Vulnerable Technological Adoption*

Sebastián Amador†

This version: September 17, 2023

Job Marker Paper

Click here for latest version

Abstract

I explore the lasting consequences of large productivity shocks, examining how they impact technological adoption. First, I develop a model in which firms face heterogeneous fixed costs of implementing and operating a technology, which they can adopt or abandon. Only large shocks can affect the share of adopters, and these impacts are bounded in the long-run. To validate the model's predictions, I employ semi-nonparametric local projections on a panel comprising 14 technologies across 18 countries, spanning 112 years. This allows to detect non-linearities without any a priori functional form assumptions. Estimates are closely aligned with the model predictions. Using a version of the model calibrated to the empirical estimates, I show that subsidizing technology operational costs can lead to significant welfare improvements, averting harm to the economy's productive potential.

JEL classification codes: E32, O11, O33, O47

*An earlier version of this paper originally circulated under the title "Hysteresis, the Big Push, and Technological Adoption".

†Department of Economics, University of California, Davis (email: jsamador@ucdavis.edu, website: https://sites.google.com/view/sebastian-amador.) I am extremely grateful to Òscar Jordà, Alan Taylor, Sanjay Singh, and Nicolas Caramp for their invaluable mentorship, support, and advice. I also want to thank Michaela Elfsbacka Schmöller, Xian Jiang, Colin Cameron, Arman Rezaee, Christopher Meissner, David Baqaee, Paula Beltran, Camilo Granados, and participants at the UC Davis Macroeconomics Lunchtime Presentation, the Bank of Finland's Villa Research Workshop, the Fall 2022 Midwest Macro Meeting, the LACEA LAMES 2022 Annual Meeting, and the AEA 2023 Annual meeting, for comments and discussion. All errors and omissions are my own. The views expressed in this paper are the sole responsibility of the author.

1 Introduction

The adoption of new technologies is a fundamental driver of economic growth (Parente and Prescott, 2002). In this paper, I argue that the transition from innovation to widespread usage can be vulnerable to temporary productivity shocks in a non-linear manner. The argument rests upon two premises. First, in the presence of demand complementarities, firms' technological choices become interlinked. This means that when a particular producer adopts an increasing returns to scale technology, it increases demand for the products of other firms, thereby increasing incentives for them to also adopt the same technology.[1] The second premise is that—if conditions are sufficiently adverse—it can be in a firm's best interest to temporarily or indefinitely abandon a costly technology. The decision to adopt requires assessing potential benefits against the associated costs. Similarly, if conditions are worse than anticipated, the expected benefits from operating a technology might fall short of its costs. In such situation, a firm might choose to abandon a technology that has not been driven into obsolescence by newer alternatives, and even if it plans to adopt again in the future. While some historical examples demonstrate technologies can be abandoned for prolonged periods[2], it is important to emphasize that complete abandonment is not a prerequisite for the argument put forth here. For abandonment to have meaningful aggregate effects, it is only necessary for some—and not all—firms to stop operating the technology after a negative shock.

Grounded in these two premises, I construct a model where intermediate input producers respond to a temporary and unexpected aggregate productivity shock by possibly adopting or abandoning a new technology, with a traditional alternative as the outside option. Both the initially implementation and the continued operation of the modern technology entail paying fixed costs, which are heterogeneous across firms. This implies the modern technology exhibits increasing returns to scale. This framework is similar to Murphy et al. (1989), in which in which if a sufficient mass of firms adopts, they can trigger a permanent increase in output, even though individual firms may find it unprofitable to adopt independently. First, I theoretically analyze the problem of the intermediate variety producer. On a first stage, I derive optimal prices and quantities given a technological choice. Relative to firms using the traditional production method, those operating the more productive technology will have a lower marginal cost, lower prices, and will produce a higher quantity of their good. Then, using the first stage as an input, I characterize the problem of choosing between the more productive—but costly—modern technology, and the less productive—but free—traditional alternative. Specifically, using the new technology entails paying a fixed adoption cost, and a

[1] Some evidence documenting complementarity in Korean firms' technology adoption decisions can be found in Choi and Shim 2022

[2] For example, supersonic airliners after the 1973 oil shock. See Appendix A for further details.

continuation cost every subsequent period. The adoption cost is assumed to be higher than the continuation cost, reflecting a situation in which replacing the traditional production method necessitates of substantial expense, owing to replacing outdated equipment with the more advanced alternative, coupled with the need for supplementary investments during the initial phases to fully exploit the advantages inherent in the new technology (as in, e.g., Brynjolfsson et al., 2021). Choosing between the two alternatives entails comparing two affine functions of the productivity shock, one corresponding to the new technology having a higher slope and a lower intercept than the one corresponding to the traditional production method. The optimal adoption-abandonment policy is expressed in terms of the realization of the aggregate productivity shock for which the firm is indifferent between the status quo and the corresponding outside option (i.e., the point in which a firm not using the technology is indifferent between adopting and staying out, and the point in which a firm using the technology is indifferent between abandoning it and staying in). These two indifference points will correspond to two thresholds, one for adoption and another for abandonment, where the former is higher than the latter. As a result, when the shock is large and positive, marginal cost falls, incentivizing them to pay the adoption cost in order to take advantage of economies of scale to increase profits. Similarly, when the shock is large and negative, marginal cost increases, and the firm can reduce average cost and raise prices (and profits) by abandoning the technology. The region between the two thresholds constitutes a "band of inaction", inside of which the magnitude of the shock is not enough to induce a change in the firm's technological choice.

Just as in the case of an individual firm, the aggregate level of technological adoption will be impacted by large shocks only, while small shocks have no effect. However, this effect is bounded in the long-run. That is, as a negative shocks dissipates over time, some firms will re-adopt the modern technology. Thus no negative shock will reduce the share of firms operating the modern technology in the long-run below a minimum sustainable level. Similarly, no positive shock will increase the number of firms using the modern technology above a maximum sustainable level, as some of them will abandon as the shock washes out. As a result, the relationship between the size of a temporary productivity shock, and changes in the share of firms using the new technology is predicted by the model to be highly non-linear, exhibiting both a Big Push in which only large positive shocks trigger enduring increases in technological adoption, and a Big Pull, in which large negative shocks can result in long-lasting reductions in the share of firms operating the new technology.

Next, using aggregate data on the adoption of 14 technologies in 18 countries, for the period 1870-2002, I estimate the relationship between total factor productivity shocks and long-run changes in technological adoption, by applying semi-nonparametric methods to local projections. While it is possible to study non-linearities in impulse responses using alternative methods (e.g., stratified local projections as in Cloyne et al., 2023), they require specifying a particular functional form to model the non-linearity. However, in the application discussed here, specifying a functional form implies determining a priori which shocks are large and which are small. Also, it is possible that the true relationship implied by the data generating process is far from the selected specification. Thus, in order to use a less demanding approach, I estimate the response of technological adoption to a productivity shock by semi-nonparametric local projections. This method entails replacing the linear term corresponding to the impulse variable in a Jordà (2005) local projection specification by a non-specified function, to be estimated non-parametrically. Results suggest that the non-linearities in the data match those predicted by the model. That is, both the Big Push and the Big Pull are present in the data, and this effect appears to be driven by the subset of technologies requiring higher capital costs. Importantly, the presence of the Big Push indicates that some temporary shocks can generate lasting benefits, a fact not encountered in previous work, in which only negative shocks are found to have significant long-term effects, (Jordà et al., 2020; Amador, 2022; Aikman et al., 2022).

Finally, using a simplified version of the model calibrated to match the non-linearities estimated via semi-nonparametric local projections, I evaluate the potential welfare gains from subsidizing the cost of continuing to operate the modern technology, in order to prevent firms from abandoning it. Unlike the decentralized economy's, this policy takes into account the aggregate demand externality due to complementarity, solving a coordination failure in the technological choices of firms' after large negative shocks. While for small shocks this policy does not improve welfare relative to the decentralized economy, it is welfare improving for large shocks. These welfare gains are increasing in the size of the shock, and are particularly large for large shocks that are followed by a period in which some firms re-adopt, incurring in high re-implementation costs.

This paper is related to the literature on the relationship between the cycle and the trend, to works on the long-run impact of temporary shocks, and to models interpreting output fluctuations as coordination failures. Each of these bodies of work will be briefly described in turn.[3]

The relationship between the cycle and trend: The long-run impact of temporary

[3] An additional work that is closely related, but does not fit any of the categories below, is Bilbiie et al. (2012). This work provides a model of fluctuations with monopolistic competition and endogenous producer entry, subject to sunk costs. Unlike the model presented here, their contribution does not include heterogeneity across producers and firm exit is exogenous. As Ghironi (2018) points out, these choices imply that the model can be solved using a log-linear approximation, at the cost of featuring long-run impacts from temporary shocks (i.e., hysteresis). The model outlined in the present paper exhibits hysteresis effects by including an exit mechanism, and does not rely on an approximated solution method.

shocks has often been referred to as hysteresis. One of the early uses of this term in the context of the relationship between the cycle and the trend can be found in Blanchard and Summers (1986). This work hypothesized that persistently high unemployment after recessions in post-war Europe was due to hysteresis effects arising from labor market frictions. Following King et al. (1988) and Stadler (1990),= several papers show how, in endogenous growth models, shocks that temporarily disrupt the growth process can have permanent effects on output. The pro-cyclicality of growth-enhancing variables such as investment, R&D, and technological adoption has been often interpreted as evidence in favor of this view (e.g., Anzoategui et al. 2019; Benigno and Fornaro 2018; Bianchi et al. 2019; Garga and Singh 2021; Elfsbacka Schmöller 2022; Ma and Zimmermann 2023). A comprehensive review of this literature can be found in Cerra et al. (2023).

Empirical evidence on the long-run impact of temporary shocks: Cerra and Saxena (2005) provide evidence on the permanent impact of shocks on the level of output. Similarly, Blanchard et al. (2015) show how roughly two thirds of all recessions have been followed by lower output, and even lower growth. A large literature has emerged on the topic (a comprehensive review can also be found in Cerra et al., 2023). Recent work has highlighted the asymmetric, non-linear properties of these phenomena. Employing a panel of 24 advanced and emerging economies since 1970, Aikman et al. (2022) find a non-linear and asymmetric relationship between various business cycle events and growth in the long-run. In particular, they detect strong scarring effects originated in contractions below the 20th percentile of the distribution of annual growth rates. In the case of identified monetary policy shocks, Jordà et al. (2020) also find asymmetric effects on GDP, persisting after more than a decade. In their sample, spanning 125 years and 17 advanced economies, they attribute this phenomenon to a decreased capital stock, and lower total factor productivity. Employing the same identification strategy, I show in Amador (2022) how technological adoption and human capital are only affected in the long-run by contractionary monetary policy. These works add important detail to a growing empirical literature revealing hysteresis effects from a variety of shocks (e.g., Furlanetto et al. 2021; Antolin-Diaz and Surico 2022; Cloyne et al. 2022).

Output fluctuations as coordination failures: In this paper, a link is established between two different strands of literature using coordination failures as the source of aggregate fluctuations. First, the Big Push proposed by Rosenstein-Rodan (1943), in which an economy can experience a permanent increase in output if a sufficient mass of firms adopts a new, more productive technology. This view emphasizes the role of coordination failures and complementarities in adoption, as individual firms may not find it profitable to adopt a new technology unilaterally, even though the whole economy would benefit from a coordinated adoption (see also Murphy et al. 1989; Ciccone 2002; Buera et al. 2021). Second, at least since Diamond (1982) a literature emerged characterizing recessions as coordination failures (Kiyotaki, 1988; Cooper and John, 1988; Durlauf, 1991; Schaal and Taschereau-Dumouchel, 2015). Typically these models feature multiple rational expectations equilibria, stemming from increasing returns to scale (e.g., Kiyotaki 1988), and strategic complementarities (e.g., Cooper and John 1988). Durlauf (1991) develops model in which strong complementarities can generate hysteresis effects, as no markets exist to coordinate production decisions. Thus, multiple stochastic equilibria exist, and aggregate productivity shocks can affect real activity indefinitely. The model presented here aligns with Schaal and Taschereau-Dumouchel (2015), who develop a theory based on coordination failures that accounts for large shocks pushing the economy into quasi-permanent recessions. The key mechanism for their result is demand complementarities, which provide a coordination motive, and feedback from aggregate demand to production decision originated in variable capacity utilization.

2 Model

In this section, I introduce the model and some results derived from it. First, I describe the setting and agents. Next, I focus on the features of the technological choice problem of individual intermediate input producers. This problem is solved sequentially. That is, conditional on a choice of technology, prices and quantities are derived. Then these prices and quantities are used to compare profits under the relevant alternatives. The optimal policy will be to adopt a the costly technology when the productivity shock is high, and to abandon it when is low. However, this will only be the case for sufficiently large shocks, as small shocks cannot trigger a change in the firm's technological choice. Finally, I define equilibrium, and characterize the effects of productivity shocks on aggregate technological choices.

2.1 Setting

Time is discrete and goes on forever. The economy is composed of a representative household, a final consumption good representative firm, and an intermediate goods sector. Each intermediate good producer is a monopolist of their own variety, and there is an unit mass of them. They purchase inelastically supplied labor from the household. Importantly, intermediate producers choose whether to operate in the modern sector, using a new more productive technology with increasing returns to scale, or in the traditional sector, where they use a less productive constant returns to scale technology. The modern sector exhibits increasing returns to scale due to the fact that each firm has to pay a fixed adoption cost j in the period they enter use the modern sector technology, and a fixed continuation cost k every period they remain operating it. These costs are heterogeneous across firms, and have to be paid in units of the final good. Firms in the modern sector have the option to abandon the new technology, and use the traditional alternative instead. Both the traditional and modern intermediate producers are subject to an unexpected aggregate autoregressive productivity shock occurring only in period one. A constant elasticity of substitution aggregator combines all intermediate inputs into a final consumption good. Profits from the sale of intermediate inputs go to the household.

2.2 Households

The preferences of the representative household are given by

$$\sum_{t=0}^{\infty} \beta^t C_t,$$

where $\beta \in (0,1)$ is the discount factor, $C_t \geq 0$ is consumption of the final good. Since there is no disutility from labor, one unit of labor is inelastically supplied every period $L_t = 1$, $\forall t$. The household takes prices as given, and it faces the following sequence of budget constraints

$$P_t C_t = W_t + \Pi_t,$$

where P_t is the price of the final good, W_t the wage rate, and Π_t the profits it receives from intermediate firms.

