NBER WORKING PAPER SERIES

# STRATEGIC COMPLEMENTARITIES IN A DYNAMIC MODEL OF TECHNOLOGY ADOPTION: P2P DIGITAL PAYMENTS

Fernando E. Alvarez
David Argente
Francesco Lippi
Esteban Méndez
Diana Van Patten

Working Paper 31280
http://www.nber.org/papers/w31280

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2023

We thank Isaac Baley, Andres Blanco, Diego Comin, Nicolas Crouzet, Joe Hazell, Rishabh Kirpalani, John Leahy, Claudio Michelacci, Ben Moll, Guillermo Noguera, a Ezra Oberfield, Franck Portier, Rob Shimer, Nancy Stokey, Daniel Xu, Yu Zhu, and seminar participants at EIEF, Oxford, Minnesota, Chicago Booth, NYU, Princeton, Dartmouth, Columbia, Richmond Fed, Philadelphia Fed, UC Berkeley, LSE, UCL, Duke- NCSU-UNC, Rochester, Toronto, Northwestern, IIES, Chicago Fed, UCSB LAEF, Central Bank of Costa Rica, Bank of Mexico, Fed Board, the Cowles Trade Summer Conference, the BSE Summer Forum, the Hydra conference, the Bank of Canada, the Bank of Portugal, the Rome Junior Conference (Pizzanomics), the SED Wisconsin, NBER SI, NBER EFG, and Warwick-CFM-Vienna Global Macro. This project received STEG grant No. 608. Lippi acknowledges financial support from the ERC grant: 101054421-DCS. The views expressed herein are those of the authors and do not necessarily represent the views of the Central Bank of Costa Rica (BCCR). All results have been reviewed by the BCCR to ensure that no confidential information is disclosed. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w31280

Strategic Complementarities in a Dynamic Model of Technology Adoption: P2P Digital Payments
Fernando E. Alvarez, David Argente, Francesco Lippi, Esteban Méndez, and Diana Van Patten
NBER Working Paper No. 31280
May 2023
JEL No. E04,E05,O1,O2

## ABSTRACT

This paper develops a dynamic model of technology adoption featuring strategic complementarities: the benefits of usage increase with the number of adopters. We study the diffusion of new means of payments, where such complementarities are pervasive. We show that complementarities give rise to multiple equilibria, suboptimal allocations, and study the planner's problem. The model generates gradualism in adoption, as individuals optimally wait for others to adopt before doing so. We apply the theory to the adoption of SINPE, an electronic peer-to-peer (P2P) payment app developed by the Central Bank of Costa Rica. Transaction-level data on the use of SINPE and several administrative data sets on the network structure allow us to exploit plausibly exogenous variation and to document sizable complementarities. A calibrated version of the model shows that the optimal subsidy pushes the economy to universal adoption.

Fernando E. Alvarez
University of Chicago
Department of Economics
1126 East 59th Street
Chicago, IL 60637
and NBER
f-alvarez1@uchicago.edu

David Argente
Pennsylvania State University
Department of Economics
403 Kern Building
University Park
State College, PA 16801
and NBER
dargente@psu.edu

Francesco Lippi
Luiss University
and EIEF
francescolippi@gmail.com

Esteban Méndez
Central Bank of Costa Rica
San José
Costa Rica
mendezce@bccr.fi.cr

Diana Van Patten
Yale School of Management
165 Whitney Avenue
New Haven, CT 06511
and NBER
diana.vanpatten@yale.edu

An online appendix and technical appendix is available at
http://www.nber.org/data-appendix/w31280

# 1 Introduction

Understanding the forces behind technology diffusion is important in several areas of economics (see e.g., Parente and Prescott (1994); Comin and Hobijn (2010); Stokey (2020)). While the literature has studied the role of learning in shaping adoption processes, less is known about how the process of diffusion is shaped by strategic complementarities, where one agent's benefit from adoption increases with the number of adopters. We develop a dynamic model of adoption to study the role of such complementarities in the diffusion of a new technology. The model allows us to analyze the efficiency of the equilibria and discuss optimal policy interventions.

In particular we study the diffusion of new means of payments, such as mobile money and other peer-to-peer (P2P) payment instruments, that have been recently propelled by digitization (see e.g., Economides and Jeziorski (2017); Aron (2018)) and appear in several plans for central bank digital currency (see e.g., Auer et al. (2020); Carapella and Flemming (2020)). A central element of our analysis is the presence of complementarities in adoption, an inherent feature of payment instruments. The applied literature on technology adoption has long recognized the presence of complementarities, whereby the probability that a new technique is adopted is an increasing function of the proportion of firms already using it (see Griliches (1957); Mansfield (1961)), but progress in this research area is hindered by the challenges that arise when modeling adoption dynamically –a large state space, non-linear decisions, multiple equilibria–, and by the lack of detailed data on technology diffusion. We present a model of technology adoption featuring heterogeneous agents, complementarities, and fully fledged dynamics: the agent's decision to adopt depends on the whole path of future adoptions. We analyze the conditions for equilibrium existence, multiplicity of equilibrium paths, multiplicity of stationary equilibria, and the local stability of stationary equilibria. We also characterize the planner's problem and its implementation through subsidies. We use the model to study the diffusion of SINPE, a digital platform developed and administered by the Central Bank of Costa Rica.[1] The platform was launched in May 2015 and over 60% of the adult population uses the app in 2021, with about 10% of the country's GDP transacted through SINPE. We leverage a battery of granular administrative datasets to characterize the adoption patterns and to document the presence of strong complementarities in adoption and usage.[2]

The model assumes the flow benefits of using the technology at time $t$ depend on the

[1] More precisely, the app is called "SINPE Móvil," although throughout we will be referring to it only as "SINPE," which stands for Costa Rica's National Electronic Payment System (by its initials in Spanish).

[2] See Björkegren (2018) for a related network-goods analysis using data on mobile phones adoption in Rwanda.

number of agents who have adopted the technology, $N(t)$, and on an idiosyncratic persistent random component, $x(t)$. Adoption entails a fixed cost and agents choose when to adopt taking the aggregate path of adoption as given. The model also includes an intensive margin for the usage of the technology. We show that when the idiosyncratic benefits are random the equilibrium features gradual adoption through a simple mechanism: agents wait for others to adopt.[3] This differs from previous contributions, discussed below, where gradualism is absent or assumed (e.g., by means of staggered adoption opportunities). While gradualism can also be generated by a learning mechanism, we see strategic complementarities as an inherent feature of means of payment, as well as important for its implications on optimal subsidies.[4] The optimal adoption rule is given by a time-dependent threshold value, denoted by $\bar{x}(t)$, such that adoption is optimal if $x(t) > \bar{x}(t)$. We assume that the economy starts with an initial measure of agents that have adopted the technology. Aggregation of the optimal adoption rule across agents yields a path for the fraction of agents that adopt the technology at each time $t$, $N(t)$. The equilibrium has a classic fixed point structure: the optimal decision path ($\bar{x}$) depends on the aggregate path ($N$), and vice-versa.

We obtain several theoretical results. First, we establish the monotonicity of the optimal decision rules and of the aggregation. In particular, we show that the optimal adoption rule for each agent, summarized by the threshold path $\bar{x}$, is a decreasing functional of the path of adoption $N$. The strength of this effect depends on the parameter that controls the strategic complementarity. Likewise, we show that the adoption path $N$ is a decreasing functional of the path $\bar{x}$, for any initial distribution of adopters. An equilibrium path $N$ is a fixed point of the composition of these two functionals. We use the monotonicity to show that, for a given initial condition, the set of equilibrium path is a lattice, i.e., the equilibria can be ordered in terms of their intensity of adoption, with a highest and a lowest equilibrium path. This means that when there is more than one equilibrium, the highest and lowest paths do not cross. Additionally, the equilibrium path $N$ increases as a function of the strength of strategic complementarity, as well as of the initial fraction of adopters. Second, we show that there is a critical mass of adopters $\underline{N}_0$ such that, if the initial measure of adopters is below $\underline{N}_0$, then there is an equilibrium where no one will adopt in the future. Third, we analyze the equilibrium model as a dynamic system for the cross sectional distribution of adoption. We show that besides the stationary distribution with no adoption the model has two additional interior stationary distributions, which we label low- and high-adoption.

[3] We also analyzed a model where $x$ is heterogeneous across agents but fixed through time. A key takeaway from this model is that, starting from a no adoption initial distribution, it features no dynamics; it is not a model of gradual diffusion, but one of "jumps." Instead, the stochastic model features gradual adoption given the option value of waiting for a high draw of the idiosyncratic benefit.

[4] See Cabral (1990); Reinganum (1981) for an early analysis of a dynamic equilibrium with externalities.

Fourth, we conduct a perturbation analysis with respect to the initial condition to study the stability of the interior stationary distributions.[5] We find that the high-adoption equilibrium is locally stable, while the low-adoption is unstable, a feature that leads us to discard it from the analysis. Fifth, all equilibria are socially inefficient: the reason is that agents do not internalize the fact that when they adopt they benefit all agents who already have the technology. We analyze the socially efficient dynamics of adoption by solving the planner's problem, and show how to decentralize the planner's solution using a time-varying subsidy paid to those that use the technology. Equivalently, it can be decentralized with a time varying adoption subsidy.

We then leverage a comprehensive set of data collected since SINPE was created to analyze the dynamics of adoption and usage, to document the presence of strategic complementarities, and to discipline the parametrization of the model. Data on users—both receivers and senders—can be linked to several relevant networks, including the employer-employee network, family networks, and spatial "neighborhood" networks. We document five empirical patterns that guide our modeling choices and inform the quantitative analysis. First, while firms can potentially use SINPE, over 95% of the transactions are between individuals: this fact aligns with our model where transactions are peer-to-peer. Second, we find that individuals belong to networks, as 75% of all transactions occur between coworkers, neighbors, or relatives. Third, we find that the technology diffuses gradually within the network. The last two facts motivate the importance of mechanisms delivering a gradual diffusion of the technology within a network. Fourth, there is evidence of selection at entry: users who adopted when adoption rates were low use the app more intensively, and that early adopters have higher wages and skills than those who adopt later. These patterns are consistent with our model where individuals with a high idiosyncratic benefit adopt the technology early on. Fifth, there is evidence of strategic complementarities: changes in the share of people within a network who adopt SINPE are associated with changes in the intensity with which users in that network use the app.[6] We use arguably exogenous variations in the network size due to mass layoffs to provide evidence in support of a causal relation between the share of agents who have adopted ($N$) and usage of the app.[7]

[5]This is a non trivial problem that involves the linearization of an infinite dimensional system, which we handle leveraging techniques from the Mean Field Game literature, developed in Alvarez, Lippi and Souganidis (2022b).

[6]This correlation persists across a battery of ways to define usage and networks. It also emerges after using a leave-one-out instrument and following a balanced panel of adopters to address concerns regarding mis-measurement selection.

[7]Namely, we focus on networks of coworkers and examine the effect of network changes on adoption, at the extensive and intensive margins, for workers displaced during a mass layoff. The intensive margin, in particular, follows workers who had already adopted the app when they were displaced, and can better disentangle the impact of complementarities from learning. We leverage our rich data to overcome the fact

We put together the theory and the data in a quantitative analysis where we calibrate the model using key moments from the data. To match the initial gradual diffusion of the technology, observed in each network, we combine our model of strategic complementarities with a random diffusion of information model, following the seminal work of Bass (1969). The calibration requires us to estimate the value of the parameter that governs the strength of the strategic complementarities. We do so by exploiting exogenous changes in the network of coworkers after mass the layoffs described above. We examine how both the extensive and the intensive margin of adoption respond to that variation. The intensive margin, in particular, allows us to tease out strategic complementarities from other channels. The calibrated model shows that the optimal subsidy moves the economy to 100% adoption.

