT Cites Non-ICT (%)	28.9	30.4

Table 23: Calibrated Parameters with Non-ICT Spillovers

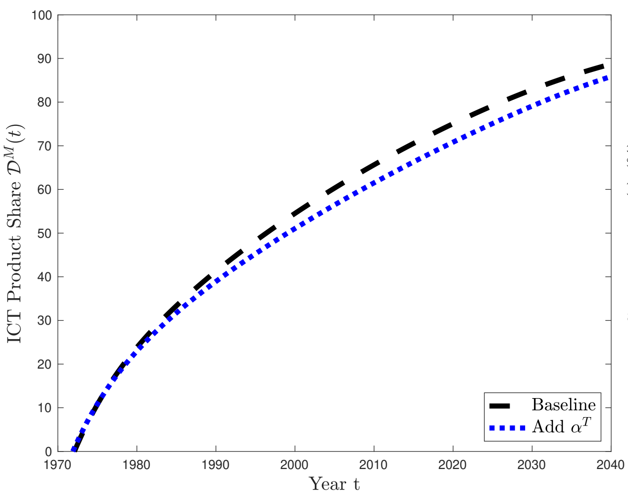
Parameter		Value
Cost of Entry	ψ_e	1.39
Cost of Non-ICT Exploration R&D	ψ_x^{non}	2.80
Cost of ICT Exploration R&D	ψ_x^{ict}	2.85
Non-ICT Step Size	λ^{non}	1.119
ICT Step Size	λ^{ict}	1.157
Level Cost of Adoption R&D	ψ_a	10.16
Rel Quality of ICT Products	η^{ict}	1.055
Cross-technology Spillovers	α^{ict}	0.261
Crv Adoption R&D	ζ_a	1.66

R&D.[36] Over time, the growth paths start to diverge as the transition path in the extended model converges to a higher balanced growth path in the long-run. The takeaway is that the assumption of $\alpha^{non} = 0$ in the baseline calibration does not substantially change the quantitative results.[37]

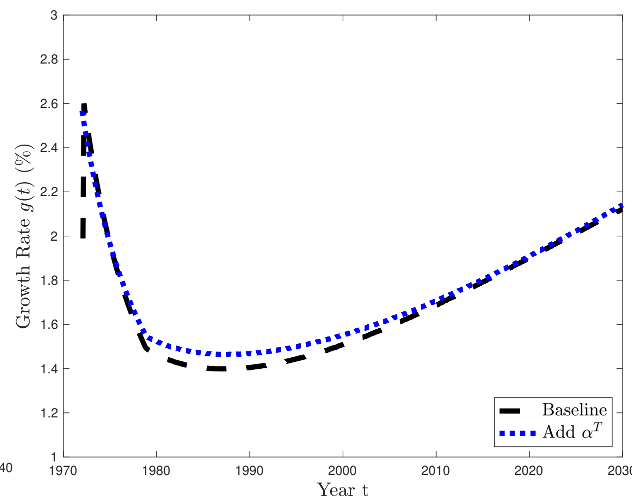
[36] The assumption that non-ICT firms build on both components of quality (i.e. $q_{j_{f'}} = \lambda^{non}(\eta^{ict} q_{j_f})$ rather than $q_{j_{f'}} = \lambda^{non} q_{j_f}$) artificially increases the growth rate over the transition periods, but this effect is small.

[37] The growth path incorporates a full recalibration of the model and should not be taken as a comparative static on the cross-technology spillover α^{non} (or α^{ict}).

Figure 11: Growth and Diffusion with Non-ICT Spillovers



(a) Diffusion



(b) Growth