2.3 Final goods producers

The final consumption good is produced by a perfectly competitive representative firm, that combines a continuum of differentiated intermediate goods using a constant elasticity of substitution production function

$$Y_t = \left(\int_k \int_j \left[y_t\left(j,k\right) \right]^{\frac{\sigma-1}{\sigma}} \mathrm{d}j \mathrm{d}k \right)^{\frac{\sigma}{\sigma-1}},$$

where each variety is indexed by the pair (j,k). The pair (j,k) identifies intermediate input producers that—as explained below—are identical in every way except for their adoption cost j, and their continuation cost, k, which are heterogenous across firms. Thus, it is convenient to index intermediate inputs with (j,k). Accordingly, $y_t(j,k)$ is the input of variety (j,k) used in the production of the final consumption good Y_t in period t. The elasticity of substitution between intermediate input varieties is denoted $\sigma > 2$. Profit maximization taking P_t and $p_t(j,k)$ as given yields the usual demand curve for each variety, and price for the final good;

$$y_t(j,k) = \left(\frac{p_t(j,k)}{P_t}\right)^{-\sigma} Y_t \tag{1}$$

Henceforth, the final consumption good will be the numéraire (i.e., $P_t = 1$).

2.4 Intermediate goods producers

Firms in the intermediate sector can choose to operate using a constant returns to scale traditional ($\mathcal{T}$ for shorthand) technology or a more productive modern (abbreviated as $\mathcal{M}$) technology with increasing returns to scale. If a traditional firm decides to modernize in t, it will have to pay a fixed adoption cost j, and a continuation cost k every subsequent period, as long it continues to operate in the modern sector. Both j and k are denoted in units of the final good.[4] I assume that $j > k$, a situation in which replacing the traditional technology entails substantial upfront costs and supplementary investments in the early stages of adoption. The resources used to pay for j and k are in units of the final consumption good, and are not used for any other purpose. As mentioned before, both the adoption and continuation costs are heterogeneous across firms, and their joint density is denoted $\delta(j,k)$. Since firms within each sector are identical in every way, except for their technological costs, it is convenient to index them using (j,k). Let $\gamma_t(j,k) = 1$ if firm (j,k) is operating in the $\mathcal{M}$-sector (and using the new technology), and $\gamma_t(j,k) = 0$ if the firm is operating in the $\mathcal{T}$-sector (and using the traditional technology). The production function of intermediate good firm (j,k) is given by

$$y_t(j,k) = \begin{cases} \exp(\epsilon_t)\,\mathcal{A} l_t(j,k) & \text{if } \gamma_t(j,k) = 1 \\ \exp(\epsilon_t)\, l_t(j,k) & \text{if } \gamma_t(j,k) = 0 \end{cases}, \tag{2}$$

where $y_t(j,k)$ is the output of firm (j,k), and $l_t(j,k)$, it's corresponding labor input. Regardless of their technology, all firms are subject to an unexpected autoregressive productivity shock, which occurs once in period $t = 1$, such that $\epsilon_t = \psi^t \varepsilon_1$, where $\varepsilon_1 \sim \mathcal{N}(0, \nu^\epsilon)$ $\forall t \geq 1$,

[4] The timing of events is the model is (i) intermediate firms choose their technology, prices, and quantities, (ii) the final good is produced, and (iii) the intermediate sector firms purchase j or k units of the final good (if applicable).

and $\psi \in (0,1)$. If the firm operates in the $\mathcal{M}$-sector then it has a higher productivity than if it did in the $\mathcal{T}$-sector, as $\mathcal{A} > 1$. Intermediate producers take the wage W_t as given. Given a technological choice $s \in \{\mathcal{T}, \mathcal{M}\}$, firms maximize gross profits (i.e., before paying technological costs, j or k);

$$\pi_t^s(j,k) = \max_{y_t^s(j,k), p_t^s(j,k), l_t^s(j,k)} p_t^s(j,k)\, y_t^s(j,k) - W_t l_t^s(j,k) \tag{3}$$

subject to demand curve (1) and production technology (2). Note that, as fixed costs do not affect marginal decisions, prices and quantities will be independent of j and k. Moreover, within each sector prices and quantities will be the same across all firms.

Proposition 1. *(Within sector symmetry of prices and quantities): Prices, $p_t^s(j,k)$, and quantities, $y_t^s(j,k)$, will be symmetric across all firms in each sector $s \in \{\mathcal{T}, \mathcal{M}\}$.*

The proof (which can be found in Appendix B) follows from the fact that, conditional on demand and technological choices, firms choose the price that maximizes gross profits (i.e., before paying j or k, if applicable). As prices, and quantities are proportional to marginal cost, they depend only on parameters, aggregate quantities, and the aggregate productivity shock ϵ_t. Thus, all firms within each sector, $\mathcal{T}$ and $\mathcal{M}$, will be symmetrical in prices and quantities. We can ignore the subindices j and k, and prices and quantities in each sector can be written as;

$$p_t^{\mathcal{T}} = \frac{\sigma}{\sigma-1}\frac{W_t}{\exp(\epsilon_t)},$$

$$p_t^{\mathcal{M}} = \frac{\sigma}{\sigma-1}\frac{W_t}{\mathcal{A}\exp(\epsilon_t)},$$

$$y_t^{\mathcal{T}} = \left(\frac{\sigma}{\sigma-1}\frac{W_t}{\exp(\epsilon_t)}\right)^{-\sigma} Y_t,$$

$$y_t^{\mathcal{M}} = \left(\frac{\sigma}{\sigma-1}\frac{W_t}{\mathcal{A}\exp(\epsilon_t)}\right)^{-\sigma} Y_t.$$

Note firms operating the $\mathcal{M}$-sector technology take advantage of their higher productivity to sell more of their variety, and at a lower price than firms in the $\mathcal{T}$-sector.

2.5 Choice of technology

Proposition 1 implies that given a value of the autoregressive shock ϵ_t, wages W_t and aggregate demand Y_t, the firm's problem can be reduced to choosing which technology to operate. First, note that by replacing prices, quantities, and Equation (2) in (3), and by Proposition 1, it follows that profits in the traditional sector are given by

$$\pi_t^{\mathcal{T}} = \frac{(\sigma-1)^{\sigma-1}}{\sigma^{\sigma}} \left(\frac{\exp(\epsilon_t)}{W_t}\right)^{\sigma-1} Y_t. \tag{4}$$

Note $\pi_t^{\mathcal{T}}$ depends on Y_t. This implies that there is feedback from aggregate demand in adoption decisions. That is, the adoption of the technology by one individual firm implies it will produce a higher quantity of their variety. Due to complementarity, this also increases demand for the products of other firms, increasing their profits, and (as shown below) the likelihood that others will adopt as well.

Equation (4) can be further simplified, using labor market clearing, so that it is only a function of the productivity shock and the share of firms operating in the $\mathcal{M}$-sector.

Proposition 2. *: Profits in the traditional sector, $\pi_t^{\mathcal{T}}$, can be written as a function of ϵ_t and m_t only. In particular;*

$$\pi_t^{\mathcal{T}} = \frac{exp(\epsilon_t)}{\sigma} \left[m_t \mathcal{A}^{\sigma-1} + (1-m_t)\right]^{\frac{2-\sigma}{\sigma-1}} \tag{5}$$

A proof of Proposition 2 can found in Appendix B. Equation (5) exhibits some interesting properties. First, profits in the $\mathcal{T}$-sector are independent of technological costs, which means all firms in that sector make the same profits (from Proposition 1). Second, $\pi_t^{\mathcal{T}}$ depends on the weighted average of the labor productivities of firms in the $\mathcal{M}$ and $\mathcal{T}$ sectors (i.e., $\exp(\epsilon_t)[m_t\mathcal{A}^{\sigma-1} + (1-m_t)]$). Third, $\pi_t^{\mathcal{T}}$ is decreasing in m_t, as $\sigma > 2$. This can be interpreted as follows. Adopting the new technology induces firms to produce a higher quantity and sell at lower prices. Because intermediate inputs are substitutable among among each other, a higher adoption share lowers profits for all firms, as adopters take market share away from higher priced firms in the $\mathcal{T}$-sector.

Similarly, the profits of firms using the $\mathcal{M}$-sector technology can also be written as a function solely depending on ϵ_t and m_t. Note profits in the $\mathcal{M}$-sector net of technological costs, denoted $\pi_t^{\mathcal{M}}(j,k)$, are given by

$$\pi_t^{\mathcal{M}}(j,k) = \begin{cases} p_t^{\mathcal{M}} y_t^{\mathcal{M}} - W_t l_t^{\mathcal{M}} - j & \text{if } \gamma_t(j,k) = 1 \text{ and } \gamma_{t-1}(j,k) = 0 \\ p_t^{\mathcal{M}} y_t^{\mathcal{M}} - W_t l_t^{\mathcal{M}} - j & \text{if } \gamma_t(j,k) = 1 \text{ and } \gamma_{t-1}(j,k) = 1 \end{cases}.$$

As in the case of profits in the $\mathcal{T}$-sector, replacing prices, quantities, using Equation (2), and by Proposition 1, it follows that profits for firm (j,k) are given by

$$\pi_t(j,k) = \begin{cases} \mathcal{A}^{\sigma-1}\pi_t^{\mathcal{T}} - j & \text{if } \gamma_t(j,k) = 1 \text{ and } \gamma_{t-1}(j,k) = 0 \\ \mathcal{A}^{\sigma-1}\pi_t^{\mathcal{T}} - k & \text{if } \gamma_t(j,k) = 1 \text{ and } \gamma_{t-1}(j,k) = 1 \\ \pi_t^{\mathcal{T}} & \text{if } \gamma_t(j,k) = 0 \end{cases}. \tag{6}$$

Equation (6) highlights the key forces that determine technological choices. As the firm can either start in one sector or the other (either $\gamma_{t-1}(j,k) = 0$ or $\gamma_{t-1}(j,k) = 1$). To maximize profits, the firm compares two functions of the productivity shock; one corresponding to $\mathcal{T}$-sector profits, and another one representing $\mathcal{M}$-sector profits. The latter features a higher slope and a lower intercept than the former. These comparisons represent the choice to either adopt the new technology and leave the $\mathcal{T}$-sector, or abandoning it and moving out of the $\mathcal{M}$-sector. Note that the productivity shock ϵ_t positively affects $\pi_t^{\mathcal{T}}$. For low values of the productivity shock, $\pi_t^{\mathcal{T}}$ will be higher than either $\mathcal{A}^{\sigma-1}\pi_t^{\mathcal{T}} - j$ or $\mathcal{A}^{\sigma-1}\pi_t^{\mathcal{T}} - k$. Conversely, for high values of ϵ_t, $\pi_t^{\mathcal{T}}$ will be lower than $\mathcal{A}^{\sigma-1}\pi_t^{\mathcal{T}} - j$. Thus, when ϵ_t is high, profits are higher using the $\mathcal{M}$-sector technology, and when it is low the $\mathcal{T}$-sector alternative is more profitable. This follows from increasing returns to scale in the $\mathcal{M}$-sector; a high productivity shock incentivizes firms to pay the adoption cost j, in order to take advantage of economies of scale, and save on production costs. Similarly, when the productivity shock is negative, the firm has incentives to save on average costs by abandoning the technology, and charging a higher price for their reduced output.

Now, let's consider the dynamic technological choice problem of firm (j,k). That is, given a particular realization of the shock ε_1 and an initial technological choice $\gamma_0(j,k) = \{0,1\}$, what is the optimal policy that determines the sequence $\{\gamma_t(j,k)\}_{t=1}^{\infty}$ that maximizes the discounted sum of profits? This optimal control problem can be characterized by the maximum of the value functions $V_t(\varepsilon_1, m_t, \gamma_{t-1}(j,k))$, as follows

$$V_t(\varepsilon_1, m_t, \gamma_{t-1}(j,k)) = \max_{\{\gamma_t(j,k)\}_{t=1}^{\infty}} \left\{ \begin{array}{c} \left[\mathcal{A}^{\sigma-1}\pi^{\mathcal{T}}(\varepsilon_1, m_t) - (1-\gamma_{t-1}(j,k))\, j - \gamma_{t-1}(j,k)\, k\right] + \beta V_{t+1}(\varepsilon_1, m_{t+1}, 1), \\ \pi^{\mathcal{T}}(\varepsilon_1, m_t) + \beta V_{t+1}(\varepsilon_0, m_{t+1}, 1) \end{array} \right\}. \tag{7}$$

Note that since $\epsilon_t = \psi^t \varepsilon_1$ for $t > 1$, all value functions for all periods can be written as functions of ε_1 instead of ϵ_t. Equation (7) implies the firm will be using the modern technology in period t if $\mathcal{A}^{\sigma-1}\pi^{\mathcal{T}}(\varepsilon_1, m_t) - (1-\gamma_{t-1}(j,k))\,j - (\gamma_{t-1}(j,k))\,k + \beta V_{t+1}(\varepsilon_1, m_{t+1}, 1) > \pi^{\mathcal{T}}(\varepsilon_1, m_t) + \beta V_{t+1}(\varepsilon_1, m_{t+1}, 0)$, and the traditional technology if $\mathcal{A}^{\sigma-1}\pi^{\mathcal{T}}(\varepsilon_1, m_t) - (1-\gamma_{t-1}(j,k))\,j - (\gamma_{t-1}(j,k))\,k + \beta V_{t+1}(\varepsilon_1, m_{t+1}, 1) < \pi^{\mathcal{T}}(\varepsilon_1, m_t) + \beta V_{t+1}(\varepsilon_1, m_{t+1}, 0)$. The optimal[5] adoption-abandonment strategy can be stated in terms of the value of ε_1 for which the firm is indifferent between alternatives. These critical values $\varepsilon^{\text{adopt}}(j,k)$, and $\varepsilon^{\text{abandon}}(j,k)$, partition the state-space into two decision regions, and are defined implicitly by;

$$\varepsilon^{\text{adopt}}(j,k) : \mathcal{A}^{\sigma-1}\pi^{\mathcal{T}}\left(\varepsilon^{\text{adopt}}, m_t\right) - j + \beta\left\{V\left(\varepsilon^{\text{adopt}}, m_{t+1}, 1\right)\right\} = \pi^{\mathcal{T}}\left(\varepsilon^{\text{adopt}}, m_t\right) + \beta\left\{V\left(\varepsilon^{\text{adopt}}, m_{t+1}, 0\right)\right\}, \tag{8}$$

$$\varepsilon^{\text{abandon}}(j,k) : \mathcal{A}^{\sigma-1}\pi^{\mathcal{T}}\left(\varepsilon^{\text{abandon}}, m_t\right) - k + \left\{V\left(\varepsilon^{\text{abandon}}, m_{t+1}, 1\right)\right\} = \pi^{\mathcal{T}}\left(\varepsilon^{\text{abandon}}, m_t\right) + \beta\left\{V\left(\varepsilon^{\text{abandon}}, m_{t+1}, 0\right)\right\}. \tag{9}$$

These thresholds imply that a sufficient large temporary shock can trigger a change in firm (j,k)'s technology choice, while small shocks are unlikely to do so. To demonstrate that this is the case, first is necessary to show that there exists a band of inaction.[6]

Proposition 3. *(Existence of band of inaction): If firm's (j,k) optimal policy is characterized by $\varepsilon^{adopt}(j,k)$ and $\varepsilon^{abandon}(j,k)$, then optimal strategy of the firm is such that $\varepsilon^{adopt}(j,k) > \varepsilon^{abandon}(j,k)$.*

The proof is simple, and relies on Equation (8), Equation (9), and the assumption that $j > k$. First, note that if $j = k$, then $\pi^{\mathcal{T}}\left(\varepsilon^{\text{adopt}}, m_t\right) = \pi^{\mathcal{T}}\left(\varepsilon^{\text{abandon}}, m_t\right)$, which is a contradiction. Also if $\pi^{\mathcal{T}}\left(\varepsilon^{\text{adopt}}, m_t\right) < \pi^{\mathcal{T}}\left(\varepsilon^{\text{abandon}}, m_t\right)$ that would imply $k > j$, which is also a contradiction.