**Contribution of the paper.** Altogether, we see our model as delivering four main contributions. First, a novel mechanism for the gradual diffusion of a new technology, based on the presence of strategic complementarities, whose empirical relevance we also document. Second, an in-depth analysis of multiple equilibria, multiple stationary equilibria, and the stability or lack or thereof (this is related to e.g., Cabral 1990, Matsuyama 1991). Third, the tractability of the model allows us to solve the planning problem. The problem is relevant for policy since the presence of complementarities implies that the equilibrium is not efficient. Our framework allows us to compute the optimal subsidy under which the equilibrium converges to the planner's solution. The optimal subsidy depends positively on the strength of strategic complementarities, which can be estimated from the data. Fourth, our paper also innovates in the use of *individual-level* data on adoption and on usage of SINPE, relying on individual earnings and demographic characteristics, considering relatives, neighbors, and coworkers. This provides a unique opportunity to understand the characteristics and relevant networks of each user, identify the strength of complementarities —at both the extensive and intensive margins— and the dynamics of adoption over a long time period.[8]

**Related Literature.** Several recent studies are related to our paper. Benhabib et al. (2021) model firms that can endogenously innovate and adopt a technology. They analyze the effect of these choices on productivity and balanced growth, but without conducting an analysis of the transition between stationary distributions; likewise, Buera et al. (2021) study policies that can coordinate technology adoption across firms. A closely related contribution is Crouzet et al. (2023), who develop a model with a unique equilibrium where the rate of

that people select into their networks and the reflection problem that arises when common shocks affect those in the network.

[8]This is related to other studies, summarized by Suri (2017), that relied on RCTs or shorter periods of time to analyze the patterns of adoption of electronic methods of payment.

adoption of electronic payment by *retailers* increases following an aggregate shock. Their analysis is motivated by 2016 Indian Demonetization, and exploits the variation in the intensity with which *firms* in Indian districts were exposed to the shock to examine the adoption of retailers. Unlike our model, which has heterogeneous agents and generates dynamics and gradual adoption *endogenously* (as agents wait for others to adopt before doing so), their model features homogeneous agents and a sluggish adjustment a' la Calvo (1983), generating gradual adoption through this imposed friction. Moreover, the heterogeneity in our model allows us to accommodate, not only aggregate shocks when we analyze transition dynamics in closed-form, but also dynamics after shocks that target particular types of agents; for instance, we compare the propagation after "giving the app" to people with high vs. low idiosyncratic benefits, which in turn can be mapped to observables like wages and skills.

# 2 The Model

We present a simple model for the adoption of a new technology. The economy is populated by a continuum of agents that differ in the potential benefits from adopting the technology. Let $N(t)$ denote the number of agents that have adopted the technology at time $t$. Let $x \in [0, U]$ be the idiosyncratic potential benefit of adopting, due to e.g., the agent's strength of connections. We assume that the flow benefit of the technology for an agent who adopts are given by

$$x(\theta_0 + \theta_n N(t)) \tag{1}$$

at time $t$, where $\theta_0, \theta_n > 0$ are parameters. The idiosyncratic potential $x$ follows a Brownian motion, independent across agents, with variance per unit of time $\sigma$, no drift, and reflecting barriers at $x = 0$ and $x = U$, so that $dx = \sigma dW$ where $W$ is a standardized Brownian motion. We let $c > 0$ be the fixed cost of adopting the technology. The time discount rate is $r > 0$, and we assume that with probability $\nu$ per unit of time agents die, so that the agents discount at rate $\rho \equiv r + \nu$. Agents that die are replaced by newborns without the technology and are given a random draw $x$ from the invariant density $f$ on $[0, U]$ which is uniformly distributed due to our reflecting barriers assumption, i.e., $f(x) = 1/U$.

## 2.1 Optimal Adoption Decisions

In this section we describe the optimal adoption decision as a function of the whole path of $N$, the fraction of agents that adopt the technology. Let $a(x, t)$ be the value function of an agent who uses the technology and has state $x$ at time $t$:

$$a(x,t) = \mathbb{E}\Big[\int_t^{\infty} e^{-\rho(s-t)}\left(\theta_0 + \theta_n N(s)\right) x(s) ds \;\Big|\; x(t) = x\Big] \tag{2}$$

for all $t \geq 0$ and $x \in [0, U]$. Note that the agent takes the path $N(s)$ as given.

For technical motives we assume that the path of $N(s)$ is constant at some given value $\bar{N}$ for $s > T$ where $T$ is given. All our results hold for finite but arbitrarily large $T$, and some of the results hold for $T \to \infty$. Later on we will focus on the case when $\bar{N}$ is the adoption rate corresponding to an invariant distribution for the model with $T = \infty$.

An agent with state $x$ that at time $t$ has not yet adopted has a value function $v(x,t)$ that solves the following stopping-time problem

$$v(x,t) = \max_{t \leq \tau} \mathbb{E}\left[e^{-\rho(\tau-t)}\left(a\left(x\left(\tau\right), \tau\right) - c\right) \Big| x(t) = x\right], \tag{3}$$

where $\tau$ denotes the time of the adoption and depends only on the information generated by the process for $x$'s and on calendar time.

**Discretized Model.** For future use, we introduce a discretized version of the model. It is defined by positive integers $I, J$ which determine step sizes for $t$ given by $\Delta_t = \frac{T}{J-1}$ and for $x$ given by $\Delta_x = \frac{U}{I-1}$. Thus $t \in \{\Delta_t(j-1) : j = 1, \ldots, J\}$ and $x(t) \in \{\Delta_x(i-1) : i = 1, \ldots, I\}$. The reflecting Brownian Motion, Poisson processes, and discounting are changed accordingly, following the scheme used in finite difference approximations. See Definition 3 in Appendix A for a detailed definition.

As a preliminary result, we show that the optimal adoption policy is a threshold rule:

Proposition 1. Fix a path $N$ and a time $t \in [0, T]$. If it is optimal to adopt at $(x_1, t)$, then it is also optimal to adopt at $(x_2, t)$ where $x_2 > x_1$. This holds for the continuous time as well as for the discretized version.

This proposition means that we can represent the optimal adoption rule at time $t$ as a threshold rule, $\bar{x}(t)$. The result is intuitive but non-trivial since the process for $x$ is persistent.

We denote $a_T(x) = a(x,T)$ and $v_T(x) = v(x,T)$, that depend only on the constant $\bar{N}$. We can now concentrate on the time interval $[0, T]$. In this interval we write the optimal decision rule as a function of the path $N : [0, T] \to [0, 1]$, and of the functions $a_T$ and $v_T$. Indeed, the optimal decision depends on the difference between $a_T$ and $v_T$ which we denote by $D_T \equiv a_T - v_T$, further discussed in Section 2.4. We denote the optimal threshold as $\bar{x} = \mathcal{X}(N; D_T)$, so that $\bar{x} : [0, T] \to [0, U]$.

## 2.2 Aggregation

In this section, we aggregate the individual adoption decisions and compute the implied path for the fraction of adopters, $N$. We start by defining the probability that an agent alive at $s$ with state $x(s) = x$ survives until time $t$, while the value of her state remains below $\bar{x}$ during this period, i.e:

$$P(x, s, t; \bar{x}) = Pr\Big[x(\iota) \leq \bar{x}(\iota), \text{ for all } \iota \in [s, t] \ \Big| \ x(s) = x\Big] \, e^{-\nu(t-s)}. \tag{4}$$

For an agent that at time $s$ has $x \leq \bar{x}(s)$, the value of $P(x, s, t; \bar{x})$ gives the probability that this agent will survive up to $t$ without adopting.

We let $m_0(x)$ be the density of the agents at time $t = 0$ without the technology. Given our assumption about $x$, we require $0 \leq m(x) \leq 1/U$ for all $x \in [0, U]$. The fraction of agents that have adopted the technology at time $t$ is thus given by

$$N(t) = 1 - \int_0^U P(x, 0, t; \bar{x}) m_0(x) dx - \int_0^t \nu \left[ \int_0^U P(x, s, t; \bar{x}) \frac{1}{U} dx \right] ds. \tag{5}$$

The second term on the right hand side is the fraction of agents who did not have the technology at time 0 and survived until time $t$ without adopting. The third term considers the cohorts of agents that are born between 0 and $t$, and for each of these cohorts computes the fraction that survived without adopting up to $t$. We note that an equivalent version of equation (5) holds in a discretized version of the model. We denote the resulting path of $N$ as a function of $\bar{x}$ (the path of the adoption threshold) and of the initial condition $m_0$, namely $N = \mathcal{N}(\bar{x}; m_0)$.

## 2.3 Equilibrium

The equilibrium is given by the fixed point between the forward looking optimal adoption decision, encoded in $\mathcal{X}$, and the backward looking aggregation, encoded in $\mathcal{N}$. To emphasize the forward looking nature of $\mathcal{X}$, note that it depends on the terminal value function $D_T = a_T - v_T$. To emphasize the backward looking nature of $\mathcal{N}$, note that it propagates the initial condition $m_0$. We then have

Definition 1. Fix an initial condition $m_0$, and a terminal value function $D_T$. An equilibrium $\{N^*, \bar{x}^*\}$ solves the fixed point :

$$N^* = \mathcal{F}\left(N^*; m_0, D_T\right) \text{ where } \mathcal{F}\left(N; m_0, D_T\right) \equiv \mathcal{N}\left(\mathcal{X}\left(N; D_T\right); m_0\right) \tag{6}$$

and the corresponding $\bar{x}^* = \mathcal{X}(N^*; D_T)$.

Note that this is a canonical definition of equilibrium, where the operator $\mathcal{F}$ combines the two operators $\mathcal{N}$ and $\mathcal{X}$ defined before. This definition holds for both the continuous time and the discretized version of the model.

## 2.4 A Recursive Formulation of the Equilibrium

The functions $a(x,t)$ and $v(x,t)$, and the optimal policy $\bar{x}(t)$, have a recursive representation in terms of Hamilton-Jacobi-Bellman (HJB) partial differential equations. We derive these equations and their boundaries in Appendix F. The information encoded in the equations can be summarized by the value function $D(x,t) \equiv a(x,t) - v(x,t)$, which satisfies:

$$\rho D(x,t) = \min \left\{ \rho c \ , \ x(\theta_0 + \theta_n N(t)) + \frac{\sigma^2}{2} D_{xx}(x,t) + D_t(x,t) \right\} \tag{7}$$

for all $x \in [0,U]$, $t \in [0,T]$ and terminal condition $D(x,T) \equiv D_T(x) = a_T(x) - v_T(x)$.

We interpret the value function $D(x,t)$ as the opportunity cost of waiting to adopt. To see why, note that $a(x,t) - c$ is the net value of adopting immediately while $v(x,t)$ is the net optimal value, that may entail adopting in the future, see equation (2) and equation (3). From here it follows that

$$D(x,t) = \mathbb{E}\Big[ \int_t^\tau e^{-\rho(s-t)} \left(\theta_0 + \theta_n N(s)\right) x(s) ds + e^{-\rho(\tau - t)} c \ \Big| \ x(t) = x \Big]. \tag{8}$$

Optimality requires that $D(x,t) \leq c$, which implies the value matching condition at the barrier. We are looking for a classical solution that satisfies:

$$\rho D(x,t) = x(\theta_0 + \theta_n N(t)) + \frac{\sigma^2}{2} D_{xx}(x,t) + D_t(x,t) \tag{9}$$

for all $x \in [0, \bar{x}(t)]$ and $t \in [0,T]$ with boundary conditions:

$$\begin{aligned} D(\bar{x}(t),t) &= c && \text{Value Matching} \\ D_x(\bar{x}(t),t) &= 0 && \text{Smooth Pasting} \\ D_x(0,t) &= 0 && \text{Reflecting} \end{aligned} \tag{10}$$

If the solution is regular, it also features smooth pasting. Finally, since $x = 0$ is a reflecting barrier, the value function has a zero derivative at that point.