Before moving on to describe how shocks may impact individual technological choices, the relationship between the shock, ε_1, and the value functions, $V(\varepsilon_1, m_{t+1}, \gamma_{t-1}(j,k))$, needs to be established. Define profits in the modern sector relative to profits in the traditional sector as

$$V^{\text{relative}}(\varepsilon_1, m_t, \gamma_{t-1}(j,k)) \equiv \left(\mathcal{A}^{\sigma-1} - 1\right)\pi^{\mathcal{T}}(\varepsilon_1, m_t) - (1-\gamma_{t-1}(j,k))\,j - (\gamma_{t-1}(j,k))\,k + \beta\left[V(\varepsilon_1, m_{t+1}, 1) - V(\varepsilon_1, m_{t+1}, 0)\right]. \tag{10}$$

[5]If for every triplet $(\varepsilon_1, m_t, \gamma_{t-1}(j,k))$ the function V_t converges to a finite number, then there exists a time invariant function $V(\cdot)$, which relates the value of the firm in period t to ϵ_t, m_t, and $\gamma_{t-1}(j,k)$. Conditions for the existence of $V(\cdot)$ are derived in Baldwin (1989). The proof requires few assumptions. In particular, the function $\pi_t^{\mathcal{T}}(\epsilon_t, m_t)$ has to be continuous in all arguments and bounded above and below. Note that since $m_t \in (0,1)$, as long as ϵ_t is bounded above and below, this assumption will hold.

[6]Or hysteresis band, in the terminology of Baldwin (1989). Similar mechanisms have been used in the international trade literature to generate hysteresis (e.g., Dixit 1989; Dumas 1989; Baldwin and Lyons 1989; Baldwin 1990; Ljungqvist 1994).

First, note that $\frac{\partial \pi^{\mathcal{T}}(\epsilon_t, m_t)}{\partial \epsilon_t} = \frac{\exp(\varepsilon_t)}{\sigma}\left[m_t \mathcal{A}^{\sigma-1} + (1-m_t)\right]^{\frac{2-\sigma}{\sigma-1}} > 0$. As the shock is autoregressive, this implies that a positive shock will weakly increase profits in the $\mathcal{T}$-sector in all subsequent periods. Also note that as $\mathcal{A}^{\sigma-1} > 1$, the continuation value of a firm in the $\mathcal{M}$-sector $(V(\varepsilon_1, m_{t+1}, 1))$, will increase more than the continuation value of a firm in the traditional sector $(V(\varepsilon_1, m_{t+1}, 0))$. Thus, $V^{\text{relative}}(\varepsilon_1, m_t)$ is a increasing function of the shock, ε_1;

$$\frac{\partial V^{\text{relative}}(\varepsilon_1, m_t)}{\partial \varepsilon_1} = \left(\mathcal{A}^{\sigma-1} - 1\right)\frac{\partial \pi^{\mathcal{T}}(\varepsilon_1, m_t)}{\partial \varepsilon_1} + \beta\left[\frac{\partial V(\varepsilon_1, m_{t+1}, 1)}{\partial \varepsilon_1} - \frac{\partial V(\varepsilon_1, m_{t+1}, 0)}{\partial \varepsilon_1}\right] > 0. \quad (11)$$

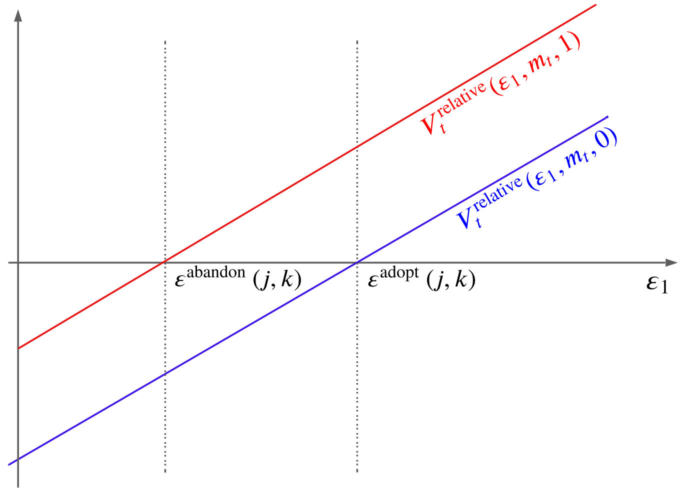
Now, lets consider how shocks of different sizes can affect the technological choices of individual firms.

Proposition 4. *(Effects of the shock of technological choices): Given that* (j,k)*'s optimal strategy is defined by a band of inaction, a sufficiently large productivity shock can trigger a permanent change in their technological choice.*

The proof is by construction of an example, summarized in Figure 1, which depicts the value of a firm in the $\mathcal{M}$-sector relative to that of a firm in the $\mathcal{T}$-sector. In the case of a firm adopting the new technology and paying the adoption cost j, the relative value is denoted $V_t^{\text{relative}}(\epsilon_1, m_t, \gamma_{t-1}(j,k) = 0)$, while for a firm already operating the new technology, and paying the continuation cost k, is denoted $V_t^{\text{relative}}(\epsilon_t, m_t, \gamma_{t-1}(j,k) = 1)$ (see Equation 10). Consider first the case of a producer initially in the $\mathcal{T}$-sector, and let the realization of the shock be such that $\varepsilon^{\text{abandon}}(j,k) < \varepsilon_1 < \varepsilon^{\text{adopt}}(j,k)$. In this case, the value of the firm will be higher staying in the $\mathcal{T}$-sector. Graphically this is reflected by the fact that the value from operating the traditional technology (the horizontal axis) is higher than the relevant outside option of adopting the modern technology (solid blue line) to the left of $\varepsilon^{\text{adopt}}(j,k)$. On the other hand, if the shock is such that $\varepsilon_1 > \varepsilon^{\text{adopt}}(j,k)$, then the value of the firm entering the $\mathcal{M}$-sector is higher that that of staying in the $\mathcal{T}$-sector (i.e., to the right of $\varepsilon^{\text{adopt}}(j,k)$ the solid blue line is above the horizontal axis). Finally, when $\varepsilon_1 \in \left(\varepsilon^{\text{abandon}}, \varepsilon^{\text{adopt}}\right)$ the firm will not change its technological choice. In the terms of Baldwin (1989), this interval is referred to as a band of inaction.

Similarly, a firm initially in the $\mathcal{M}$-sector will only abandon the modern technology if $\varepsilon_1 < \varepsilon^{\text{abandon}}(j,k)$ (i.e., the red line is below zero to left of the $\varepsilon^{\text{abandon}}(j,k)$ threshold). More sophisticated examples may be interesting, but this simple case is sufficient to prove Proposition 4. In sum, Figure 1 illustrate the basic logic behind technological choices at the level of an individual intermediate firm. Small shocks will be unlikely to trigger changes in

Figure 1: **Effect of productivity shock on technological choices**



Note: The figure depicts relative profits in the modern sector $\left(V^{\text{relative}}(\varepsilon_1, m_t, \gamma_{t-1}(j,k)) = \left(\mathcal{A}^{\sigma-1} - 1\right)\pi_t^{\mathcal{T}}(\varepsilon_1, m_t) - (1-\gamma_{t-1}(j,k))\,j - (\gamma_{t-1}(j,k))\,k + \beta\left[V(\varepsilon_1, m_{t+1}, 1) - V(\varepsilon_1, m_{t+1}, 0)\right]\right)$ for a firm adopting in period t and its currently paying the adoption cost j (the blue line, denoted $V^{\text{relative}}(\varepsilon_1, m_t, 0)$), and a firm which adopted prior to period t, and is currently paying the continuation cost k (the red line, denoted $V^{\text{relative}}(\varepsilon_1, m_t, 1)$). The horizontal represents the value of the unexpected aggregate productivity shock, ε_1.

the use of technology. However, large changes have the potential to break the firm out of the band of inaction, triggering adoption or abandonment, depending on the relevant case.

2.6 Equilibrium

So far, optimality conditions have been discussed regarding pricing, quantities and technological choices. However, it is yet to be determined how to aggregate the continuum of intermediate producers to obtain the equilibrium share of firms operating the modern technology. To being, lets define equilibrium in the economy described above. A list of equilibrium conditions can be found in Appendix B.

Definition. *(Equilibrium): Given an initial technological choice for each firm,* $\gamma_0(j,k)$*, and a sequence of shocks* $\{\epsilon_t\}_{t=1}^{\infty}$*, an equilibrium is household policies* $\{C_t(\epsilon_t, m_t), L_t(\epsilon_t, m_t)\}_{t=1}^{\infty}$*; policies for firms,* $1\{\gamma_{j,k,t}(\epsilon_t, m_t)\}_{t=1}^{\infty}$*,* $\left\{y_t^{\mathcal{T}}(\epsilon_t, m_t), y_t^{\mathcal{M}}(\epsilon_t, m_t)\, l_t^{\mathcal{T}}(\epsilon_t, m_t), l_t^{\mathcal{M}}(\epsilon_t, m_t)\right\}_{t=1}^{\infty}$*; prices* $\left\{p_t^{\mathcal{T}}(\epsilon_t, m_t), p_t^{\mathcal{M}}(\epsilon_t, m_t), P_t(\epsilon_t, m_t), W_t(\epsilon_t, m_t)\right\}_{t=1}^{\infty}$*; and a measure of intermediate firms operating in the* $\mathcal{M}$*-sector,* $\{m_t(\epsilon_t)\}_{t=1}^{\infty}$*, such that (i) the household maximizes utility, (ii) all intermediate producers maximize their profits net of technological costs (if applicable), (iii)* *the final good producer's optimality conditions are met, (iv) prices clear all markets, (v) the measure of firms in using the modern technology, m_t, satisfies;*[7]

$$m_t = \int_k \int_j \delta(j,k)\gamma_t\,(j,k)\,\mathrm{d}j\mathrm{d}k, \tag{12}$$

where $\delta(j,k)$ is the joint density of adoption, j, and continuation, k, costs, and;

$$\gamma_t\,(j,k) = \begin{cases} 1 & \text{if firm } (j,k) \text{ is operating the modern technology in period } t \\ 0 & \text{if firm } (j,k) \text{ is operating the traditional technology in period } t \end{cases}, \tag{13}$$

and (vi) the following resource constraints hold;

$$L_t = 1 = m_t l_t^{\mathcal{M}} + (1-m_t)\,l_t^{\mathcal{T}}, \tag{14}$$

$$Y_t = C_t + J_t + K_t. \tag{15}$$

2.7 Model implications

In order to characterize equilibrium in this economy, first define $\pi^{\text{adopt}}\,(j,k,m_t) \equiv \pi^{\mathcal{T}}\left(\varepsilon^{\text{adopt}}\,(j,k)\,,m_t\right)$ as the level of profits in the traditional sector evaluated at the value of the aggregate productivity shock that makes firm (j,k) indifferent between adopting the modern technology and remaining in the traditional sector (i.e., evaluated at $\varepsilon^{\text{adopt}}\,(j,k)$). Similarly, let $\pi^{\text{abandon}}\,(j,k,m_t) \equiv \pi^{\mathcal{T}}\left(\varepsilon^{\text{abandon}}\,(j,k)\,,m_t\right)$ be the level of profits in the traditional sector evaluated at the value of the aggregate productivity shock that makes firm (j,k) indifferent between abandoning the modern technology and remaining in it (i.e., evaluated at $\varepsilon^{\text{abandon}}\,(j,k)$). To determine in which sector firm (j,k) will operate, it suffices to compare the current value of $\pi_t^{\mathcal{T}}$ to the thresholds $\pi^{\text{adopt}}\,(j,k,m_t)$ and $\pi^{\text{abandon}}\,(j,k,m_t)$, depending on the previous status of the firm, $\gamma_{t-1}\,(j,k)$. Let $\Delta\left(\pi^{\text{adopt}}\right)$ and $\Delta\left(\pi^{\text{abandon}}\right)$ be the cumulative density functions of $\pi^{\text{adopt}}\,(j,k,m_t)$ and $\pi^{\text{abandon}}\,(j,k,m_t)$. That is, the share of firms that would adopt the technology if $\pi_t^{\mathcal{T}} = \pi^{\mathcal{T}}\left(\varepsilon^{\text{adopt}}\,(j,k)\,,m_t = \Delta\left(\pi^{\text{adopt}}\right)\right)$ and the share of firms that would abandon de technology if $\pi_t^{\mathcal{T}} = \pi^{\mathcal{T}}\left(\varepsilon^{\text{abandon}}\,(j,k)\,,m_t = \Delta\left(\pi^{\text{adopt}}\right)\right)$.

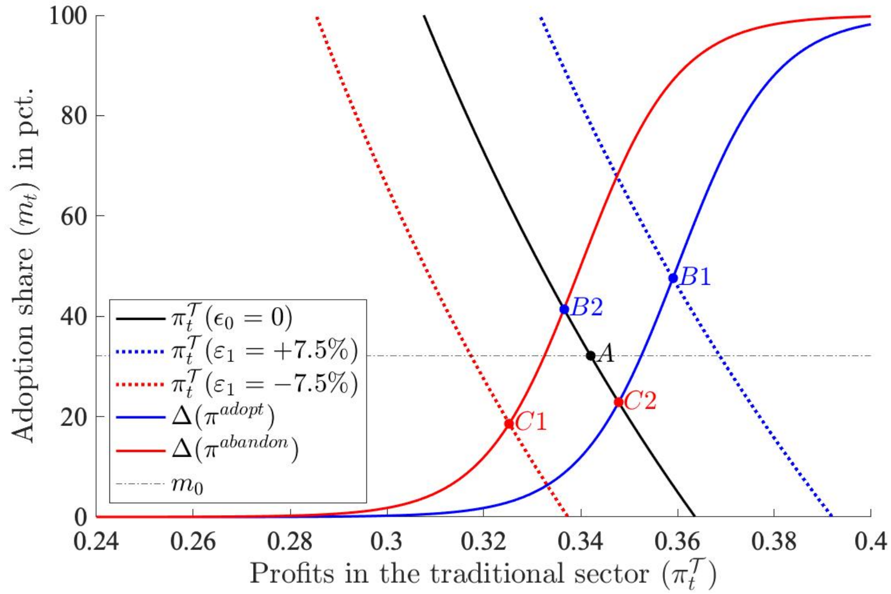
[7] Expressions similar to Equation (12) are often used in the context of mathematical models of hysteresis (see Mayergoyz 2003), Although the models presented in Mayergoyz (2003) have been mainly used in the physical sciences, some authors have emphasized their generality, and the possibility of applying them to economic phenomena (for instance, Mayergoyz and Korman 2021). The model presented here includes some elements from these works, and thus can be considered the first to apply their insights in general equilibrium. For a review of this literature see Göcke (2002).

To illustrate how shocks may affect m_t in equilibrium, consider the numerical examples contained in Figures 2, 3, and 4. These describe how m_t evolves after productivity shocks of large, intermediate, and small magnitudes, respectively (the definition of small, intermediate, and large shocks will be clarified momentarily). First, Figure 2, considers an scenario in which a large negative or positive shock occurs. Profits in the $\mathcal{T}$-sector under no shock are denoted $\pi_t^{\mathcal{T}}(\epsilon_0 = 0)$ (solid black line), after a negative 7.5% shock denoted $\pi_t^{\mathcal{T}}(\varepsilon_1 = -7.5\%)$ (dashed red line), and after a positive 7.5% shock denoted $\pi_t^{\mathcal{T}}(\varepsilon_1 = +7.5\%)$ (dashed blue line). These three curves are Equation (5) evaluated at the given values of ε_1. The solid blue line represents the cumulative density function $\Delta\left(\pi^{\text{adopt}}\right)$. That is, each each point in the solid blue curve is the percentage of firms which would adopt the $\mathcal{M}$-sector technology (vertical axis) at that level of profits in the $\mathcal{T}$-sector (horizontal axis). Similarly, the red curve represents $\Delta\left(\pi^{\text{adopt}}\right)$, the cumulative density function of $\pi^{\text{abandon}}(j,k)$ (i.e., the percentage of firms that would stay in the modern sector at that level of profits in the $\mathcal{T}$-sector). In these numerical examples, both CDFs are assumed to follow a logistic distribution.[8] Equilibria are determined by the intersection of these curves, as explained below. Note that from Proposition 3 it follows that $\Delta\left(\pi^{\text{adopt}}\right)$ will be to the right of $\Delta\left(\pi^{\text{abandon}}\right)$. This is the case because for every firm $\varepsilon^{\text{adopt}}(j,k) > \varepsilon^{\text{abandon}}(j,k)$.

First, let us consider the situation before the shock in Figure 2. The initial equilibrium in this case is given by profits in the $\mathcal{T}$-sector before the shock (solid black line). Note that any point to the left of the intersection of the solid black and solid red curves (point B2) cannot be a valid equilibrium, as profits would be too low to sustain that level of firms operating in the $\mathcal{M}$-sector. Similarly, points to the right of the intersection of the solid black and solid blue lines (point C2) cannot be equilibria, as more firms would enter the $\mathcal{M}$-sector at that level of profits. Thus, any point on the solid black line between points B2 and C2 can be the initial equilibrium. Consider for example point A; no firm would like to adopt as profits are not sufficient to induce the entry of any additional firm into the $\mathcal{M}$-sector at the initial adoption share m_0 (dashed black line), and no firm would like to abandon the technology, as profits are higher than the level that would trigger the abandonment (intersection between dashed black line and solid red curve). Assuming the initial adoption share is given by point A, when a positive shock occurs, new firms will adopt the $\mathcal{M}$-sector technology if the intersection of profits during the shock period (dashed blue line) and $\Delta\left(\pi^{\text{adopt}}\right)$ (solid blue line) is above the initial adoption share m_0 (black dashed line). The adoption share and $\mathcal{T}$-sector profits in the shock period will be higher than in the initial period. This can be seen graphically by the fact

[8]The assumption of a logistic distribution is supported by the empirical literature on the diffusion of technologies, which concludes that S-shaped curves, such as the logistic, provide a good approximation to technological adoption on the extensive margin (e.g., Comin et al., 2008).

Figure 2: **Equilibrium determination for large shocks**

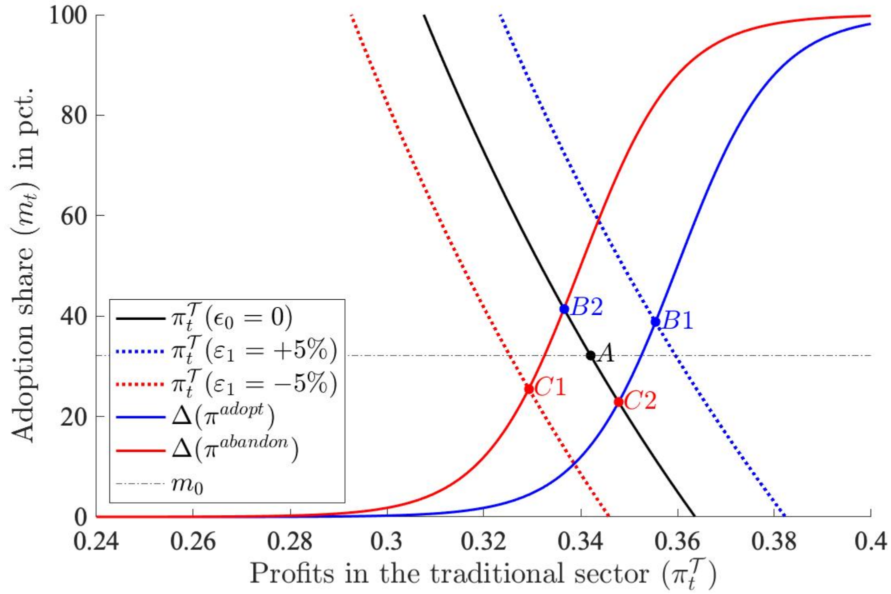


Note: The figure depicts how the equilibrium values of the adoption share can be determined for large shocks, both positive and negative. See text for details.

than point B1 is higher than point A in both the m_t and the $\pi_t^{\mathcal{T}}$ axes. As the shock dissipates, the dashed blue line will converge to the solid black line. After a number of periods the intersection of the dashed blue line and the solid red line will result in an adoption share lower than the adoption share of point B1. Then, some of the new entrants will start abandoning the $\mathcal{M}$-sector technology. This will continue until the shock completely dissipates, and the economy will converge to point B2, with a higher m_t, and lower $\pi_t^{\mathcal{T}}$ than the initial point A. The same logic applies for negative shocks. First, the negative shock reduces profits from the initial black line to the dashed red line during the shock period, inducing the exit of some firms, and moving the economy from point A to point C1. Then as the shock dissipates, the dashed red line converges to the solid black line. As profits increase this induces re-adoption for some firms. When the shock dissipates, equilibrium will be the intersection of the solid black and blue lines, at point C2, with a lower m_t and higher $\pi_t^{\mathcal{T}}$ than in the initial point A.

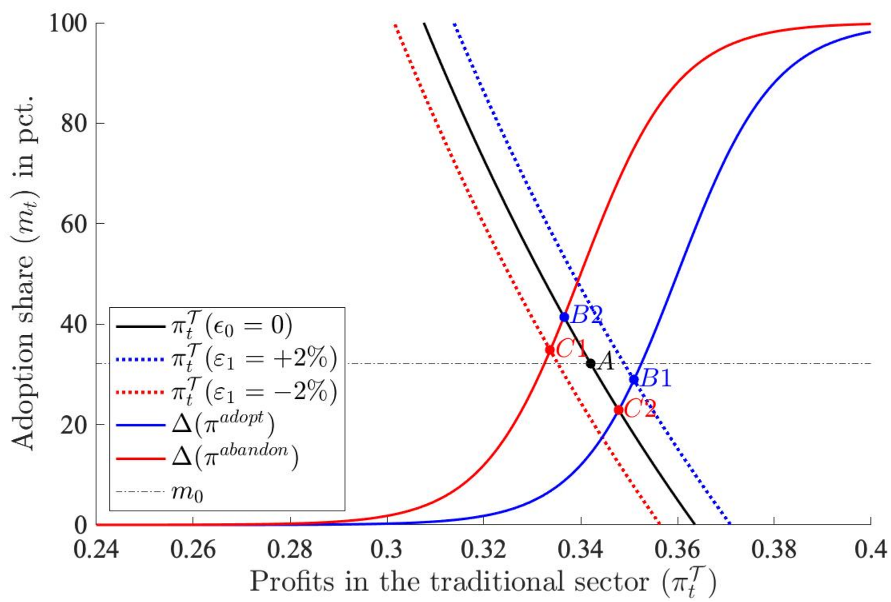
Figure 3 demonstrates how shocks of intermediate size can also have a permanent impact on the adoption share, without some firms re-adopting or re-abandoning as the shock dissipates. Note that in this case the adoption share for point B1 is below the adoption share of point B2. This implies no firm re-abandons after the shock dissipates. Similarly, the adoption share for point B1 is above that of point B2. This implies no firm re-adopts as the shock dissipates, and m_t stays at the same level as in point B1. The case of a negative shock of the same magnitude is represented by points C1 and C2.

Figure 3: **Equilibrium determination for intermediate shocks**



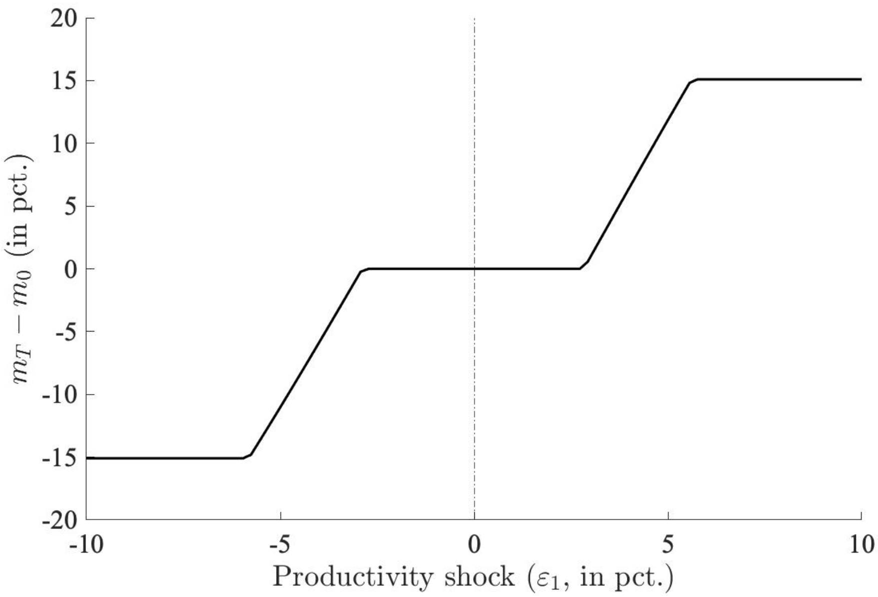
Note: The figure depicts how the equilibrium values of the adoption share can be determined for shocks of intermediate size, both positive and negative. See text for details.

Figure 4: **Equilibrium determination for small shocks**



Note: The figure depicts how the equilibrium values of the adoption share can be determined for small shocks, both positive and negative. See text for details.

Figure 5: Permanent impact on adoption share as a function of shock size



Note: The figure depicts the long-run change in the adoption share, $m_T - m_0$, as a function of the productivity shock ε_1. Se text for details.

Finally, Figure 4 illustrates that sufficiently small shocks do not have permanent effects on the adoption share. Note that B1 has an adoption share lower than the initial one, and that C1 has a higher m_t than that of point A. In both cases the shock does not trigger any change in the adoption share.

The relationship between the size of the shock and the long-run change in the adoption share (i.e., $m_T - m_0$, where T tends to infinity) illustrated in Figures 2, 3, and 4, is summarized in Figure 5, which shows productivity shocks from -10 to $+10\%$ on the horizontal axis and the corresponding change in the adoption share, denoted $m_T - m_0$ (where $T \to \infty$). First, note that small shocks have no impact on the adoption share, as in Figure 4. Second, shocks of intermediate magnitude have a linearly increasing long-run impact on the adoption share, as in Figure 3. Finally, the long run impact of shocks is bounded above and below by points B2 and C2 in Figures 2, 3, and 4.

The main implications of the model are summarized in Figure 5. First, there is a maximum sustainable level of technological adoption, characterized by the adoption share at point B2, and a corresponding minimum sustainable level, characterized by m_t at point C2 (note points B2 and C2 are common across Figures 2, 3, and 4). Any level of adoption between these two can be sustained in the long run. Second, the long-run impact of shocks of intermediate sizes will be linear on the size of the shock. Third, as in the individual case, there will be an aggregate band of inaction, in which there will be no change in technological choices. The definition of large, intermediate, and small shocks is a result of the model and it is clear from Figure 5; small shocks are those which are not large enough to break the economy out of the aggregate band of inaction, intermediate shocks are those that push the economy out of the band of inaction, triggering a long-run change in the adoption share, but without a subsequent period re-adoption or re-abandonment of the technology. In other words, in addition to the Big Push, there is also a Big Pull in which large negative shocks have to potential to bring the economy to an equilibrium with a lower level of technological adoption. Finally, large shocks are those that trigger a change in the adoption share, followed by a period of re-adoption or re-abandonment, with the economy finally settling on the maximum or minimum sustainable level of adoption, depending on the case.

Although the results of Figure 5 are intuitive, it is important to highlight that they constitute only an particular example of how the relationship between the shock and the adoption share may look in the long-run. It may well be the case that shocks do not affect the share of firms using the $\mathcal{M}$-sector technology in long-run if, for example, $\pi^{\text{adopt}}(j,k) = \pi^{\text{abandon}}(j,k)$ for all firms. In such scenario, the economy would always converge back to point A over time, and the shock would not have lasting effects on the adoption share m_t (other than the direct impact the shock would have if it was very persistent). Similarly, if the distribution of the shock ε_1 is bounded above and below, such that the adoption share of point B1 is always below m_0, and the adoption share of point C1 is always above m_0, then no shock will have a permanent impact, and Figure 5 would correspond to the horizontal axis (i.e., shocks are bounded to be too small to have a permanent impact).