Let $m(x,t)$ denote the density of the agents with $x$ that have not adopted at $t$. The law of motion of $m$ for all $t \geq 0$ is:

$$
\begin{aligned}
m_t(x,t) &= \nu\left(\frac{1}{U} - m(x,t)\right) + \frac{\sigma^2}{2} m_{xx}(x,t) \text{ if } 0 \leq x \leq \bar{x}(t) \\
m(x,t) &= 0 \quad \text{ for } x \in [\bar{x}(t), U] \\
m_x(0,t) &= 0
\end{aligned}
\tag{11}
$$

and initial condition $m_0(x) = m(x,0)$ for all $x \in (0,U)$. The p.d.e. is the standard Kolmogorov forward equation (KFE). The density of non-adopters is zero to the right of $\bar{x}(t)$, since this is an exit point. The last boundary condition is obtained from our assumption that $x$ reflects at $x = 0$.

The fraction of agents that have adopted the technology is thus given by

$$
N(t) = 1 - \int_0^{\bar{x}(t)} m(x,t)dx. \tag{12}
$$

We use these equations to provide an equilibrium definition, equivalent to Definition 1, which emphasizes the dynamic nature of the equilibrium.

Definition 2. An equilibrium is given by the functions $\{D, m, \bar{x}, N\}$ satisfying the coupled p.d.e.'s for $D$ and $m$ given in equation (9) and equation (11), and the boundary conditions given in equation (10), equation (11) and equation (12).

We note that this system of p.d.e.'s is involved for two reasons. First, the equations are coupled through $\bar{x}$ and $N$. Second, the equations feature a time-varying free boundary, which is known to be non trivial.

# 3 Equilibrium of the Stochastic Baseline Model

In this section we establish equilibrium existence. We first give a normalization of the primal problem that is useful for empirical applications.

Lemma 1. The problem with parameters $\{c, \rho, \nu, \sigma, \theta_0, \theta_n, U\}$, initial condition $m_0$, $f(x) = 1/U$ and equilibrium objects $\{\bar{x}(t), N(t), a(x,t), v(x,t)\}$ for $x \in [0,U]$ and $t \in (0,T)$ is equivalent to the following normalized problem $\left\{\frac{c}{U\theta_0}, \rho, \nu, \frac{\sigma}{U}, 1, \frac{\theta_n}{\theta_0}, 1\right\}$ for a normalized variable $z \equiv \frac{x}{U} \in (0,1)$ and $t \in (0,T)$ with initial condition $m_0(z) = U\, m_0(x)$, $f(z) = 1$ and equilibrium objects $\left\{\frac{\bar{x}(t)}{U}, N(t), \hat{a}(z,t), \hat{v}(z,t)\right\}$ where $\hat{a}(z,t) \equiv \theta_0 a(zU,t)$ and $\hat{v}(z,t) \equiv \theta_0 v(zU,t)$.

The lemma shows that the problem features 5 independent parameters as $U$ and $\theta_0$ can be normalized without affecting the nature of the solution as the dynamics of the technology diffusion are unchanged.

## 3.1 Monotonicity and Existence of Equilibrium

The next proposition shows that the function $\mathcal{X}$, giving the path of the optimal threshold $\bar{x}$ as a function of the path $N$, is monotone decreasing. Thus an agent facing a higher path of adoption will choose to adopt earlier. Moreover, the proposition shows that an agent facing larger values of $\theta_0$ and/or $\theta_n$, will also adopt earlier.

Proposition 2. Fix the terminal value function $D_T = a_T - v_T$ and $\theta_n \geq 0$. Let $\bar{x}$ be the threshold path implied by $N(t)$. Consider two paths such that $N'(t) \geq N(t)$ for all $t \in [0, T]$, then $\bar{x}'(t) \leq \bar{x}(t)$. Moreover, let $\theta = (\theta_0, \theta_n)$ with the corresponding optimal threshold path $\bar{x}$. If $\theta' \geq \theta$ then $\bar{x}'(t) \leq \bar{x}(t)$.

Proposition 2 also holds if we replace the continuous time model by a discrete-time, discrete-state, approximation to it. For instance, it holds for a finite difference approximation, which we use for some computations, and which converges to the continuous-time version. The reason the proof holds is that we verify the conditions to use Topkis (1978). Thus, once we reformulate the problem in terms of stopping times, we can apply the monotone comparative statics logic developed by Milgrom and Shannon (1994) to characterize the policy function.

Next we show that for the same initial condition $m_0(x)$, if the path $\bar{x}(t) \leq \bar{x}'(t)$ then $N'(t) \leq N(t)$ for all $t$. We need to show that the fraction of non-adopters is decreasing in $\bar{x}(t)$. This implies that $\mathcal{N}$ is monotone decreasing.

Proposition 3. Fix $m_0$ and consider two path of thresholds $\bar{x}, \bar{x}'$ satisfying $\bar{x}'(t) \geq \bar{x}(t)$ for all $t \in [0, T]$. Let $N' = \mathcal{N}(\bar{x}'; m_0)$ and $N = \mathcal{N}(\bar{x}; m_0)$. Then $N'(t) \leq N(t)$ for all $t \in [0, T]$. Moreover, fix a threshold $\bar{x}$, and consider two initial measures with $m'_0(x) \geq m_0(x)$ for all $x \in [0, U]$, then $N' = \mathcal{N}(\bar{x}; m'_0)$ and $N = \mathcal{N}(\bar{x}; m_0)$. Then $N'(t) \leq N(t)$ for all $t \in [0, T]$.

The next theorem uses the monotonicity of $\mathcal{X}$ and $\mathcal{N}$, established in Proposition 2 and Proposition 3, which by the definition in equation (6) implies that $\mathcal{F}$ is monotone. This allows us to use Tarski's theorem. For technical reasons the theorem applies to a finite horizon, discretized version of the model introduced in Section 2.1 where the time domain $[0, T]$ is divided into $J$ segments and the state $[0, U]$ is divided into $I$ segments (see Definition 3 in Appendix A).[9] We have:

[9]The reason is the completeness of the lattice in which $\mathcal{F}$ is defined.

Theorem 1. Consider a finite horizon, discrete time - discrete state version of the model and $\theta_n \geq 0$. Fix an initial condition $m_0 \in \mathbb{R}_+^I$ and a terminal value function $D_T \in \mathbb{R}_+^I$.
(i) The equilibria of this model are a non-empty lattice. Hence the model has a smallest equilibrium, $\{\bar{x}^L, N^L\}$, and a largest one, $\{\bar{x}^H, N^H\}$, and any equilibrium path $\{\bar{x}, N\}$ satisfies $N^L \leq N \leq N^H$ and $\bar{x}^L \geq \bar{x} \geq \bar{x}^H$.
(ii) Let $\theta' \geq \theta$ and $m_0' \leq m_0$. Consider the equilibrium $\{\bar{x}', N'\}$ with the largest $N'$ corresponding to $\{\theta', m_0'\}$ and the equilibrium $\{\bar{x}, N\}$ with largest $N$ corresponding to $\{\theta, m_0\}$. Then $\bar{x}' \leq \bar{x}$ and $N' \geq N$.

The first statement of the theorem establishes existence of the equilibrium for the finite horizon - discrete time version of the model. The result holds for an arbitrary small length of the time period, and for an arbitrary large horizon $T$. An important consequence of the theorem is that the equilibrium set, for a given initial distribution of non-adopters $m_0$ and terminal valuation $D_T = a_T - v_T$, is a lattice. Moreover, we can compute the value of the extreme equilibria by iterating on $N^{k+1} = \mathcal{F}(N^k; D_T, m_0)$ for $k = 0, 1, \ldots,$ starting from $N^0(t) = 1$ or from $N^0(t) = 0$, for all $t$. The theorem ensures that the limit converges to a fixed point. If the two sequences converge to the same limit, then the equilibrium is unique. The second statement of the theorem establishes a useful comparative statics result: considering a model with a larger $\theta$ or with a smaller $m_0$ implies that the high-adoption equilibrium is larger (more agents adopt).

## 3.2 No Adoption Equilibrium

We briefly analyze the equilibrium in which there is no adoption i.e., $\bar{x}(t) = U$ for all $t$. For simplicity we focus on the case where $T = \infty$. This case is particularly easy because agents decision are in a corner. We find the basin of attraction for such equilibrium, i.e., we find a threshold for the number of adopters $\underline{N}$, so that a no adoption equilibrium exists if and only if at $t = 0$ there are fewer agents with the technology than $\underline{N}$.

Proposition 4. A no-adoption equilibrium with $\bar{x}(t) = U$ and $N(t) = N(0)e^{-\nu t}$ for all $t \geq 0$ exists if and only if $1 - \int_0^U m_0(x)dx \leq \underline{N}$, where

$$\frac{\rho c}{U} = \theta_0 \left[1 + g(\eta U)\right] + \underline{N} \frac{\rho \theta_n}{\rho + \nu} \left[1 + g(\eta' U)\right] \tag{13}$$

$$\eta = \sqrt{\frac{2\rho}{\sigma^2}}\,,\ \eta' = \sqrt{\frac{2(\rho+\nu)}{\sigma^2}} \text{ and } g(y) \equiv \frac{\operatorname{csch}(y) - \coth(y)}{y} \in (-\tfrac{1}{2}, 0)\ . \tag{14}$$

Note that $\underline{N} > 0$ if and only if $\frac{\rho c}{U} > \theta_0 \left[1 + g(\eta U)\right]$. Moreover, if $\underline{N} > 0$ we have:

(i) $\underline{N}$ is an increasing function of $\sigma$, satisfying

$$\frac{\rho + \nu}{\rho \theta_n} \left( \frac{\rho c}{U} - \theta_0 \right) \leq \underline{N} \leq \frac{\rho + \nu}{\rho \theta_n} \left( 2 \frac{\rho c}{U} - \theta_0 \right), \tag{15}$$

where the two limits are reached as $\sigma \to 0$ and as $\sigma \to \infty$, respectively.
(ii) $\underline{N}$ is a decreasing function of $\theta_n$.

An immediate corollary of this proposition is that $m_0(x) = 1/U$ is an invariant distribution provided that $\underline{N} \geq 0$, i.e., under this condition if we start with no adoption, then we stay with no adoption. The fact that $\underline{N} > 0$ requires $\theta_0$ to be small is intuitive: when this condition is violated then agents with a large $x$ will find it profitable to adopt regardless. Likewise, the effect of $\sigma$ is intuitive since, for a given $U$, a large $\sigma$ makes the process to revert to the mean faster. Finally, if $\theta_n$ is large then it is more profitable to coordinate on high $N$ and then the basin of attraction is smaller.

# 4 Stationary Equilibria

In this section we let $T = \infty$ and analyze the stationary version of the model and its corresponding stationary distributions. We look for an initial condition $m_0$ such that the distribution is invariant, and that both $\bar{x}(t) = \bar{x}_{ss}$ and $N(t) = N_{ss}$ are constant through time.

## 4.1 Stationary Equilibria in the Deterministic Model ($\sigma = 0$)

We begin by studying the deterministic case where $\sigma = 0$, so that the agent's valuation $x$ does not change. This case is useful to relate to the existing literature studying technology diffusion (e.g., Stokey (2020); Buera et al. (2021); Crouzet et al. (2023)), and it unveils the basic forces at work in the adoption problem.