3 Taking the model to the data

Does the data support the model outlined above? Specifically, is the the relationship between productivity shocks and long-run changes in the share of firms using costly, more productive technologies closely approximated by Figure 5? This section tackles these questions, by estimating the long-run response of technological adoption to total factor productivity (TFP) shocks using semi-nonparametric local projections.

3.1 Semi-nonparametric local projections

Let $m_{i,c,t}$ denote a variable measuring the share of firms using technology i in country c, and period t; let $\epsilon_{c,t}$ be a variable representing an aggregate TFP shock affecting country c in period t; and let $\boldsymbol{x}_{c,t}$ denote a vector of control variables. We are interested in characterizing how a TFP shock $\epsilon_{c,t}$ affects $m_{i,c,t}$ in subsequent periods $t+h$, relative to a baseline of no shock. Formally, define an impulse response as

$$\mathcal{R}(h) \equiv E\left[m_{i,c,t+h}|\epsilon_{c,t}=\varepsilon;\boldsymbol{x}_{c,t}\right] - E\left[m_{i,c,t+h}|\epsilon_{c,t}=0;\boldsymbol{x}_{c,t}\right]; \quad h=\{0,1,\ldots,H\}, \tag{16}$$

where ε is a particular realization of the shock. Let $m_{i,c,t}$ be characterized by the following equation

$$m_{i,c,t+h} = \alpha_{i,c} + \mu_h(\epsilon_{c,t}) + \Gamma_h \boldsymbol{x}_{c,t} + \nu_{i,c,t+h} \tag{17}$$

where $\alpha_{i,c}$ are country-technology fixed effects, $\nu_{i,c,t+h}$ the error term, and $\mu_h(\cdot)$ an unspecified function relating the TFP shock to the dependent variable at each horizon h. An estimate of the impulse response in period h would be $\hat{\mathcal{R}}(h) = \hat{\mu_h}(\varepsilon) - \hat{\mu_h}(0)$. Now, consider the following local projection (Jordà, 2005)

$$m_{i,c,t+h} - m_{i,c,t-1} = \Omega_h(\epsilon_{c,t}, \epsilon_{c,t-1}) + \Gamma_h \Delta\boldsymbol{x}_{c,t} + u_{i,c,t+h} \tag{18}$$

where $\Delta\boldsymbol{x}_{c,t} = \boldsymbol{x}_{c,t} - \boldsymbol{x}_{c,t-1}$, $u_{i,c,t+h} = \Delta\nu_{i,c,t+h}$, and $\Omega_h(\epsilon_{c,t}, \epsilon_{c,t-1}) = \mu_h(\epsilon_{c,t}) - \mu_h(\epsilon_{c,t-1})$. In order to consistently estimate $\mu_h(\cdot)$, I follow the approach of Baltagi and Li (2002) which propose approximating $\mu_h(\epsilon_{c,t})$ by the series $\lambda^q(\epsilon_{c,t})$, where $\lambda^q(\epsilon_{c,t})$ are the first q terms of a sequence of functions $\{\lambda_1(\epsilon_{c,t}), \lambda_2(\epsilon_{c,t}), \ldots\}$. This implies $\Omega_h(\epsilon_{c,t}, \epsilon_{c,t-1})$ is approximated by $\lambda^q(\epsilon_{c,t}) - \lambda^q(\epsilon_{c,t-1})$. We can now rewrite Equation (18) as

$$m_{i,c,t+h} - m_{i,c,t-1} = \beta\left(\lambda^q(\epsilon_{c,t}) - \lambda^q(\epsilon_{c,t-1})\right) + \Gamma_h \Delta\boldsymbol{x}_{c,t} + u_{i,c,t+h} \tag{19}$$

which can be consistently estimated using OLS. Then, using $\hat{\Gamma}_h$ and $\hat{\beta}$, the fixed effects $\hat{\alpha}_{i,c}$ are computed, and used in turn to obtain $\hat{z}_{i,c,t+h} = m_{i,c,t+h} - \hat{\alpha}_{i,c} - \hat{\Gamma}_h \Delta\boldsymbol{x}_{c,t} = \mu_h(\epsilon_{c,t}) + \nu_{i,c,t+h}$. Then the curve $\mu_h(\epsilon_{c,t})$ is fitted by regressing $\hat{z}_{i,c,t+h}$, employing a non-parametric regression estimator. In the following estimations, I will use a local quadratic regression based on an Epanechnikov kernel to estimate $\hat{\mu}_h(\epsilon_{c,t})$. Finally, I use $\hat{\mu}_h(\epsilon_{c,t})$ to compute the implied estimated non-parametric local projection impulse response $\hat{\mathcal{R}}(h) = \hat{\mu_h}(\varepsilon) - \hat{\mu_h}(0)$.

The estimation approach outlined above presents several advantages. First, using a non-parametric regression to estimate $\hat{\mu_h}(\varepsilon)$ does not impose any assumption on the functional form of $\mu_h(\cdot)$. While the model in Section 2 suggests a particular non-linear shape for $\mu_h(\cdot)$ (summarized by Figure 5), a priori there is no obvious way to impose assumptions on where the various kinks may be located. Moreover, if the model is not a good approximation of the impulse response in the long-run, then the true data generating process may be far from

Figure 5. In this case, any estimate relying on assumptions about $\mu_h(\cdot)$ could be subject to substantial misspecification bias. The non-parametric procedure I propose allows for verifying whether Figure 5 is a good approximation for the relationship present in the data, $\hat{\mu}_h(\cdot)$, without imposing any functional form assumptions. Second, as pointed out by Piger and Stockwell (2023), the cumulative local projection of Equation (18) has a better performance finite sample relative to an specification in levels (such as directly estimating Equation 17). In the small samples frequently employed in empirical macroeconomics, OLS estimates of impulse responses via local projections can be biased and produce incorrect inference (Kilian and Kim, 2011; Herbst and Johannsen, 2021). Piger and Stockwell (2023) show by simulation that estimating local projections in differences (as in Equation 18) substantially reduce bias and improve confidence interval accuracy with respect to an specification in levels (such as Equation 17).

To illustrate the properties of this estimator in finite samples Monte Carlo experiments are presented in Appendix C. Results suggest that when the true data generating process is linear, a non-parametric estimator can recover the linear impulse response at long horizons, even in small samples. On the other hand, if the true data generating process is non-linear (illustrated by a cubic example), a non-parametric estimator can reveal the presence of non-linearities, but at the cost of requiring more data to reduce finite sample biases. The simulations suggest that, for the sample used below, the impact of finite sample biases should be limited.

3.2 Data

This subsection describes the data used in the semi-nonparametric local projections estimation. To obtain accurate impulse responses at long horizons, time series spanning as much time as possible are required, preferably featuring a wide panel of countries to increase statistical power. Therefore, the sources were selected to maximize both the time series and panel dimensions. Three databases meeting these requirements are used in the subsequent analyses. All of them reaching back to at least 1890, and comprising 18 advanced economies.[9]

The first one is the Macrohistory database Jordà et al. (2017) covers 18 advanced economies since 1870 on an annual basis.[10] It contains nominal and real macroeconomic series, such as output, interest rates, inflation, credit, and other relevant controls.

The second source is the Cross-country Historical Adoption of Technology (CHAT) dataset from Comin and Hobijn (2009). Comin and Hobijn (2004, 2009, 2010) introduced historical

[9]These are Australia, Belgium, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, and the United States

[10]Available online at `http://www.macrohistory.net/data/`.

Table 1: **Description of technologies used**

Technology	Measure	Capital intensity	Invention date
Railroad	Km of track installed	High	1825
Telegram	Number of telegrams sent	High	1835
Telephone	Number of telephones connected	High	1875
Electricity production	Kw/Hr produced	High	1882
Electric arc steel	Tons produced	High	1907
Blast furnace steel	Tons produced	High	1950
Cell phones	Number of users	High	1973
Passenger cars	Number in operation	Low	1885
Commercial Vehicles	Number in operation	Low	1885
Tractors	Number in operation	Low	1903
Radio	Number in operation	Low	1920
TV	Number in operation	Low	1927
Computers	Number in operation	Low	1973
MRI machines	Number in operation	Low	1977

Source: Comin and Nanda (2019).

data on the adoption of major technologies over the period 1750-2003 for over 150 countries. From this database, it is possible to construct a country-technology-year panel, which measures the evolution over time of the intensity of adoption of each technology in every country. As Comin and Nanda (2019), I focus on a subset of 14 general purpose technologies, presented in Table 1.

A potential issue is the heterogeneity among the different technologies. Some of them represent technical change embodied in capital goods (e.g., number of passenger cars in circulation), others are production technologies and are measured by output (e.g., tons steel produced in electric arc furnaces), and the remainder by the number of users (e.g., number of cellphone users). How can these different variables be transformed, so they are expressed in the same units measuring technological adoption? The answer is first to take logarithms of the per capita technology variables. This effectively removes the units, transforming each variable into a technology diffusion curve measured in percent (see Comin and Hobijn 2010 for details). Second, as technologies at some point are fully adopted or become obsolete, the data is censored when the level of adoption becomes stable in each country. This is specially important in the case of this paper, as I focus on technologies abandoned for reasons unrelated to obsolescence. Finally, in the local projections specification below, technology-country fixed effects are included to account for adoption lags and constant unobserved factors.

Capital intensity is a particularly relevant characteristic of the technologies considered here. The adoption of certain technologies is more costly than others, because of the relatively high cost of the capital goods which embody them, or due to the need for a expensive support infrastructure. As the effect of shocks is likely to be different depending on the capital intensity two sets of analyses are performed; one considering all technologies, and another one taking into account only those exhibiting high capital intensity. The measure of capital intensity employed here is based on the cost of adopting each technology. For instance, the railroad is considered to be a capital intensive technology, as it requires of installing a network of tracks, stations, and equipment before being operational. As Comin and Nanda (2019) remark, capital intensity is a purely technological attribute, stable across time and space, thus facilitating the analysis (for details on how the capital intensity of each technology was obtained see Table 8 in Comin and Nanda, 2019).

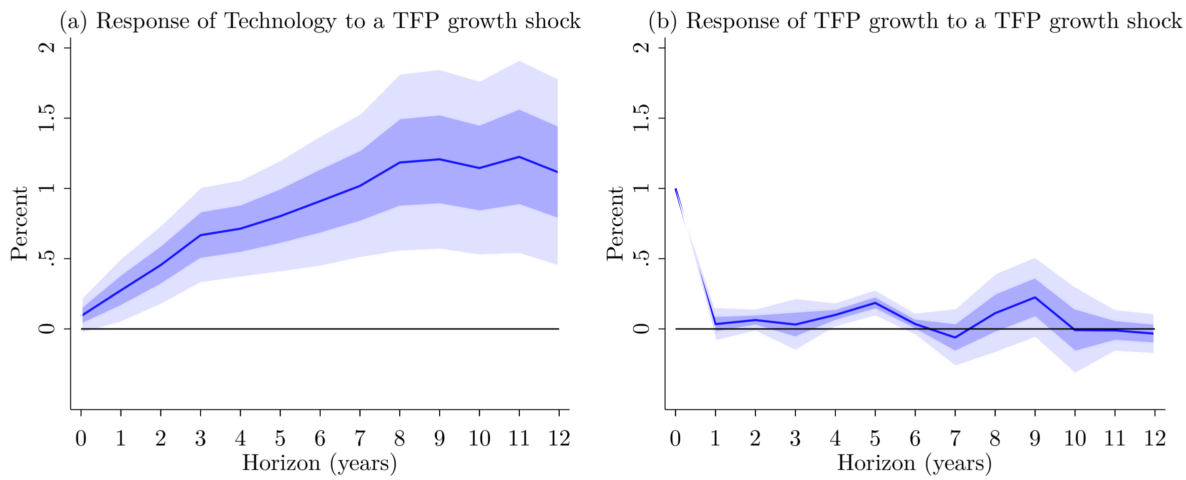
Finally, the third data source is the Long Term Productivity (LTP) database[11] of Bergeaud et al. (2016), which includes data on TFP per hour worked, for the 18 countries in sample, from 1890 to 2015. Variables in this database are that it is expressed in terms of purchasing power parity, and it uses consistent assumptions across countries and time for the construction of the series. Thus, the TFP shocks used in the estimation are comparable across countries. TFP is computed as the ratio of GDP to an aggregation of the two considered production factors, capital, and labor. The capital stock is computed by the perpetual inventory method from gross capital formation data on machinery, equipment, and buildings, each with is own depreciation assumptions (for details see subsection 2.2 in Bergeaud et al. 2016). Assuming a Cobb-Douglas production function, then $TFP_t^{LTP} = \frac{GDP_t}{K_{t-1}^{\alpha} L_t^{\beta}}$, where $\alpha + \beta = 1$. The parameters α and β represent output elasticities with respect to different factors. Thus, these can be estimated by their share of their remuneration on total income. As in the sample labor costs represent around two thirds of income, it is assumed that $\alpha = 0.3$.

3.3 Estimation results

This subsection presents linear and semi-nonparametric local projection estimates of the response of technological adoption to TFP shocks. The semi-nonparametric local projection estimates are based on Equation (18), and use the method outlined in Section 3.1 to obtain $\hat{\mu}_h(\epsilon_{c,t})$. Then the estimated impulse response is computed as $\hat{\mathcal{R}}(h) = \hat{\mu}_h(\varepsilon) - \hat{\mu}_h(0)$. The shock variable is the log difference of TFP from the LTP Database. Standard linear local projection estimates are also presented as a benchmark. The vector of controls, $\Delta \boldsymbol{x}_{c,t}$. includes up to two lags of the dependent and impulse variables, consumer price index, investment as a percentage of GDP, government expenditure as a percentage of GDP, population, and the long and short term interest rates (variables in percentage are included in first differences, while all others are in log first differences). The shock variable is trimmed at the 1st and 99th

[11] Available online at http://www.longtermproductivity.com.