We specialize equation (7) to the stationary equilibrium of the deterministic model. Since $\sigma^2 = 0$ then $D_{xx}\sigma^2 = 0$ and, since we focus on the stationary case, $D_t = 0$. The equation becomes

$$\rho \tilde{D}(x) = \min \left\{ \rho c \ , \ x(\theta_0 + \theta_n N_{ss}) \right\} \tag{16}$$

for all $x \in [0, U]$. The stationary threshold $\bar{x}_{ss}$ is the value of $x$ solving

$$\rho c = \bar{x}_{ss}(\theta_0 + \theta_n N_{ss}). \tag{17}$$

Using equation (11), imposing the $\sigma^2 = 0$ and the stationary condition $m_t = 0$ gives one equation for the invariant distribution of agents without the technology which is given by $\tilde{m}(x) = \frac{1}{U}$ for $x \in [0, \bar{x}_{ss}]$ and $\tilde{m}(x) = 0$ for $x \in [\bar{x}_{ss}, U]$ so that we have

$$N_{ss} = 1 - \frac{\bar{x}_{ss}}{U}. \tag{18}$$

Solving this simple system for $\bar{x}_{ss}$ gives a quadratic equation that can have zero, one or two interior stationary distributions. We have the following result:

Proposition 5. There are two cases. Case (i): If $\rho c < \theta_0 U$, then there is a unique stationary distributions with $\bar{x}_{ss}$ with $0 < \bar{x}_{ss} < U$. Case (ii): If $\theta_0 U < \rho c$, then there is always a no activity stationary equilibrium with $\bar{x}_{ss} = U$. In this case, there is threshold value for $\theta_n^*$ such that, if $\theta_n < \theta_n^*$, there is no other stationary equilibrium, whereas if $\theta_n > \theta_n^*$ there are two additional interior stationary equilibrium.

A key insight of the proposition is that multiple interior stationary equilibria occur when the complementarities are strong relative to the intrinsic value of the technology, i.e., when $\theta_0$ is small and $\theta_n$ is large. We concentrate on the stationary equilibria of the deterministic model for two reasons. First, for small $\sigma$ they provide a good benchmark for the stationary equilibria of the stochastic model analyzed next. Second, we omit the treatment of the dynamics of this model because for a non-pathological set of initial conditions the model converges immediately to the corresponding stationary distribution.[10]

## 4.2 Stationary Equilibria in the Stochastic Model ($\sigma > 0$)

Next we analyze the stationary equilibria of the stochastic version of the model. We will show that convergence to these equilibrium must be gradual, i.e., unlike the deterministic case, it is not possible to "jump" to the equilibrium given a generic initial condition.

A stationary equilibrium is given by two constant values of $N_{ss}$ and $\bar{x}_{ss}$ that solve the time invariant version of the partial differential equations presented in Section 2.4. Given $N_{ss}$ we can derive the value function $D(x,t) = \tilde{D}(x)$ and $\bar{x}(t) = \bar{x}_{ss}$. Given $\bar{x}_{ss}$ we can derive the invariant distribution $m(x,t) = \tilde{m}(x)$, from which we derive $N_{ss}$. Given $N_{ss}$, we find $\tilde{D}$

[10]Indeed, in Appendix J we show that if the initial condition is such that at time zero no agent with low valuation has adopted the technology (while some high valuation agents may have done so), the equilibrium of the deterministic problem has no dynamics. This implies that adoption occurs instantaneously and that the fraction of adopters is a constant $N(t) = N_{ss}$.

and $\bar{x}_{ss}$ that solve:

$$\begin{aligned}
\rho\tilde{D}(x) &= x(\theta_0+\theta_n N_{ss}) + \frac{\sigma^2}{2}\tilde{D}_{xx}(x) \text{ if } x\in[0,\bar{x}_{ss}] && \text{Value of Adoption}\\
\tilde{D}_x(0) &= 0 && \text{Reflecting}\\
\tilde{D}(\bar{x}_{ss}) &= c && \text{Value Matching}\\
\tilde{D}_x(\bar{x}_{ss}) &= 0 && \text{Smooth Pasting}
\end{aligned}$$

Given $\bar{x}_{ss}$ solve for $\tilde{m}$

$$\begin{aligned}
0 &= -\nu\tilde{m}(x) + \nu\frac{1}{U} + \frac{\sigma^2}{2}\tilde{m}_{xx}(x) && \text{KFE if } x\le\bar{x}_{ss}\\
\tilde{m}(\bar{x}_{ss}) &= 0 \text{ and } \tilde{m}_x(0)=0 && \text{Exit and Reflecting}
\end{aligned}$$

and given $\tilde{m}(x)$ and $\bar{x}_{ss}$, we define the fixed point

$$N_{ss} = 1 - \int_0^{\bar{x}_{ss}} \tilde{m}(s)dx.$$

We begin by solving $\tilde{D}(x)$, and $\bar{x}_{ss}$ given a value for $N_{ss}$. The details of the solution can be found in Appendix I.1. Using the solutions for $\tilde{D}$ we can solve for $\mathcal{X}_{ss}:[0,1]\to[0,U]$, a function that gives the *optimal* stationary threshold as a function of a given $N_{ss}$. The monotonicity properties of the function $\tilde{D}$ on the parameters $N_{ss},\theta_n,c$ and $\theta_0$ give the following characterization of the threshold $\mathcal{X}_{ss}$.

Lemma 2. The function $\mathcal{X}_{ss}$ is decreasing in $N_{ss}$, strictly so at the points where $0<\bar{x}_{ss}<U$. Fixing a value of $N_{ss}$, the function $\mathcal{X}_{ss}$ is strictly increasing in $c$, strictly so at the points where $0<\bar{x}_{ss}<U$. Fixing a value of $N_{ss}$, the function $\mathcal{X}_{ss}$ is strictly decreasing in $\theta_0$ and $\theta_n$ at the points where $0<\bar{x}_{ss}<U$. Moreover we have the following expansion: $\mathcal{X}_{ss}(N_{ss}) = \frac{\rho c}{\theta_0+\theta_n N_{ss}} + \frac{\sigma}{\sqrt{2\rho}} + o(\sigma)$.

Since the function $\mathcal{X}_{ss}(N_{ss})$ is decreasing in $N_{ss}$, it has an inverse, which we denote by $\mathcal{X}_{ss}^{-1}$, and is given by:

$$\mathcal{X}_{ss}^{-1}(\bar{x}_{ss}) = \frac{1}{\theta_n}\left[\frac{\rho c}{\left(\bar{x}_{ss}+\bar{A}_1 e^{\eta\bar{x}_{ss}}+\bar{A}_2 e^{-\eta\bar{x}_{ss}}\right) - \frac{\left(1+\eta\left(\bar{A}_1 e^{\eta\bar{x}_{ss}}-\bar{A}_2 e^{-\eta\bar{x}_{ss}}\right)\right)\left(e^{\eta\bar{x}_{ss}}+e^{-\eta\bar{x}_{ss}}\right)}{\eta\left(e^{\eta\bar{x}}-e^{-\eta\bar{x}_{ss}}\right)}} - \theta_0\right] \text{ where}$$

$$\bar{A}_1 \equiv \frac{1}{\eta}\frac{\left(1-e^{-\eta U}\right)}{\left(e^{-\eta U}-e^{\eta U}\right)},\ \bar{A}_2 \equiv \frac{1}{\eta}\frac{\left(1-e^{\eta U}\right)}{\left(e^{-\eta U}-e^{\eta U}\right)} \text{ and } \eta\equiv\sqrt{2\rho/\sigma^2}. \tag{19}$$

Note that, from the expansion given in Lemma 2, fixing $\bar{x}_{ss}$, then $\mathcal{X}_{ss}^{-1}(\bar{x}_{ss})$ is increasing in

$\sigma$ in a neighborhood of $\sigma = 0$, provided that $\theta_n > 0$, we have

$$\mathcal{X}_{ss}^{-1}(\bar{x}_{ss}) \approx \frac{1}{\theta_n}\left(\frac{c\rho}{\bar{x}_{ss} - \sigma/\sqrt{2\rho}} - \theta_0\right).$$

Next we can solve the Kolmogorov forward equations for $\tilde{m}(x)$, given a barrier $\bar{x}_{ss}$ subject to an exit point and to the conditions coming from the reflecting barriers. We denote the corresponding value of the fraction that have adopted as $\mathcal{N}_{ss}(\bar{x}_{ss})$. The details of the solutions can be found in Appendix I.2. Solving this equation we obtain

$$\mathcal{N}_{ss}(\bar{x}_{ss}) = 1 - \frac{\bar{x}_{ss}}{U} + \frac{\tanh(\gamma\bar{x}_{ss})}{U\gamma} \text{ where } \gamma \equiv \sqrt{2\nu/\sigma^2}. \tag{20}$$

As it is intuitive, the value of $\mathcal{N}_{ss}(\bar{x}_{ss})$ is *decreasing* in the level of the barrier $\bar{x}$. The next lemma, obtained by analyzing equation (20) gives a characterization of $\mathcal{N}_{ss}$.

Lemma 3. Fix $\gamma > 0$, then $\mathcal{N}_{ss}(\bar{x})$ is strictly decreasing in $\bar{x}_{ss}$. Fixing $\bar{x} > 0$, then $\mathcal{N}_{ss}$ is strictly increasing in $\gamma$, and hence strictly decreasing in $\sigma$. Moreover, we have the expansion: $\mathcal{N}_{ss}(\bar{x}) = 1 - \frac{\bar{x}_{ss}}{U} + \frac{\sigma}{U\sqrt{2\nu}} + o(\sigma)$.

The system given by equation (19) and equation (20) determines $\bar{x}_{ss}$ and $N_{ss}$. In particular, a stationary equilibrum is described by the pair $\{\bar{x}_{ss}, N_{ss}\}$, which solves

$$N_{ss} \equiv \mathcal{N}_{ss}(\bar{x}_{ss}) = \mathcal{X}_{ss}^{-1}(\bar{x}_{ss}).$$

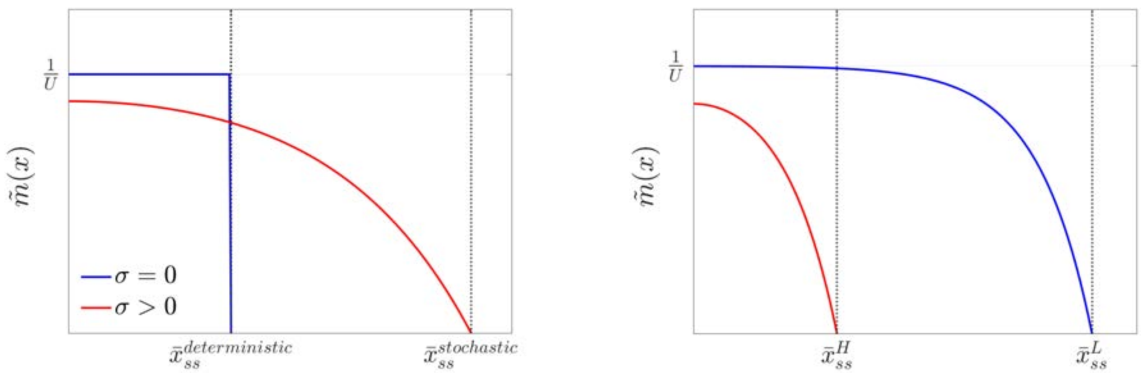
Next, we summarize the behavior of the stationary equilibrium for small values of $\sigma$. We label the stationary equilibrium with superscripts $\{H, L\}$ to hint at the associated High or Low level of adoption, so that $\bar{x}^H < \bar{x}^L$.

Proposition 6. Assume that $\nu > 0$ and that the parameters $\theta_0, \theta_n, c$ and $\rho$ are such that there are two interior stationary equilibria in the deterministic case of $\sigma = 0$, and label them as $\bar{x}_{ss}^H < \bar{x}_{ss}^L$. Then, (i) there exists a $\bar{\sigma} > 0$ such that for all $\sigma \in (0, \bar{\sigma})$ there are two interior stationary equilibria with $\bar{x}_{ss}^H < \bar{x}_{ss}^L$. (ii) The thresholds for each stationary equilibria is continuous with respect to $\sigma$ at $\sigma = 0$. (iii) The sign of the comparative static differs across stationary equilibria, with

$$\frac{\partial \bar{x}_{ss}^H}{\partial c} > 0 > \frac{\partial \bar{x}_{ss}^L}{\partial c} \quad \text{and} \quad \frac{\partial \bar{x}_{ss}^L}{\partial \theta_0} > 0 > \frac{\partial \bar{x}_{ss}^H}{\partial \theta_0}$$

The proposition shows that the high adoption stationary equilibria behaves in an intuitive way, with more adoption (a lower $\bar{x}_{ss}^H$) associated to a smaller adoption cost ($c$), or to a larger intrinsic value of the technology ($\theta_0$). The comparative statics for the low adoption stationary state are just the opposite.

Figure 1: Stochastic Stationary Equilibria: Density of non-adopters: $m(x)$




(a) Deterministic and Stochastic Stat. Eqbm (b) High and Low Adoption Stat. Eqbm

Panel (a) of Figure 1 compares the density of stationary distribution of non-adopters for the deterministic case ($\sigma = 0$) with the one for the stochastic case ($\sigma > 0$). The key difference is that in the invariant equilibrium of the stochastic case there are agents with low benefits, namely with $x(t) < \bar{x}_{ss}$, who have the technology. These are agents who adopted the technology in the past (for some $t' < t$ when $x(t') > \bar{x}(t')$, and whose stochastic $x$ decreased over time. As a result, $m(x) < 1/U$ when $\sigma > 0$, and the density of non-adopters below $\bar{x}_{ss}$ is not uniform. Given that the density takes time to adjust, the stochastic model features the presence of dynamics in the adoption of a new technology; while the distribution of the deterministic model can be generated instantaneously (with the agents above $\bar{x}$ immediately adopting), the distribution of the stochastic model takes time to change from the initial uniform distribution to the invariant distribution, as agents adopt when $x(t) > \bar{x}(t)$ and it takes times for these $x's$ to crawl back below the stationary threshold. Panel (b) shows the densities of the invariant distribution of the high and the low adoption stationary equilibria in the stochastic model ($\sigma > 0$). Naturally, $\bar{x}_{ss}^H < \bar{x}_{ss}^L$, so that both equilibria have adopters below the stationary thresholds, although fewer of them in the low adoption stationary equilibrium.

# 5 Stability of Stationary Equilibria

In this section we analyze the local stability of the stationary equilibria. We explore the question using a perturbation of the distribution of adopters in each of the two interior invariant distributions using techniques from the Mean Field Game literature developed in Alvarez, Lippi and Souganidis (2022b). For the purpose of this section we model the equilibrium as from Definition 2. This dynamical system is infinite-dimensional because the state at every time $t$ is given by the entire density $m(x,t)$.

We consider the stationary equilibrium given by $\tilde{m}$, and we ask whether starting from an initial condition $m_0$ close to $\tilde{m}$ the economy converges back to $\tilde{m}$. Because the system is infinite-dimensional there are many deviations that are possible. Any initial condition can be described by $m_0(x) = \tilde{m}(x) + \epsilon\omega(x)$, for some $\omega$ satisfying $\int_0^U \omega(x)dx = 0$. The sense in which the analysis is local is that we evaluate the derivative of the system with respect to $\epsilon$ and evaluate it at $\epsilon = 0$. The alert reader will notice that the local dynamics of a system in $\mathbb{R}^q$ are encoded in a $q \times q$ matrix. The analogous infinite dimensional object is a linear operator that will be analytically presented below.

We begin with the approximation of $\bar{x}(t) = \mathcal{X}(N)(t)$. We take the directional derivative (Gateaux) with respect to an arbitrary perturbation $n$ of a constant path $N$. In particular, we consider paths defined by $N(t) = N_{ss} + \epsilon\, n(t)$ around the stationary value $N_{ss}$. We will denote this Gateaux derivative by $\bar{y}$.

**Proposition 7.** Fix a stationary equilibrium with interior $\bar{x}_{ss}$, and its corresponding $N_{ss}$. Let $D_T$ be equal to the stationary value function $\tilde{D}$ corresponding to that stationary equilibrium. Let $n : [0,T] \to \mathbb{R}$ be an arbitrary perturbation. Then

$$\bar{y}(t) \equiv \lim_{\epsilon \downarrow 0} \frac{\mathcal{X}(N_{ss} + \epsilon n; \tilde{D})(t) - \mathcal{X}(N_{ss}; \tilde{D})(t)}{\epsilon}$$
$$= \frac{\theta_n}{\tilde{D}_{xx}(\bar{x}_{ss})} \int_t^T G(\tau - t) n(\tau) d\tau, \tag{21}$$

where

$$G(s) \equiv \sum_{j=0}^{\infty} c_j e^{-\psi_j s} \geq 0\,,\ \psi_j \equiv \rho + \frac{\sigma^2}{2}\left(\frac{\pi(\frac{1}{2}+j)}{\bar{x}_{ss}}\right)^2 \text{ and } c_j \equiv 2\left(1 - \frac{\cos(\pi j)}{\pi(j+\frac{1}{2})}\right),$$

where $\tilde{D}_{xx}(\bar{x}_{ss}) < 0$ is the second derivative of the stationary value function:

$$\tilde{D}_{xx}(\bar{x}_{ss}) = \frac{\rho c - \bar{x}_{ss}\left[\theta_0 + \theta_n N_{ss}\right]}{\sigma^2/2}\,,\ N_{ss} = 1 - \frac{\bar{x}_{ss}}{U} + \frac{\tanh(\gamma \bar{x}_{ss})}{\gamma U} \text{ and } \gamma = \sqrt{\frac{2\nu}{\sigma^2}}.$$

Thus, we can write $\bar{x}(t) = \bar{x}_{ss} + \epsilon\bar{y}(t) + o(\epsilon)$. Note that $G$ is positive and $D_{xx}$ is negative, so the effect of the future path on the current value is negative, which is consistent with the property that $\mathcal{X}$ is decreasing. Also note that it is proportional to $\theta_n$, so if $\theta_n = 0$, then the threshold will be constant. Thus, the approximation of $\bar{x}(t)$ depends on the perturbation of the path of $N$ from $t$ to $T$, given by $n(s)$ for $s = [t, T]$. The proof of the proposition is obtained by jointly differentiating with respect to $\epsilon$ the system defined by $D$ and $\bar{x}$ in equation (9) and equation (10). This yields a new p.d.e., and new boundary conditions. The expression for $\bar{y}$ is obtained once we solve this new p.d.e., see the proof in Appendix C.1.

Now we turn to the perturbation for the fraction of the adopters as a function of the thresholds and of a perturbation of the initial condition. We approximate $N(t) = \mathcal{N}(\bar{x}, m_0)(t)$ by taking the directional derivative (Gateaux) with respect to an arbitrary perturbation $\bar{y}$ of a constant path $\bar{x}$ and a perturbation $\omega$ on the stationary density $\tilde{m}$. In particular, we consider paths defined by $\bar{x}(t) = \bar{x}_{ss} + \epsilon\, \bar{y}(t)$ around the stationary threshold $x_{ss}$, and $m_0(x) = \tilde{m}(x) + \epsilon\omega(x)$. We will denote this Gateaux derivative by $n$.

Proposition 8. Fix the interior threshold $\bar{x}_{ss}$ of a stationary equilibrium and its corresponding $N_{ss}$, and let $\tilde{m}$ be the corresponding invariant distribution of non-adopters. Let $\omega : [0, \bar{x}_{ss}] \to \mathbb{R}$ be an arbitrary perturbation to the distribution, and let $\bar{y} : [0, T] \to \mathbb{R}$ be an arbitrary perturbation of the threshold. Then

$$\begin{aligned} n(t) &\equiv \lim_{\epsilon \downarrow 0} \frac{\mathcal{N}(\bar{x}_{ss} + \epsilon \bar{y}; \tilde{m} + \epsilon w)(t) - \mathcal{N}(\bar{x}_{ss}; \tilde{m})(t)}{\epsilon} \\ &= n_0(\omega)(t) + \frac{\tilde{m}_x(\bar{x}_{ss})\sigma^2}{\bar{x}_{ss}} \int_0^t J(t-\tau)\bar{y}(\tau)d\tau \end{aligned} \tag{22}$$

where

$$J(s) = \sum_{j=0}^{\infty} e^{-\mu_j s} \text{ with } \mu_j = \nu + \frac{1}{2}\sigma^2 \left( \frac{\pi(\frac{1}{2}+j)}{\bar{x}_{ss}} \right)^2 \tag{23}$$

$$n_0(\omega)(t) \equiv -\sum_{j=0}^{\infty} \frac{\bar{x}_{ss}}{\pi(\frac{1}{2}+j)} \frac{\langle \varphi_j, \omega \rangle}{\langle \varphi_j, \varphi_j \rangle} e^{-\mu_j t}, \tag{24}$$

$$\varphi_j(x) \equiv \sin\left( \left( \frac{1}{2} + j \right) \pi \left( 1 - \frac{x}{\bar{x}_{ss}} \right) \right) \text{ for } x \in [0, \bar{x}_{ss}] \tag{25}$$

$$\frac{\langle \varphi_j, \omega \rangle}{\langle \varphi_j, \varphi_j \rangle} = \frac{2}{\bar{x}_{ss}} \int_0^{\bar{x}_{ss}} \varphi_j(x)\omega(x)dx \text{ and } \tilde{m}_x(\bar{x}_{ss}) = -\frac{\gamma}{U}\tanh(\gamma \bar{x}_{ss}).$$

Thus, we can write $N(t) = N_{ss} + \epsilon n(t) + o(\epsilon)$. This formula encodes the effect of two perturbations: $\omega$ and $\bar{y}$. The former is the perturbation on the initial condition $m_0$, whose effect is in the term $n_0(\omega)(t)$. We note that $n_0(\omega)(t)$ is the effect at time $t$ on the path $N(t)$ triggered by a perturbation of the initial condition keeping the threshold rule $\bar{x}$ fixed. The function $n_0(\omega)$ can be further reinterpreted by considering the limiting case of a perturbation $\omega$ given by a distribution concentrated at $x = \hat{x} \leq \bar{x}_{ss}$, i.e., a Dirac's delta function as

$\omega(x) = \delta_{\hat{x}}(x)$. In this case,

$$n_0(\delta_{\hat{x}})(t) = -\sum_{j=0}^{\infty} 2\frac{\sin\left(\left(\frac{1}{2}+j\right)\pi\left(1-\frac{\hat{x}}{\bar{x}_{ss}}\right)\right)}{(\frac{1}{2}+j)\pi} e^{-\mu_j t}.$$

The effect of the perturbation, $\bar{y}$, on the path of the threshold, $\bar{x}(s)$, is captured by the second term in equation (22). This term gives the effect at time $t$ on the path $N(t)$ of a perturbation of the threshold rule $\bar{x}$, keeping the initial condition $\tilde{m}$ fixed. Also, consistent with our general result for $\mathcal{N}$, the effect of the threshold is negative, as $J > 0$ and $\tilde{m}_x(\bar{x}_{ss}) < 0$.

For future reference it is useful to understand the behavior of $n_0(t)$ as function of time. In particular, the rate at which the perturbation $\omega$ to the initial distribution converges back to the stationary distribution, while keeping $\bar{x}(t) = \bar{x}_{ss}$. This rate is given by the value of $\mu_0 = \nu + \frac{\sigma^2}{8}\left(\frac{\pi}{\bar{x}_{ss}}\right)^2$, i.e., the dominant eigenvalue. The proof's strategy is in Appendix C.2 and resembles the one for the previous proposition.

The next step is to use the last two propositions to derive one equation for the linearized equilibrium as a function of the perturbed initial distribution $m_0(x) = \tilde{m}(x) + \epsilon\omega(x)$. We combine equation (21) and equation (22) to arrive to a single linear equation that $n(t)$ must solve as a function of $\omega$.