Figure 6: Linear response of technological adoption to a TFP shock



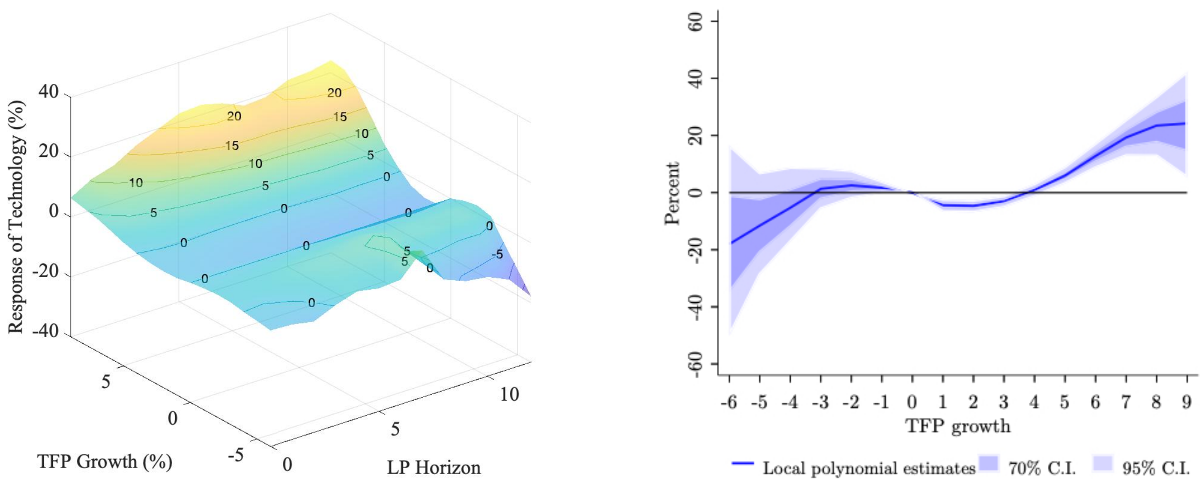
Note: One and two standard deviation confidence bands for each estimate shown as shaded blue areas. Standard errors clustered at the technology and country levels in Panel (a) and at the country level only in Panel (b). Panel (a) sample: 1893-2012 (5702 observations). Panel (b) sample: 1894-2019 (1124 observations). See text for details.
Source: Author's calculations

percentiles, to avoid estimating $\mu_h(\epsilon_{c,t})$ in regions where data is sparse. All standard errors are computed clustering at the technology and country levels.

Before presenting semi-nonparametric local projection results, two preliminary linear local projections are presented. Two facts are worthy of mention. First, a linear estimate of Equation (18) (that is, $\mu_h(\epsilon_{c,t}) = \delta^h \epsilon_{c,t}$) is presented in Figure 6, Panel (a). The accumulated IRF estimates reveal that 12 years after impact a 1 percent shock to TFP growth is associated to an increase in adoption of about 1 percent. These estimates are symmetrical, which implies a negative one percent shock would result in a minus one percent impact on adoption, 12 years after impact. Additionally, regardless of the size of the shock, the impact would always be determined by the same coefficient (at the relevant horizon). Second, it is important to verify that the long-run responses in Panel (a) are not a simple mechanical result of an unusually persistent response of TFP growth to its own shock. Panel (b) shows that TFP growth returns to zero in the year after impact. This verifies that the long-run responses in Panel (a) are not a mechanical result of a very persistent response of TFP to its own shock.

Now, lets verify if there are any non-linearities present in the impulse response. Figure 7 shows the non-linear responses of technological adoption to a TFP shock, estimated by semi-nonparametric local projections. Panel (a) presents a 3D-surface plot of $\hat{\mathcal{R}}(h) = \hat{\mu_h}(\varepsilon) - \hat{\mu_h}(0)$, for each horizon $h = \{0, 1, 2, \ldots, 12\}$ and value of the TFP shock, while Panel (b) shows the response 12 years after impact, along with bootstrapped confidence bands clustered at the

Figure 7: $\hat{\mathcal{R}}\,(h)$ **- Full sample**



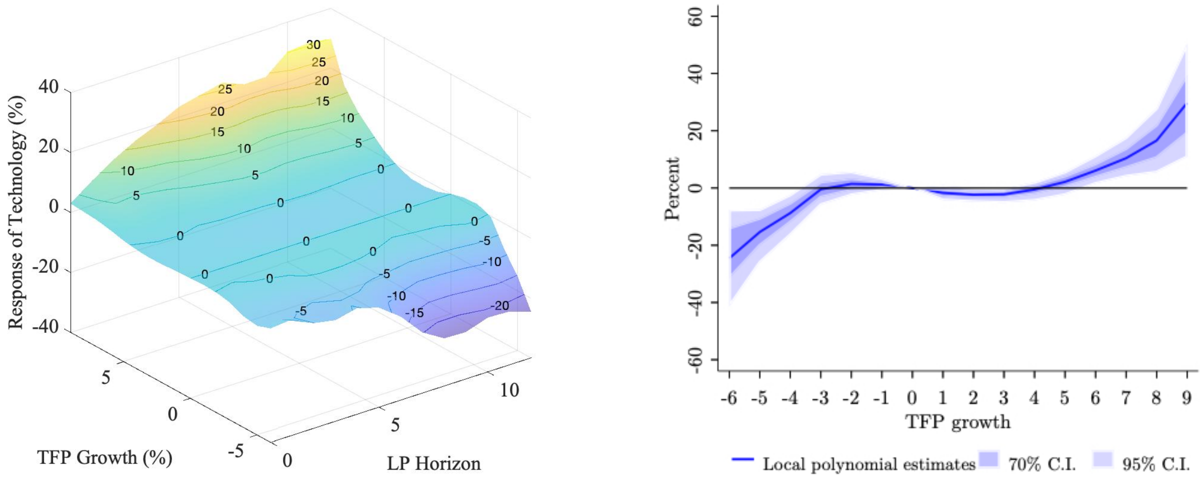
Panel (a) Surface plot **Panel (b)** $\hat{\mathcal{R}}\,(12)$

Note: Response of technological adoption to a TFP shock. Local polynomial regression based on an Epanechnikov kernel. The shock variable was trimmed at the 1st and 99th percentiles. Sample 1893-2002 (5640 observations). See text for details.
Source: Author's calculations

technology and country level. The main result is that at all horizons large shocks have a substantially larger impact than small shocks, with the latter having responses closer to zero).

To test if the cost of adoption is the mechanism driving the non-linearities, Figure 8 presents the same estimates as Figures 7, but focusing on the sub-sample of high capital intensity technologies (as identified by Comin and Nanda, 2019). As before, Panel (a) presents a surface plot with the impulse responses at $h = \{0, 1, 2, \ldots, 12\}$, while Panel (b) exhibits the response 12 years after impact, and the corresponding bootstrapped confidence bands clustered at the technology and country level. Both panels show steeper slopes for larger shocks, and values closer to zero from small shocks, relative to Figure 7. Although this sub-sample comprises roughly half of the sample, confidence bands are noticeably tighter, indicating increased precision in the estimates. Additionally, both large positive and negative values are statistically different from zero. This suggests that in the data, the mechanism explaining the non-linearities in the response of technological adoption to TFP shocks matches that of the model presented above; the costs of initially implementing technologies into production processes.

Figure 8: $\hat{\mathcal{R}}(h)$ **- High capital intensity sample**



Panel (a) Surface plot **Panel (b)** $\hat{\mathcal{R}}(12)$

Note: Response of technological adoption to a TFP shock. Local polynomial regression based on an Epanechnikov kernel. The shock variable was trimmed at the 1st and 99th percentiles. Sample 1893-2002 (2782 observations). See text for details.
Source: Author's calculations

4 Policy

Using the model and estimates presented above, this section aims to verify if—after a negative shock—it is optimal for a central planner to subsidize the cost of operating the modern technology. Since output and consumption are increasing in the share of firms using the more productive technology, this policy has the potential to avoid the loss of the economy's productive potential after a negative TFP shock. However, would this policy be welfare increasing? I answer this question in a simplified three period version of the model, in which I compare the representative consumer's welfare in a decentralized economy and under central planner choosing to subsidize the continuation cost k of all firms in the modern sector during the shock period. The shorter time horizon simplifies the computation of the adoption and abandonment thresholds (i.e., $\pi^{\text{adopt}}(j,k)$, and $\pi^{\text{abandon}}(j,k)$) as functions of j and k, for each firm (j,k).[12] The social planner funds the technological costs of firms by levying a lump sum tax from the household. Thus the budget constrain in the shock period $(t=1)$ will be

$$C_1(\varepsilon_1, m_0) = W_1(\varepsilon_1, m_0) + m_0\mathcal{A}^{\sigma-1}\pi_1^{\mathcal{T}}(\varepsilon_1, m_0) + (1-m_0)\,\pi_t^{\mathcal{T}}(\varepsilon_1, m_0) - \tau \tag{20}$$

[12]Specifically, the shorter time horizon allows the continuation values to be easily computed for each firm (j,k), this allows for computing the corresponding thresholds, $\pi^{\text{adopt}}(j,k)$, and $\pi^{\text{abandon}}(j,k)$ directly from Equations 8 and 9.

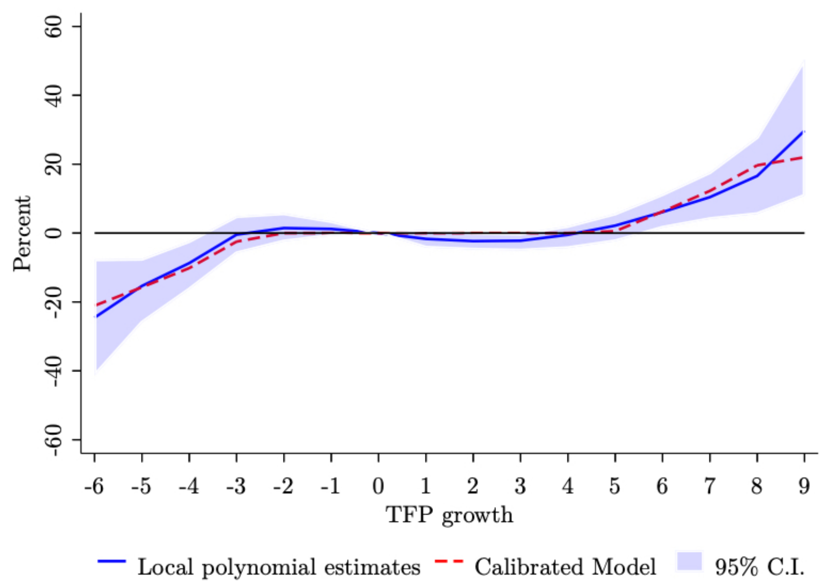
where m_0 is the initial ($t = 0$) adoption rate, ε_1 the TFP shock (which reverts back to zero in period 2), and τ a lump sum tax. To fund this policy, the lump sum tax should satisfy

$$\tau = K_1 = \int_k \int_j \gamma_{j,k,1} \gamma_{j,k,0} k \mathrm{d}j \mathrm{d}k. \tag{21}$$

Note the social planner can potentially achieve a higher welfare than the decentralized economy. This is due to the fact that the competitive equilibrium is inefficient, because of monopolistic distortions, and the fact that firms do not internalize that adopting the modern technology increases demand for other producers, therefore failing to coordinate on the efficient level of technological adoption.

The timing of events in this simplified 3-period model is as follows. In period 1, the shock occurs, potentially reducing the share of firms operating the modern technology, m_1. In period 2, the shock completely disappears. In this period some firms may readopt and pay the cost j. Finally, in period 3 the firms which re-entered the modern sector switch to paying the continuation cost k.

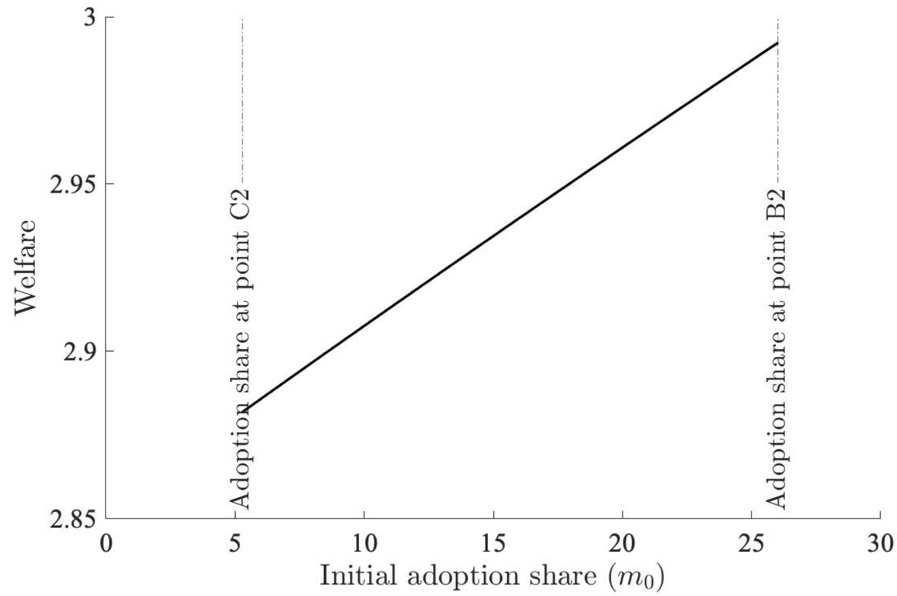
Figure 9: $\hat{\mathcal{R}}\,(12)$ **- Estimates and model calibration - High Capital Intensity Sample**



Note: Local polynomial regression based on an Epanechnikov kernel. The shock variable was trimmed at the 1st and 99th percentiles. Sample 1893-2002 (2782 observations). Model calibration; $\mathcal{A} = 1.25$, $\sigma = 2.75$, $\pi^{\text{adopt}} \sim$ Logistic(Mean $=$ 0.34, Std. dev.=0.004), and $\pi^{\text{abandon}} \sim$ Logistic(Mean $=$ 0.38, Std. dev.=0.007). See text for details.
Source: Author's calculations

Results are calibrated to match the estimates of Panel (a) Figure 8. That is, the productivity parameter $\mathcal{A}$, the elasticity of substitution between intermediate input varieties, and the distribution of $\pi^{\text{adopt}}(j,k)$, and $\pi^{\text{abandon}}(j,k)$ are calibrated[13] to match the non-linear relationship between the size of TFP shocks and the response of the adoption share after 12 years (the solid blue line in Figure 8, Panel b). Figure 9 overlays the calibrated relationship and the semi-nonparametric estimates for the high capital intensity sample. The selected calibration closely matches the estimated impulse response at $h = 12$.

Figure 10: **Welfare for different initial values of the adoption share**



Note: The figure depicts welfare in the decentralized economy in the simplified 3 period version of the model, when no shock occurs, for different initial values of the share of firms using the modern technology. Parameter values are calibrated to match the empirical relationship presented in Figure 8. See text for details.

Before moving on to analyze the welfare implications of shocks under the decentralized and planner economies, it is important to note that a higher percentage of firms using the $\mathcal{M}$-sector technology is always preferred in terms of consumer welfare. Figure 10 presents welfare in the calibrated three-period version of the model for decentralized economies with different initial values of the adoption share, m_0. It depicts a situation in which no shock occurs, for values of m_0 between the minimum and maximum sustainable levels (i.e., those corresponding to points C2 and B2 in Figures 2, 3, and 4). Welfare is linearly increasing with the share of firms using the $\mathcal{M}$-sector technology. This means that the discounted sum of consumption minus aggregate continuation costs among the set of possible long-run

[13] The calibration is $\mathcal{A} = 1.25$, $\sigma = 2.75$. The joint distribution $\delta(j,k)$ is set produce the following distributions for the adoption and abandonment thresholds; $\pi^{\text{adopt}} \sim \text{Logistic}(\text{Mean} = 0.34,\ \text{Std. dev.}{=}0.004)$ and $\pi^{\text{abandon}} \sim \text{Logistic}(\text{Mean} = 0.38,\ \text{Std. dev.}{=}0.007)$. Alternative calibrations yield qualitatively similar results.

decentralized equilibria is maximized at the adoption share at point B2. This is a common characteristic in Big Push models, where coordination failures may keep the economy away from the most efficient equilibrium (e.g., Murphy et al. 1989). Since I focus on the welfare impacts of negative shocks, studying the optimal policies subsidizing adoption (potentially into an adoption share far above that of point B2) is beyond the scope of this paper.[14]

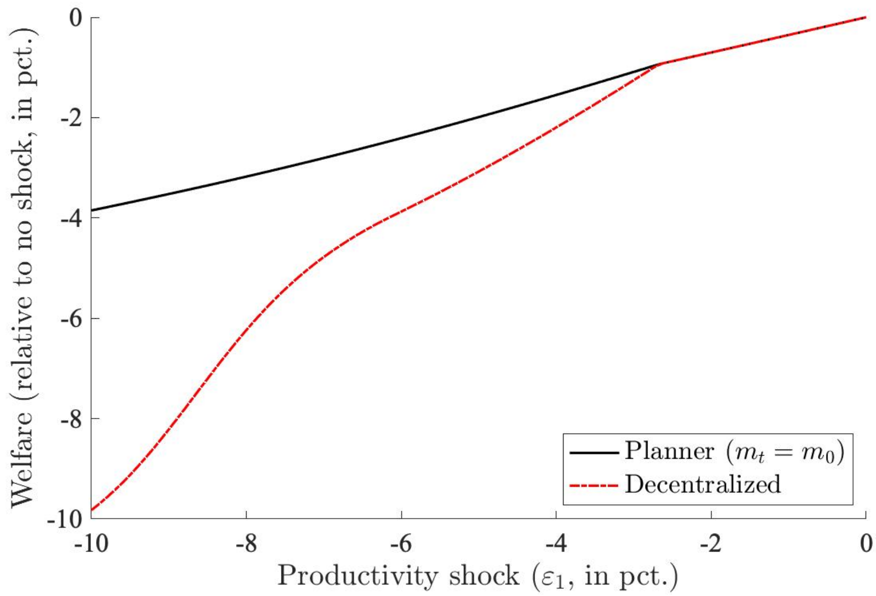
To illustrate the effect of subsidizing technological costs, I compute welfare (defined as the discounted sum of consumption minus aggregate adoption and continuation costs) for negative TFP shocks of different magnitudes. Figure 11 presents welfare in both the decentralized and planner economies, relative to a no-shock benchmark. First, note that since all shocks are negative, in every case welfare is lower than the no-shock benchmark, and decreasing in the magnitude of the shock. Second, since small shocks do not trigger a change in the percentage of firms operating the modern technology, welfare is the same in the planner and decentralized economies (this corresponds to the scenario depicted in Figure 4). Third, for shocks of intermediate magnitude, welfare is higher under the central planner. In these cases, welfare losses in the decentralized economy are not only explained by the shock, but also by the fact that the adoption share falls in period one and stays at that level even after the shock dissipates (see Figure 3). Because final good output and consumption are increasing in the share of intermediate firms operating the modern technology, consumption would be lower in all periods in the decentralized economy. Finally, for large shocks there is an additional welfare cost associated to the re-adoption of the modern technology (see Figure 2). Larger shocks will trigger more abandonments, and consequently more re-adoptions. Since by assumption $j > k$, this entails a significant cost in period 2, which is increasing on the size of the shock.

5 Conclusion

In this paper, I presented a model in which firms can use or stop using an increasing returns to scale technology, subject to heterogenous implementation and operation costs. The implications of this model shed light on how temporary shocks may have long-lasting impacts on the economy, through the endogenous technological adoption and abandonment, with a traditional constant returns to scale alternative as the outside option. The core implication of the model is that—unlike smaller episodes— large productivity shocks have the potential to affect the share of firms using the technology, both positively and negatively. However, these impacts are bounded above and below in the long-run, by maximum and minimum

[14] For an analysis of how temporary adoption subsidies can have permanent effects, see Choi and Shim (2022).

Figure 11: **Welfare analysis of subsidizing technological costs**



Note: The figure depicts welfare in the decentralized economy, and under a planner that subsidizes the cost of operating the modern technology in the shock period, in order to maintain the initial (pre-shock) level of adoption. Parameter values are calibrated to match the empirical relationship presented in Figure 8. See text for details.

sustainable levels of technological adoption. This nuanced characterization differs from works in which shocks have linearly long-lasting impacts on technological adoption, by highlighting the key role the magnitude of shocks. Additionally, I provide empirical evidence revealing that both positive and negative shocks can have enduring effects on the technological choices of firms, in a non-linear pattern that matches the model's predictions. To my knowledge, this paper is the only work showing that positive transitory shocks can have long-lasting benefits on the economy via increased technological adoption. Finally, welfare analysis using version of the model calibrated using the empirical results shows that subsidizing the cost of can be welfare improving after large negative shocks.

These analyses have some direct policy implications. Specifically, they provide a rationale for policies aiming to prevent the loss of productive potential by subsidizing the costs associated to the operation of increasing returns to scale technologies, in a context in which there is substantial amplification due to demand complementarities. However, such subsidies are welfare increasing for large shocks only, and determining if a shock falls in that category entails having information on specific implementation and operational costs associated to the technologies being adopted in the economy, how these are different across firms, and the characteristics of the alternatives they are replacing. Additionally, the fact that the long-run impact of shocks on the share of firms using the modern technology is bounded above suggests that there might be limits to policies seeking to push the economy closer to the technological frontier. Determining where those limits are located is a promising avenue for future research.

References

Aikman, David, Mathias Drehmann, Mikael Juselius, and Xiaochuan Xing, "The scarring effects of deep contractions," *Bank of Finland Research Discussion Paper*, 2022, (12). 1

Amador, Sebastián, "Hysteresis, endogenous growth, and monetary policy," *University of California, Davis, Department of Economics Working Paper*, 2022, (No. 348). 1

Antolin-Diaz, Juan and Paolo Surico, "The long-run effects of government spending," 2022. 1

Anzoategui, Diego, Diego Comin, Mark Gertler, and Joseba Martinez, "Endogenous technology adoption and R&D as sources of business cycle persistence," *American Economic Journal: Macroeconomics*, 2019, *11* (3), 67–110. 1

Baldwin, Richard, "Sunk-Cost Hysteresis," March 1989, (2911). 5, 6, 2.5

_ , "Hysteresis in trade," *Empirical Economics*, 1990, *15*, 127–142. 6

_ **and Richard Lyons**, "Exchange rate hysteresis," 1989. 6

Baltagi, Badi and Dong Li, "Series estimation of partially linear panel data models with fixed effects," *Annals of economics and finance*, 2002, *3* (1), 103–116. 3.1

Benigno, Gianluca and Luca Fornaro, "Stagnation traps," *The Review of Economic Studies*, 2018, *85* (3), 1425–1470. 1

Bergeaud, Antonin, Gilbert Cette, and Rémy Lecat, "Productivity trends in advanced countries between 1890 and 2012," *Review of Income and Wealth*, 2016, *62* (3), 420–444. 3.2

Bianchi, Francesco, Howard Kung, and Gonzalo Morales, "Growth, slowdowns, and recoveries," *Journal of Monetary Economics*, 2019, *101*, 47–63. 1

Bilbiie, Florin, Fabio Ghironi, and Marc Melitz, "Endogenous entry, product variety, and business cycles," *Journal of Political Economy*, 2012, *120* (2), 304–345. 3

Blanchard, Olivier and Lawrence Summers, "Hysteresis and the European unemployment problem," *NBER macroeconomics annual*, 1986, *1*, 15–78. 1

— , **Eugenio Cerutti, and Lawrence Summers**, "Inflation and activity–two explorations and their monetary policy implications," Technical Report No. W21726 2015. 1

Boom-Supersonic, "Boom Supersonic Announces Significant Overture Program Advances Selects key tier-one structure," 2023. 5

Brynjolfsson, Erik, Daniel Rock, and Chad Syverson, "The productivity J-curve: How intangibles complement general purpose technologies," *American Economic Journal: Macroeconomics*, 2021, *13* (1), 333–372. 1

Buera, Francisco, Hugo Hopenhayn, Yongseok Shin, and Nicholas Trachter, "Big push in distorted economies," 2021. 1

Cerra, Valerie and Sweta Saxena, "Did output recover from the Asian crisis?," *IMF Staff Papers*, 2005, *52* (1), 1–23. 1

— , **Antonio Fatás, and Sweta Saxena**, "Hysteresis and business cycles," *Journal of Economic Literature*, 2023, *61* (1), 181–225. 1

Choi, Jaedo and Younghun Shim, "Technology adoption and late industrialization," *Available at SSRN 4308957*, 2022. 1, 14

Ciccone, Antonio, "Input chains and industrialization," *The Review of Economic Studies*, 2002, *69* (3), 565–587. 1

Cloyne, James, Joseba Martinez, Haroon Mumtaz, and Paolo Surico, "Short-term tax cuts, long-term stimulus," Technical Report, National Bureau of Economic Research 2022. 1

— , **Òscar Jordà, and Alan Taylor**, "State-dependent local projections: Understanding impulse response heterogeneity," Technical Report, National Bureau of Economic Research 2023. 1

Comin, Diego and Bart Hobijn, "Cross-country technology adoption: making the theories face the facts," *Journal of monetary Economics*, 2004, *51* (1), 39–83. 3.2

— **and** — , "The CHAT dataset," Technical Report, National Bureau of Economic Research 2009. 3.2

— **and** — , "An exploration of technology diffusion," *American economic review*, 2010, *100* (5), 2031–2059. 3.2

— **and Ramana Nanda**, "Financial development and technology diffusion," *IMF Economic Review*, 2019, *67*, 395–419. 1, 3.2, 3.2, 3.3

— **, Bart Hobijn, and Emilie Rovito**, "A new approach to measuring technology with an application to the shape of the diffusion curves," *The Journal of Technology Transfer*, 2008, *33*, 187–207. 8

Cooper, Russell and Andrew John, "Coordinating coordination failures in Keynesian models," *The Quarterly Journal of Economics*, 1988, *103* (3), 441–463. 1

Diamond, Peter, "Aggregate demand management in search equilibrium," *Journal of political Economy*, 1982, *90* (5), 881–894. 1

Dixit, Avinash, "Hysteresis, import penetration, and exchange rate pass-through," *The Quarterly Journal of Economics*, 1989, *104* (2), 205–228. 6

Dumas, Bernard, "Perishable Investment and Hysteresis in Capital Formation," Technical Report, National Bureau of Economic Research 1989. 6

Durlauf, Steven, "Multiple Equilibria and Persistence in Aggregate Fluctuations," *American Economic Review*, 1991, *81* (2), 70–74. 1

Furlanetto, Francesco, Antoine Lepetit, Ørjan Robstad, Juan Rubio Ramírez, and Pål Ulvedal, "Estimating hysteresis effects," 2021. 1

Garga, Vaishali and Sanjay Singh, "Output hysteresis and optimal monetary policy," *Journal of Monetary Economics*, 2021, *117*, 871–886. 1

Ghironi, Fabio, "Macro needs micro," *Oxford Review of Economic Policy*, 2018, *34* (1-2), 195–218. 3

Göcke, Matthias, "Various concepts of hysteresis applied in economics," *Journal of economic surveys*, 2002, *16* (2), 167–188. 7

Herbst, Edward and Benjamin Johannsen, "Bias in local projections," 2021. 3.1

Heritage-Concorde, "Concorde orders and options," 2010. 5

Jordà, Òscar, "Estimation and inference of impulse responses by local projections," *American economic review*, 2005, *95* (1), 161–182. 1, 3.1

— **, Moritz Schularick, and Alan Taylor**, "Macrofinancial history and the new business cycle facts," *NBER macroeconomics annual*, 2017, *31* (1), 213–263. 3.2

__ , **Sanjay Singh, and Alan Taylor**, "The long-run effects of monetary policy," 2020, (No. W26666). 1

Kilian, Lutz and Yun Jung Kim, "How reliable are local projection estimators of impulse responses?," *Review of Economics and Statistics*, 2011, *93* (4), 1460–1466. 3.1

King, Robert, Charles Plosser, and Sergio Rebelo, "Production, growth and business cycles: II. New directions," *Journal of Monetary Economics*, 1988, *21* (2-3), 309–341. 1

Kiyotaki, Nobuhiro, "Multiple expectational equilibria under monopolistic competition," *The Quarterly Journal of Economics*, 1988, *103* (4), 695–713. 1

Ljungqvist, Lars, "Hysteresis in international trade: a general equilibrium analysis," *Journal of International Money and Finance*, 1994, *13* (4), 387–399. 6

Ma, Yueran and Kaspar Zimmermann, "Monetary Policy and Innovation," *Working paper*, 2023. 1

Mayergoyz, Isaak D, *Mathematical models of hysteresis and their applications*, Elsevier, 2003. 7

__ **and Can E Korman**, "Economic Hysteresis and Its Mathematical Modeling," *arXiv preprint arXiv:2107.10639*, 2021. 7

Murphy, Kevin, Andrei Shleifer, and Robert Vishny, "Industrialization and the big push," *Journal of political economy*, 1989, *97* (5), 1003–1026. 1, 4

Parente, Stephen and Edward Prescott, *Barriers to riches*, MIT press, 2002. 1

Piger, Jeremy and Thomas Stockwell, "Differences from Differencing: Should Local Projections with Observed Shocks be Estimated in Levels or Differences?," *Available at SSRN 4530799*, 2023. 3.1

Rosenstein-Rodan, Paul, "Problems of industrialisation of eastern and south-eastern Europe," *The economic journal*, 1943, *53* (210-211), 202–211. 1

Schaal, Edouard and Mathieu Taschereau-Dumouchel, "Coordinating business cycles," *Available at SSRN 2567344*, 2015. 1

Schmöller, Michaela Elfsbacka, "Endogenous technology, scarring and fiscal policy," *Bank of Finland Research Discussion Paper*, 2022, (13). 1

Stadler, George, "Business cycle models with endogenous technology," *The American Economic Review*, 1990, pp. 763–778. 1

Appendix A

The Concorde

An illuminating example of absolute technological abandonment can be found on the Franco-British supersonic airliner Concorde. The first two airlines scheduled to receive the plane were Air France and British Airways, whose governments had footed the bill to develop and build the aircraft. After these two, Pan Am Airlines signed non-binding purchase options for six planes in 1963. Other airlines also ordered the plane in order to compete in the new supersonic flight market. At the time of its first flight, orders from 16 airlines were received, comprising 74 planes. However, by 1973 there was significant disagreement between Aérospatiale/British Aircraft Corporation (the manufacturers) and Pan Am, which subsequently canceled its orders over its operating costs. In particular, rising oil prices were making the plane increasingly unprofitable soured its prospects (in addition to concerns about the effect it's sonic boom would have over populations on the ground). This represented a particularly large shock to the operational cost of the Concorde, as the reduction in supply by the OPEC triggered an increase in the price of crude of 214 percent between 1972 and 1974. However, this was a temporary shock, as by 1986 the real price of crude oil was back at the level it had prior to 1973 (for details, see Heritage-Concorde 2010).