Theorem 2. Fix an interior threshold $\bar{x}_{ss}$ for an stationary state, with its corresponding $N_{ss}$, and let $\tilde{m}$ be the corresponding invariant distribution of non-adopters. Let $m_0(x) = \tilde{m}(x) + \epsilon\omega(x)$. Let $D_T$ be equal to the value function $\tilde{D}$ corresponding to that stationary equilibrium. The linearized equilibrium must solve

$$n(t) = n_0(\omega)(t) + \Theta\int_0^T K(t,s)n(s)ds, \tag{26}$$

where $n_0(\omega)(t)$ is given in Proposition 8 and $\Theta \equiv \frac{\tilde{m}_x(\bar{x}_{ss})\sigma^2\theta_n}{\bar{x}_{ss}\tilde{D}_{xx}(\bar{x}_{ss})} > 0$. The kernel $K$ is given by

$$K(t,s) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} c_j e^{-\mu_i t - \psi_j s}\left[\frac{e^{(\mu_i+\psi_j)\min\{t,s\}} - 1}{\mu_i + \psi_j}\right] > 0. \tag{27}$$

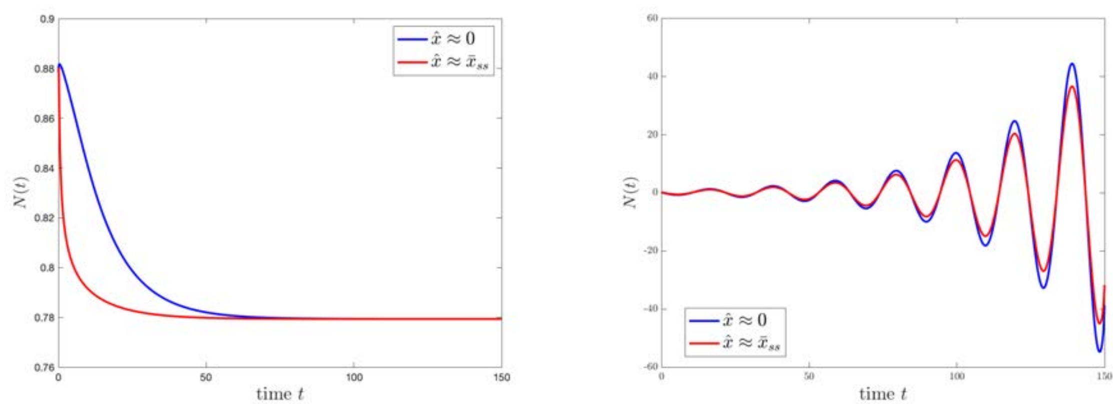
Moreover, $\mathrm{Lip}_K \equiv \sup_t \int |K(t,s)|ds \leq \left(\frac{\bar{x}_{ss}^2}{\sigma^2}\right)^2$. Furthermore, if $\Theta\,\mathrm{Lip}_K < 1$ there exists a unique bounded solution to equation (26) which is the limit of

$$n = \left[I + \Theta\mathcal{K} + \Theta^2\mathcal{K}^2 + \dots\right] n_0(\omega) \;\; \text{where} \;\; \mathcal{K}(g)(t) \equiv \int_0^T K(t,s)g(s)ds, \tag{28}$$

and where $\mathcal{K}^{j+1}(g)(t) \equiv \int_0^T K(t,s)\,\mathcal{K}^j(g)(s)\,ds$ for any bounded $g:[0,T] \to \mathbb{R}$.

A few remarks are in order. First, note that $K$ depends on $\theta_n$ because $\mu_j, \psi_j$ are a function of $\bar{x}_{ss}$, which is itself a function of $\theta_n$. The coefficient $\Theta$ depends on $\theta_n$ directly and indirectly through $\bar{x}_{ss}$. Hence the equation (26) and its solution depend on which steady equilibria we focus on. Second, if we discretize time so that $t \in \{\Delta_t(j-1) : j = 1, \ldots, J\}$ for $\Delta_t = \frac{T}{J-1}$, as done in Section 2.1, then the operator $\mathcal{K}$ is a $J \times J$ matrix with elements $K(t_i, t_j)$, and $n_0, n$ are $J \times 1$ vectors, so that equation (26) becomes the linear equation $n = n_0 + \Theta \mathcal{K}\, n$. Third, the fact that $\Theta\mathcal{K} > 0$ implies that the terms $\Theta\mathcal{K} + \Theta^2\mathcal{K}^2 + \ldots$ in equation (28) give the amplification over and above $n_0$, due to the time-varying path of the barrier $\bar{x}$.

Figure 2: Perturbation of Stationary Equlibria




(a) High Adoption Stat. Eqbm.

(b) Low Adoption Stat. Eqbm.

Figure 2 illustrates the stability of the high and low adoption equilibria, respectively, in Panels (a) and (b). Each panel considers two shocks that displace a small mass of agents away from the invariant distribution of non-adopters and endows them with the app. The shocks differ in the direction in which the mass is displaced. The blue line depicts the case where the app is given to agents with low benefit, namely with $x \approx 0$, while the red line considers a perturbation where the app is assigned to agents with a high benefit, namely with $x \approx \bar{x}_{ss}$. Two remarks are due. First, the high adoption equilibrium is locally stable, as displayed in Panel (a): for all shocks considered, the system returns to its invariant distribution. We also note that the half life of the shock is much shorter when the perturbation assigns the app to agents with a high benefit ($x \approx \bar{x}_{ss}$), as these agents were going to get the app soon anyways. Second, it is apparent that the low adoption equilibrium is unstable in Panel (b): the dynamics of the system following a perturbation are explosive, showing that the sequence in equation (28) does not converge; i.e., that the system does not return to the the invariant distribution after being shocked. To see the explosive nature of the path nearby the low activity stationary equilibrium, note the difference in the scale of the two panels.

# 6 The Planning Problem

This section sets up the planning problem in the stochastic version of the model ($\sigma > 0$). We first state the planning problem, provide a characterization of its solution, and show how it can be decentralized as an equilibrium with a subsidy. Section D.1 characterizes the stationary solution of this problem.[11]

The planner solves a non-trivial dynamic problem since the state of the economy is an entire distribution. At time zero the planner solves:

$$\max_{\{\bar{x}(t)\}} \left\{ \int_0^\infty e^{-rt} \int_0^U \underbrace{(1/U - m(x,t))}_{\text{Density of adopters}} \underbrace{x\left(\theta_0 + \theta_n N(t)\right)}_{\text{Flow benefit}} dx\, dt \right.$$
$$\left. - \underbrace{\int_0^\infty e^{-rt} c\left(N_t(t) + \nu N(t)\right) dt}_{\text{Flow of adoption cost: gross new adoptions}} \right\}$$

subject to

$$N(t) = 1 - \int_0^{\bar{x}(t)} m(z,t)dz \quad \text{for all } t$$
$$m_t(x,t) = -\nu\left(m(x,t) - 1/U\right) + \frac{\sigma^2}{2} m_{xx}(x,t) \quad \text{for } x \in (0, \bar{x}(t)) \text{ and all } t \geq 0 \quad \text{KFE}$$
$$m(x,t) = 0 \quad \text{for } x \in [\bar{x}(t), U] \text{ and all } t \geq 0 \quad \text{Adoption}$$
$$m_x(0,t) = 0 \quad \text{for all } t \geq 0 \quad \text{Reflecting}$$
$$m(x,0) = m_0(x) \text{ for all } x \quad \text{initial condition}$$

The objective function of the planner integrates the lifetime utility of agents using as a weight the discount factor $e^{-rt}$ for the cohort born at $t$. The first term contains the utility flows of all those using the technology. The second term subtracts the cost of adoption across time, where $N_t(t) + \nu N(t)$ is the gross cost of adoption at time $t$. The planner decides at each time a threshold $\bar{x}(t)$ which determines adoption, and takes as given the initial condition $m_0(x)$. The planner takes as given the law of motion of the density $m$ that is only affected through the choice of $\bar{x}$. The first constraint defines $N(t)$, the second one is the KFE of the density of non-adopters. As before, the density of non-adopters is zero to the right of $\bar{x}(t)$,

[11] Appendix D.5 uses a linearized version of the problem to analyze dynamics around its invariant distribution.

there is an exit point at $\bar{x}(t)$, and there is a boundary conditions from reflection at zero.

To characterize the solution we write the lagrangian for this problem. We denote the lagrange multiplier of the KFE equation by $e^{-rt}\lambda(x,t)$ and replace $N(t)$ and $N_t(t)$ by the corresponding integrals. To derive the p.d.e's for non-adopters, we first adapt the planning problem to a discrete-time discrete-state using a finite-difference approximation. In this set up we allow for a more general policy, i.e., not necessarily a threshold rule. We obtain the first order conditions for a problem in finite dimensions and take limits to find the corresponding p.d.e's. We provide details of this derivation in Appendix D.3. The p.d.e's corresponding to the planning problem are summarized in the following proposition.

Proposition 9. A planner problem is given by $\{\bar{x}(t), \lambda(x,t), m(x,t)\}$ such that adoption occurs for $x \geq \bar{x}(t)$, and the Lagrange multiplier $\lambda$, and the density of non-adopters $m$ solve the p.d.e. for non-adopters:

$$\begin{aligned}\rho\lambda(x,t) &= x\big(\theta_0 + \theta_n[1 - \int_0^{\bar{x}(t)} m(z,t)dz]\big) + \theta_n\big(\tfrac{U}{2} - \int_0^{\bar{x}(t)} m(z,t)z\,dz\big) \\ &\quad + \tfrac{\sigma^2}{2}\lambda_{xx}(x,t) + \lambda_t(x,t) \text{ for } x \leq \bar{x}(t) \text{ and } t \geq 0 \end{aligned} \tag{29}$$

$$\begin{aligned}\lambda(x,t) &= c \text{ for } x \geq \bar{x}(t) \text{ and } t \geq 0 \\ \lambda_x(\bar{x}(t),t) &= 0 \text{ for } t \geq 0 \\ \lambda_x(0,t) &= 0 \text{ for } t \geq 0\end{aligned} \tag{30}$$

and

$$\begin{aligned}m_t(x,t) &= \nu\big(1/U - m(x,t)\big) + \tfrac{\sigma^2}{2}m_{xx}(x,t) \text{ for } x < \bar{x}(t) \text{ and } t \geq 0 \\ m(x,t) &= 0 \text{ for } x \geq \bar{x}(t) \text{ and } t \geq 0 \\ m_x(0,t) &= 0 \text{ for } t \geq 0 \\ m(x,0) &= m_0(x) \text{ for all } x\end{aligned}$$

This proposition has two important consequences. First, it allows us to compute the solution of the planning problem following similar steps as for the computation of the equilibrium described in Section 3.1. Second, it indicates how to decentralize the optimal allocation as an equilibrium. Define $Z(t) \equiv \frac{U}{2} - \int_0^{\bar{x}(t)} m(x,t)x\,dx \geq 0$ and note that this non-negative magnitude is the difference between the average $x$ in the population, $U/2$, and the average $x$ among those who have not adopted the technology (the integral term). Comparing the p.d.e. for the Lagrange multiplier $\lambda$ in equation (29) with the p.d.e. for $D$ which characterizes the equilibrium in equation (9), we see that these equations only differ in the term $\theta_n Z(t)$ in the flow. Thus, if agents that adopt the technology were given a flow subsidy $\theta_n Z(t)$ every period after they have adopted, then the planner allocation would be an equilibrium. Note that

$\theta_n Z(t)$ contains the inframarginal valuation of the technology for those that use it. So, this subsidy corrects the externality. We summarize this discussion in the following proposition.

PROPOSITION 10. Fix an initial condition $m_0$ and the solution of the planner's problem $\{\bar{x}, \lambda, m\}$. The planner's allocation coincides with an equilibrium for the same initial conditions with a time varying subsidy paid to adopters. The flow subsidy paid at time $t$ to those that have adopted at $t$ or before is given by $\theta_n Z(t)$ where

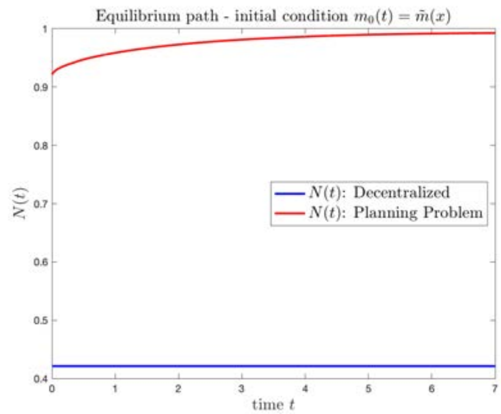
$$Z(t) \equiv \tfrac{U}{2} - \int_0^{\bar{x}(t)} m(x,t)\, x\, dx \quad \text{for all } t \geq 0 \tag{31}$$

The subsidy $\theta_n Z$ is independent of $x$.