Just as competitive pressures made other airlines order the plane after Pan Am, their cancellation also triggered the withdrawal of the most of the purchase agreements. By the end of 1973 all but six orders were canceled. Ultimately only Air France and British Airways operated the aircraft, until its retirement in 2003. Although this technology can be considered abandoned, there is significant hope for its re-adoption. Three start-ups are currently developing supersonic passenger jets. As of August 2022, American, United and Japan Airlines have secured options to purchase 130 supersonic jets from Boom Technology Inc. Unlike the Concorde, this time a significant portion of the orders are binding. Additionally, they plan to optimize their engines to use sustainable aviation fuel (for details, see Boom-Supersonic 2023).

The example of supersonic airliners reveals how the adoption of a non-obsolete technology can be indefinitely delayed due to a large temporary shock that makes its operation unprofitable. It also illustrates how adoption and abandonment decisions are complementary.

Appendix B

Proof of Proposition 1

Proof. Given a sequences of technological choices $\{\gamma_{j,k,t}\}_{t=0}^{\infty}$, and of aggregate productivity shocks $\{\epsilon_t\}_{t=0}^{\infty}$, the cost function for an intermediate firm in sector $s \in \{\mathcal{T}, \mathcal{M}\}$ is given by;

$$\min_{l_{j,k,t}^{s},} \left\{ W_t l_{j,k,t}^{s} + \left[\gamma_{j,k,t}(1-\gamma_{j,k,,t-1})\right] j + \left[\gamma_{j,k,t}\gamma_{j,k,t-1}\right] k \right\} \;\; \text{s.t.,}\; y_{j,k,t}^{s} = \left[(\gamma_{j,k,t}\mathcal{A}) + (1-\gamma_{j,k,t})\right] e^{\epsilon_t} l_{j,k,t}^{s}$$

The Lagrange multiplier of this problem ϕ_t^s, corresponds to the marginal cost in terms of the numéraire for firms in each sector. Note that the subscripts j and k are omitted, since these marginal costs are the same for all firms within a sector (i.e., they depend only on the wage, the new technology productivity parameter, and the aggregate productivity shock);

$$\phi_t^{\mathcal{T}} = \frac{W_t}{\exp(\epsilon_t)} \tag{22}$$

$$\phi_t^{\mathcal{M}} = \frac{W_t}{\mathcal{A}\exp(\epsilon_t)} = \frac{\phi_t^{\mathcal{T}}}{\mathcal{A}} \tag{23}$$

Given prices, real profits net of technological costs in the $\mathcal{M}$-sector can be written as;

$$\tilde{\pi}_{j,k,t}^{\mathcal{M}} = \begin{cases} (p_{j,k,t}^{\mathcal{M}} - \frac{\phi_{i,t}^{\mathcal{T}}}{\mathcal{A}}) \left(p_{j,k,t}^{\mathcal{M}}\right)^{-\sigma} Y_t - k & \text{if } \gamma_{j,k,t-1} = 1 \\ (p_{j,k,t}^{\mathcal{M}} - \frac{\phi_{i,t}^{\mathcal{T}}}{\mathcal{A}}) \left(p_{j,k,t}^{\mathcal{M}}\right)^{-\sigma} Y_t - j & \text{if } \gamma_{j,k,t-1} = 0 \end{cases} \tag{24}$$

Similarly, profits for firm in the traditional sector are given by;

$$\pi_{j,k,t}^{\mathcal{T}} = (p_{j,k,t}^{\mathcal{T}} - \phi_t^{\mathcal{T}}) \left(p_{j,k,t}^{\mathcal{T}}\right)^{-\sigma} Y_t \tag{25}$$

From (24) and (25) the following are the optimality conditions for prices;

$$p_{j,k,t}^{\mathcal{T}} = \frac{\sigma}{\sigma-1}\phi_t^{\mathcal{T}}$$

$$p_{j,k,t}^{\mathcal{M}} = \frac{\sigma}{\sigma-1}\frac{\phi_{i,t}^{\mathcal{T}}}{\mathcal{A}}$$

Replacing these prices in (1), it is possible to write expressions for the output of firms in each sector;

$$y_{j,k,t}^{\mathcal{T}} = \left(\frac{\sigma}{\sigma-1}\phi_t^{\mathcal{T}}\right)^{-\sigma} Y_t \tag{26}$$

$$y^{\mathcal{M}}_{j,k,t} = \left(\frac{\sigma}{\sigma-1}\frac{\phi^{\mathcal{T}}_t}{\mathcal{A}}\right)^{-\sigma} Y_t \tag{27}$$

Note that both prices and quantities depend only on marginal costs, which in turn depend only on aggregate quantities, and not on firm specific characteristics. Thus for every firm in sector s, prices and quantities will be the same, and the j and k subscripts can be omitted.

□

Proof of Proposition 2

Proof. As there are only two types of firms, the final good output can be written as;

$$Y_t = \left[m_t\left(y^{\mathcal{M}}_t\right)^{\frac{\sigma-1}{\sigma}} + (1-m_t)\left(y^{\mathcal{T}}_t\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \tag{28}$$

Since the consumption good is the numéraire ($P_t = 1$),

$$P_t = 1 = \left[m_t\left(p^{\mathcal{M}}_t\right)^{1-\sigma} + (1-m_t)\left(p^{\mathcal{T}}_t\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}, \tag{29}$$

Replacing the expressions for prices and marginal costs, and solving for W_t, yields the following expression;

$$W_t = \left(\frac{\sigma-1}{\sigma}\right)\exp(\epsilon_t)\left[m_t\mathcal{A}^{\sigma-1} + (1-m_t)\right]^{\frac{1}{\sigma-1}}, \tag{30}$$

Finally, replacing the wage in $\pi^{\mathcal{T}}_t$;

$$\pi^{\mathcal{T}}_t = \frac{(\sigma-1)^{\sigma-1}}{\sigma^{\sigma}}\left(\frac{1}{\left(\frac{\sigma-1}{\sigma}\right)\left[m_t\mathcal{A}^{\sigma-1} + (1-m_t)\right]^{\frac{1}{\sigma-1}}}\right)^{\sigma-1} Y_t$$

Simplifying;

$$\pi^{\mathcal{T}}_t = \frac{1}{\sigma}\left(m_t\mathcal{A}^{\sigma-1} + (1-m_t)\right)^{-1} Y_t \tag{31}$$

Replace labor demands on market clearing for labor;

$$1 = m_t\left(\frac{\sigma}{\sigma-1}\frac{\phi^{\mathcal{T}}_t}{\mathcal{A}}\right)^{-\sigma}\frac{Y_t}{\mathcal{A}\exp(\epsilon_t)} + (1-m_t)\left(\frac{\sigma}{\sigma-1}\phi^{\mathcal{T}}_t\right)^{-\sigma}\frac{Y_t}{\exp(\epsilon_t)}$$

Solving for Y_t

$$Y_t = \frac{(\exp(\epsilon_t))^{1-\sigma}}{\left[\mathcal{A}^{\sigma-1}m_t + (1-m_t)\right]}\left(\frac{\sigma}{\sigma-1}W_t\right)^{\sigma}$$

Replacing W_t;

$$Y_t = \exp(\epsilon_t)\left[m_t\mathcal{A}^{\sigma-1} + (1-m_t)\right]^{\frac{1}{\sigma-1}}$$

Replacing Y_t in $\pi_t^{\mathcal{T}}$;

$$\pi_t^{\mathcal{T}} = \frac{\exp(\epsilon_t)}{\sigma}\left[m_t\mathcal{A}^{\sigma-1} + (1-m_t)\right]^{\frac{2-\sigma}{\sigma-1}} \tag{32}$$

□

List of variables and equations

Table 2: List of variables and corresponding equations

Name	Notation	Equation
Consumption	C_t	$C_t = Y_t - J_t - K_t$
Labor supply	L_t	$L_t = 1$
Adoption operator for firm (j,k)	$\gamma_t(j,k)$	$\gamma_t(j,k) = \begin{cases} 1 & \text{if firm } (j,k) \text{ is operating in } \mathcal{M}\text{-sector in period } t \\ 0 & \text{if firm } (j,k) \text{ is operating in } \mathcal{T}\text{-sector in period } t \end{cases}$
Output of firms in $\mathcal{T}$-sector	$y_t^{\mathcal{T}}$	$y_t^{\mathcal{T}} = \left(\frac{\sigma}{\sigma-1}\phi_t^{\mathcal{T}}\right)^{-\sigma} Y_t$
Output of firms in $\mathcal{M}$-sector	$y_t^{\mathcal{M}}$	$y_t^{\mathcal{M}} = \left(\frac{\sigma}{\sigma-1}\frac{\phi_t^{\mathcal{T}}}{\mathcal{A}}\right)^{-\sigma} Y_t$
Labor demand of firm in $\mathcal{T}$-sector	$l_t^{\mathcal{T}}$	$l_t^{\mathcal{T}} = \left(\frac{\sigma}{\sigma-1}\phi_t^{\mathcal{T}}\right)^{-\sigma} \frac{Y_t}{\exp(\epsilon_t)}$
Labor demand of firm in $\mathcal{M}$-sector	$l_t^{\mathcal{M}}$	$l_t^{\mathcal{M}} = \left(\frac{\sigma}{\sigma-1}\frac{\phi_t^{\mathcal{T}}}{\mathcal{A}}\right)^{-\sigma} \frac{Y_t}{\mathcal{A}\exp(\epsilon_t)}$
Price of firm in $\mathcal{T}$-sector	$p_t^{\mathcal{T}}$	$p_t^{\mathcal{T}} = \frac{\sigma}{\sigma-1}\phi_t^{\mathcal{T}}$
Price of firm in $\mathcal{M}$-sector	$p_t^{\mathcal{M}}$	$p_t^{\mathcal{M}} = \frac{\sigma}{\sigma-1}\frac{\phi_t^{\mathcal{T}}}{\mathcal{A}}$
Real marginal cost in $\mathcal{T}$-sector	$\phi_t^{\mathcal{T}}$	$\phi_t^{\mathcal{T}} = \frac{W_t}{\exp(\epsilon_t)}$
Real marginal cost in $\mathcal{M}$-sector	$\phi_t^{\mathcal{M}}$	$\phi_t^{\mathcal{M}} = \frac{W_t}{\mathcal{A}\exp(\epsilon_t)} = \frac{\phi_t^{\mathcal{T}}}{\mathcal{A}}$
Gross profits of $\mathcal{T}$-sector firm	$\pi_t^{\mathcal{T}}$	$\pi_t^{\mathcal{T}} = \frac{\exp(\epsilon_t)}{\sigma}\left[m_t\mathcal{A}^{\sigma-1} + (1-m_t)\right]^{\frac{2-\sigma}{\sigma-1}}$
Gross profits of $\mathcal{M}$-sector firm	$\pi_t^{\mathcal{M}}$	$\pi_t^{\mathcal{M}} = \mathcal{A}^{\sigma-1}\pi_t^{\mathcal{T}}$
Consumption good price	P_t	$P_t = 1$
Wage	W_t	$W_t = \left(\frac{\sigma-1}{\sigma}\right)\exp(\epsilon_t)\left[m_t\mathcal{A}^{\sigma-1} + (1-m_t)\right]^{\frac{1}{\sigma-1}}$
Measure of firms in the $\mathcal{M}$-sector	m_t	$m_t = \int\limits_k\int\limits_j \delta(j,k)\gamma_t(j,k)\mathrm{d}j\mathrm{d}k$
Aggregate net profits	Π_t	$\Pi_t = m_t\mathcal{A}^{\sigma-1}\pi_t^{\mathcal{T}} + (1-m_t)\pi_t^{\mathcal{T}} - J_t - K_t,$
Aggregate adoption costs	J_t	$J_t = \int\limits_k\int\limits_j \gamma_t(j,k)(1-\gamma_{t-1}(j,k))j\mathrm{d}j\mathrm{d}k$
Aggregate continuation costs	K_t	$K_t = \int\limits_k\int\limits_j \gamma_t(j,k)\gamma_{t-1}(j,k)k\mathrm{d}j\mathrm{d}k$
Aggregate productivity shock	ϵ_t	$\epsilon_t = \psi^t\varepsilon_1$, where $E[\varepsilon_1] = 0$, and Std. dev. $[\nu^\epsilon]$
Aggregate output	Y_t	$Y_t = \left[m_t\left(y_t^{\mathcal{M}}\right)^{\frac{\sigma-1}{\sigma}} + (1-m_t)\left(y_t^{\mathcal{T}}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$

Appendix C

This Appendix shows a simple Monte Carlo experiment in which a semi-nonparametric local projection estimator is computed to obtain the impulse response of the following cubic data generating process

$$y_{i,t} = \rho y_{i,t-1} + \epsilon_{i,t}^3 \,, \tag{33}$$

where $\rho = 0.9$. Figure 12 shows the implied cumulative impulse response function depending on the shock $\epsilon_{i,t}$ twelve periods after impact, along with Monte Carlo averages from estimates of using a local polynomial regression based on an Epanechnikov kernel, for samples of 100, 1000, and 5000 observations. Each experiment is repeated one thousand times. Results indicate that finite sample biases are lower for the 5000 observations sample, which is a size similar to that of the sample used in the semi-nonparametric estimations of subsection 3.3.

Figure 12: Non-parametric estimates of cubic processes at a long horizon