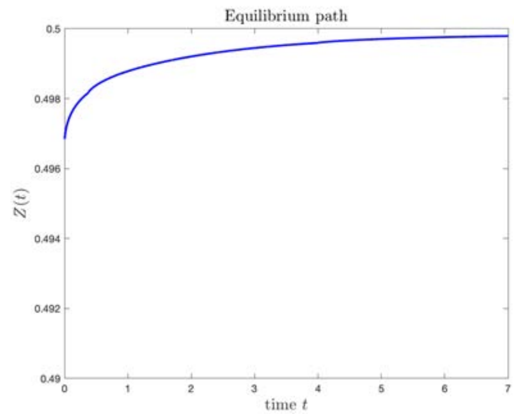
For future reference, we define $Z \equiv \mathcal{Z}(\bar{x}; m_0)$ as the solution of the path for $Z$ defined in equation (31). In particular, given $\bar{x}$ and $m_0$, using the KFE one solves for the path of $m$, and computing the integral in equation (31) gives $Z$. Consider the path $\bar{x}$ that solves the p.d.e. $\rho\lambda(x,t) = x\,(\theta_0 + \theta_n N(t)) + \theta_n Z(t) + \frac{\sigma^2}{2}\lambda_{xx}(x,t) + \lambda_t(x,t)$ with the three boundaries given in equation (30) given the paths of $N$ and $Z$ and terminal condition $\lambda(x,T) = \lambda_T(x)$. For future reference, we define $\bar{x} = \mathcal{X}^P(N, Z; \lambda_T)$ to denote the functional, which is defined as the $\mathcal{X}$ in Section 2.1 and where the superscript $P$ denotes the planning problem. Note that, using the definitions for $\mathcal{X}^P, \mathcal{Z}$ and $\mathcal{N}$ the planner's problem must satisfy the fixed point $\bar{x}^* = \mathcal{H}(\bar{x}^*, \lambda_T, m_0)$ where $\mathcal{H}(\bar{x}; \lambda_T, m_0) \equiv \mathcal{X}^P(\mathcal{N}(\bar{x}; m_0), \mathcal{Z}(\bar{x}; m_0); \lambda_T)$. We can use the same type of analysis, based on monotonicity, to characterize the solution to this fixed point problem, and to compute it.

Figure 3 illustrates an application of the optimal subsidy to reach a high adoption equilibrium. In Panel (a) of the figure, we plot the time path of the share of adopters, $N(t)$, for the stationary high-adoption equilibrium and for the planning problem. Since the initial distribution of non-adopters corresponds to the stationary equilibrium, the path of $N(t)$ is constant. Instead, the path for $N(t)$ in the planning problem jumps on impact (at the time the subsidy appears) and converges gradually to the stationary distribution of the planning problem (see Appendix D.1 for a characterization of this stationary distribution). Panel (b) shows the time path of the optimal subsidy, $Z(t)$, which starts at the value $Z(0) = \frac{U}{2} - \int_0^{\bar{x}_H} \tilde{m}(x) x dx$ and increases over time thereafter. In this example, although the high-adoption equilibrium has an interior solution, the planning problem mandates close to full adoption in stationary distribution.

Figure 3: Planning Problem: $m_0(x) = \tilde{m}(x)$




(a) $N(t)$




(b) $Z(t)$

# 7 Application: SINPE, A Digital Payments Platform

In May 2015, the Central Bank of Costa Rica launched SINPE Móvil (hereafter, SINPE), a digital platform that allows users to make money transfers between each other using their mobile phones.[12] To use SINPE, users must have a bank account at a financial entity and link this account to their mobile number. According to the Central Bank of Costa Rica, SINPE's main goal was to become a mass-market payment mechanism that could reduce the demand for cash as a method of payment. As such, SINPE was originally designed to be used for relatively small transfers, which are not subject to any fee as long as they do not exceed a daily sum. The maximum daily amount transferred without a fee varies by bank; for most users, it is approximately \$310, although some banks have lower limits of \$233 and \$155.[13] The average size of transactions in SINPE is about \$50, and has slowly decreased over time, as shown in Figure G2.

## 7.1 Data

**SINPE Transactions** Our data on SINPE usage is comprehensive: For each user in the country, we have official records on the *exact date* when she adopted the technology, along with records on each transaction made. In particular, for each transaction, the data records the *amount transacted* along with the individual identifier of *the sender and the receiver* of the money. Records also include the sender's and the receiver's bank. Importantly, this

[12]SINPE is an acronym for the initials of "National Electronic Payment System" (*Sistema Nacional de Pagos Electrónicos*), in Spanish.
[13]Respectively, these limits in dollars correspond with approximately 200,000; 150,000; and 100,000 Costa Rican colones.

information is available, not only for individuals, but also for firms.

**Family Networks and Demographics** Data on nationwide family networks is available from Costa Rica's National Registry. In particular, these data records, for each citizen, if he or she is married, to whom, and who their children are. Thus, it is possible to reconstruct each person's family tree. We find that the average number of first-degree, second-degree, and third-degree relatives is 6.4 (median 5), 10.9 (median 9), and 22.0 (median 18), respectively. The data includes individual identifiers that can be linked to SINPE. The data is dynamic, meaning that we can see how family networks are changing over time between 2015 and 2021. The same data source provides details on individual demographics.

**Networks of Coworkers, Income, and Occupation** Matched employer-employee data is obtained from the Registry of Economic Variables of the Central Bank of Costa Rica, which tracks the universe of formal employment and labor earnings. The data include *monthly* details on each employee, including her occupation, earnings, and employment history spanning SINPE's lifetime (2015-2021).[14] The average number of coworkers in our sample is 4.7 (median 1). Using this data, we can identify which people are working at the same firm in a given month to construct networks of coworkers which can be matched to SINPE records. Networks of coworkers vary at a monthly frequency as people change employers.

**Networks of Neighbors and Residential Location** We construct networks of neighbors for all adult citizens in the country leveraging data from the National Registry and the Supreme Court of Elections. The data consist of official records on the residence of each citizen, along with his or her identifier. While the records include each person's district of residence, and there are 488 districts across the country, they also include the voting center which is closest to the citizen's residence, with 2,059 centers in total. Thus, we leverage the latter to get a more precise notion of a person's neighborhood. Approximately, 1,670 adults are assigned to each voting center, on average (median 613).

**Firm-Level Data** We leverage data on corporate income tax returns from the Ministry of Finance, which cover the universe of formal firms in the country and contain typical balance sheet variables, including sales, input costs, and net assets. The data start in 2005 to 2021 and includes details on each firm's sector and location.

[14] It is worth noting that informal workers are a relatively small share of all workers in Costa Rica (27.4%), which is significantly below the Latin American average of 53.1% (ILO, 2002).

## 7.2 From Model to Data: Stylized Facts

As described in the previous section, we obtained (i) individual-level data on networks of neighbors, coworkers, and relatives from official sources; and (ii) transaction-level data including information on the senders and receivers who took part in each transaction since the app's inception. Moreover, crucially, we can link identifiers in (i) and (ii). We leverage this substantial data effort to construct measures of networks ($N$) for *each* individual and to obtain individual-level measures of adoption at the extensive and intensive margins. This will enable us to document six stylized facts consistent with the model, along with evidence of selection ($x$) and strategic complementarities ($\theta_n$).

*Fact 1: Most transactions are peer-to-peer.* In theory, firms are allowed to adopt SINPE and conduct transactions within the app. In practice, however, transactions involving firms represent a small fraction of all payments. In fact, as shown in Figure G3, individual-to-individual transactions account for over 95% of all transactions, regardless of the time period considered.[15] This motivates us to study adoption through the lens of our model while focusing on peer-to-peer transactions where small agents trade with each other, rather than one with a few non-atomistic players (large firms).

*Fact 2: Individuals "belong" to networks.* We can identify different types of networks for each user. In particular, we could identify which transactions take place within an individual's network of neighbors, coworkers, or relatives. To do so, we construct the network of neighbors of each user—which would correspond with the people assigned to her voting center—and calculate the number and total value of SINPE transactions involving another user who also resides in the same neighborhood. Similarly, we construct the network of coworkers for each employed user based on employer-employee data. Finally, we construct family networks taking into account relatives up to a third-degree of kinship.

In Table 1, we document that most transactions involve a counterpart who belongs to at least one of these networks:[16] 43% of all transactions have a neighbor as counterpart, 44% of all transactions are among coworkers, and 28% are conducted with relatives. We can also consider the *union* of all three networks described above, and document that about three-quarters of all transactions take place with someone within at least one of the three types of networks. Moreover, we also document that users have relatively few peers with whom they transact. Before 2019, each user had less than two distinct connections per month,

[15] This finding holds if we instead consider unweighted number of transactions, as shown in Figure G4.

[16] Table 1 calculates shares using 2018 data; the midpoint of our sample period. Results remain quite similar if, instead, we consider the average shares of transaction for the entire sample period, as shown in Table G1.

both as a sender and as a receiver. By the end of 2021, this number had increased; each user had just over six distinct *monthly* connections and the average *total* number of distinct connections per user was 44, i.e., people do not necessarily transact with the *same* six peers each month.[17] The average transaction size is \$46 (median \$48), and has decreased slowly over time, as shown in Figure G2.

Table 1: Share of Transactions Within Network

| | (1) Neighborhood | (2) Firm | (3) Family | (4) Union of all three |
|---|---|---|---|---|
| Neighborhood | 0.43 | | | |
| Firm | 0.62 | 0.44 | | 0.72 |
| Family | 0.55 | 0.66 | 0.28 | |

*Notes:* We construct average shares using data on transactions per user from 2018, i.e., the middle of our sample period. Shares using the entire sample—from May 2015, when the technology was introduced, to December 2021—are shown in Table G1.

*Fact 3: The adoption of the technology within each network was gradual.* We classify networks (i.e., neighborhoods, families, firms) according to their level of adoption. In particular, we calculate the share of individuals within a network who had adopted SINPE by December 2021, the last period available in our data set. We then compute percentiles of this share across networks to generate a distribution. Figure 4 displays the diffusion path of the technology for the median neighborhood and the median family network in Panels (a) and Panel (b), respectively. The panels show that the technology was adopted gradually within networks.[18] While Figure 4 is computed based on networks of neighbors and relatives, the same patterns emerge when analyzing networks of coworkers, as shown in Figure G7.

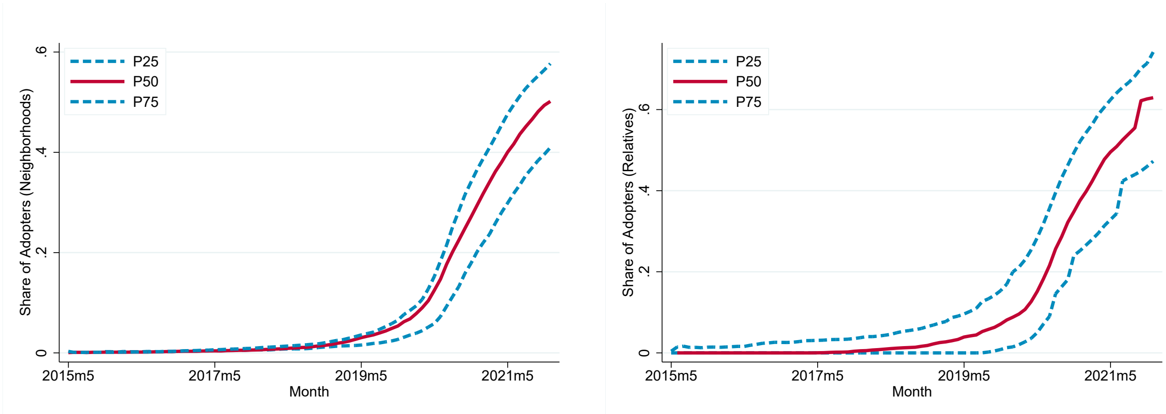
*Fact 4: There is evidence of selection at entry.* Through the lens of our model, early adopters—who started using the technology even when the network was small—should be more intense users (with higher $x$). Consistent with this notion, we document that early adopters have distinct characteristics as compared with users who adopted later. For this exercise, and throughout the entire paper, we classify an individual as an adopter starting from the time when she first used the app. First, as shown in Figure 5, we find that early adopters have a higher average wage as compared with individuals who adopted later (Panel (a)), and are on average more high-skill (Panel (b)).[19] Early adopters are also younger, on

[17] Average monthly patterns are documented in Figure G5.

[18] The diffusion of technology is displayed after controlling for the prevalence of COVID-19 at the local level over time.

[19] We classify an occupation as high-skill if it requires education or training beyond a high-school diploma. The dashed vertical line in each figure denotes the beginning of the pandemic, which just as in Figure G1

Figure 4: Gradual Diffusion Within Networks




(a) Networks of Neighbors (b) Family Networks

*Notes*: The figures show the patterns of diffusion of the technology within networks across different percentiles of the distribution of networks. Percentiles are calculated in the last period of the sample using the share of individuals that had adopted the technology. Panel (a) defines networks as neighborhoods, while Panel (b) considers family networks.

average, than later adopters, as shown in Figure G6.

Second, we can more closely bring the model to the data by interpreting the flow benefit of agents who adopt the technology as being proportional to how intensively they use SINPE. Specifically, suppose SINPE users choose the intensity with which they use the application, $\xi_t$, maximizing the following expression:

$$\xi_t^*(x_t, N_t) = \arg\max_{\xi_t} \frac{1+p}{p}\left[\beta(x_t, N_t)\xi_t - \frac{\xi_t^{1+p}}{1+p}\right],$$

where $p > 0$ so that the problem is convex and $\beta(x_t, N_t) > 0$. The first order condition describes the optimal intensity in which the technology is used: $\xi_t^*(x_t, N_t) = \beta(x_t, N_t)^{1/p}$, and we can choose the function $\beta(x_t, N_t)$ such that the indirect utility function gives the specified flow benefit, i.e:

$$[\theta_0 + \theta_n N_t]x_t = \max_{\xi_t} \frac{1+p}{p}\left[\beta(x_t, N_t)\xi_t - \frac{\xi_t^{1+p}}{1+p}\right] \text{ for all } x_t \in [0, U] \text{ and } N_t \in [0, 1].$$

The solution is given by $\beta(x_t, N_t) = [(\theta_0 + \theta_n N_t)\, x_t]^{\frac{p}{p+1}}$; combining this expression with the first-order condition and taking logs, we obtain:

$$\ln \xi_t^* = \frac{1}{1+p} \ln\left[(\theta_0 + \theta_n N_t)\right] + \frac{1}{1+p} \ln x_t. \tag{32}$$

did not have a major impact on overall trends.

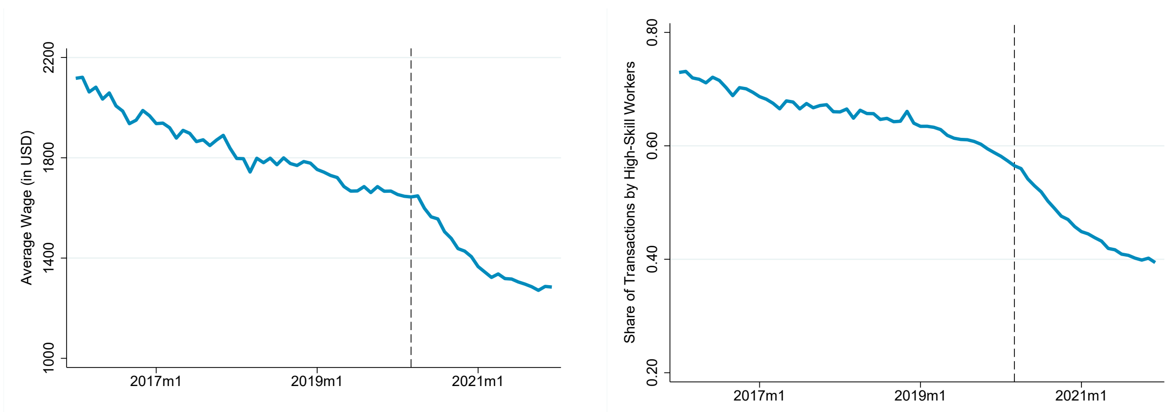
In this equation, note that if we were to remove the network×time variation, then $\ln \xi_{it}^{n}$ would proxy for $\ln x_t$, as through the lens of the model only the idiosyncratic variation would remain. Moreover, the model also predicts that users with a higher $x$ would adopt the technology earlier. Thus, we can obtain a relation between intensity of usage ($\xi_{it}^{n}$) and the share of user's $i$ network who had adopted *when she used the app for the first time* ($N_{i,entry}^{n}$):

$$\ln \xi_{it}^{n} = \gamma + \zeta N_{i,entry}^{n} + \lambda_t^n + \nu_{it}^{n},$$

where $n \in$ {neighbors, coworkers, relatives} and $\xi_{it}^{n}$ is defined as number of transactions of user $i$ each month $t$. Our model predicts that $\zeta < 0$, as users who adopted the app ("entered") when the network was smaller should have a higher idiosyncratic taste for the app and use it more intensively—note that the inclusion of the network-time fixed effect, $\lambda_t^n$, prevents this relationship from being mechanical.

We estimate $\hat{\zeta}$ to be $-2.7$ when defining a network as a neighborhood. This relationship is shown in Column (1) of Table 2, and while suggestive, points to the presence of selection at entry. The relation is also robust to defining networks using coworkers and relatives, as shown in Columns (2) and (3) in Table 2. The relation also holds if, instead of the total number of transactions, we consider the value of transactions as our dependent variable, as reported in Table G2.

Figure 5: Average Wage and Skill at the Time of Adoption




(a) Wage

(b) High Skill

*Notes*: Panel (a) shows the cross-sectional distribution of SINPE users' monthly wages in USD. Panel (b) shows the cross-sectional distribution of SINPE users' skills. High skill users are those that are in an occupation that requires more than a high school degree. Both panels show averages weighted by the number of transactions of each user. Both figures include a vertical dashed line to mark the start of the COVID-19 pandemic (March 2020).

*Fact 5: There is evidence of strategic complementarities.* The core idea behind strategic complementarities is that usage benefits increase in the size of an user's network. To explore

Table 2: Number of Transactions and Size of Network at Entry

| *Dependent variable: Number of Transactions (IHS)* | | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Size of Neighbors' Network at Entry | -2.730*** | | |
| | (0.005) | | |
| Size of Coworkers' Network at Entry | | -1.300*** | |
| | | (0.005) | |
| Size of Family Network at Entry | | | -1.181*** |
| | | | (0.006) |
| Observations | 34,409,818 | 16,138,736 | 14,700,288 |
| Network×Time/Cohort FE | Yes | Yes | Yes |
| Adjusted R-squared | 0.234 | 0.304 | 0.199 |

*Notes:* The dependent variable in this estimation is the number of transactions each month for each user transformed using the inverse hyperbolic sine function. Coefficients describe the effect of increasing the share of an individual's network who had adopted the app at the time when she used it for the first time. All regressions control for network size (in levels) and use data from May 2015, when the technology launched, to December 2021. Standard errors, clustered by individual, are in parenthesis.

this notion further, recall the model-derived expression in equation (32). Under this interpretation of the model, the intensity with which the application is used, which is observable in the data (e.g., number or value of transactions), is proportional in logs to the flow benefit of adopting the application as described in the model. After taking the first order Taylor expansion of $\ln(\theta_0 + \theta_n N_t)$ around $N^*$ and plugging it into equation (32), we obtain:

$$\ln \xi_t^* = \ln(\theta_0 + \theta_n N^*) + \frac{1}{1+p}\frac{\theta_n(N_t - N^*)}{\theta_0 + \theta_n N^*} + \frac{1}{1+p}\ln x_t. \tag{33}$$

Moreover, taking first differences, it follows that:

$$\Delta \ln \xi_t^* = \beta \Delta N_t + \nu_t, \tag{34}$$

where $\beta \equiv \frac{1}{1+p}\frac{\widetilde{\theta}}{1+\widetilde{\theta}N^*}$ and $\nu_t \equiv \frac{1}{1+p}\Delta \ln x_t$. Further, in the case of a quadratic adjustment costs (i.e., $p = 1$), then $\widetilde{\theta} = \frac{2\beta}{1-2N^*\beta}$. Thus, throughout all the tables in this section, we can evaluate $N^*$ at its mean value to recover $\widetilde{\theta}$ from each $\beta$; these are our coefficients of interest since strategic complementarities in the adoption of the technology exist if $\beta > 0 \iff \widetilde{\theta} > 0 \iff \theta_0 > 0$ and $\theta_n > 0$. Note that equation (34) is in *differences*, therefore, *any individual or network characteristics which are time invariant would cancel out.*

Now, in the data, we have many networks—for example many neighborhoods across the country—thus so we consider the following version of equation (34):

$$\Delta \ln \xi_{it}^n = \gamma + \beta \Delta N_t^n + \psi \Delta X_t^n + \lambda_t + \lambda_c + \epsilon_{it}^n, \tag{35}$$

where $\ln \xi_{it}^n$, the intensity with which individual $i$ in network $n$ uses the technology, and can be interpreted as either the value or the number of SINPE transactions in a given month $t$, $N_t^n$ is share of user $i$'s network that has adopted the app, $\Delta X_t^n$ is a vector of controls at the network×time level, which includes the change in the size of network $n$ in *levels* and the change in the number of COVID-19 cases; we also include time fixed-effects, $\lambda_t$, which captures the fact that aggregate adoption is increasing over time along with cohort fixed-effects, $\lambda_c$, which through the lens of the model controls for selection into the app. Again, networks can be defined in different ways, and as such $n \in$ {neighbors, coworkers, relatives}. This regression considers only the intensive margin of adoption, and thus allows us to isolate the effect of strategic complementarities from any other learning externalities which might be active when studying the extensive margin of adoption.[20]

Table 3 shows results when considering $n$ as a user's network of neighbors, coworkers, and relatives. The dependent variable refers to the number of SINPE transactions transformed using the inverse hyperbolic sine (IHS) function. Across specifications, we find that $\beta$ remains positive and statistically significant. Further, the coefficients corresponding with each network remain stable when considering all of them simultaneously in Column (4) of Table 3. Findings remain unchanged if we consider either alternative transformations to IHS or the monthly value of transactions of each user as our dependent variable, instead of the number of transactions, as reported in Table G3 and Table G4.[21]

It is also possible to use an alternative measure of $N$, which *by construction* comprehends all transactions. Our starting point is the last period in our sample (December 2021)—in which most adults have already adopted. Then, we look back in time at all transactions which have occurred, and, for each individual, we define her network as the collection of people with whom she transacted at some point in time. Thus, for instance, the share of adopters in someone's network in 2016 will have all her connections who have adopted in the numerator, and all her past *and future* connections in the denominator. Table G5 shows the results of estimating equation (35) using this alternative network and the number of transactions per user as our dependent variable: a positive correlation between changes in usage and in share of adopters within network is always present across specifications.[22]

[20] For instance, an individual might be more likely to learn about the existence of the app if she has more friends who have adopted the app.

[21] Panel (a) of Table G3 runs the regression with a dependent variable in logs. Panel (b) repeats the exercise transforming the value of transactions following Davis and Haltiwanger (1992) (i.e. $\Delta x_t = 2\frac{x_t - x_{t-1}}{x_t + x_{t-1}}$). We prefer the inverse hyperbolic sine function over a transformation using logs, as it is frequent to find zero transactions for individuals in a given month, even after they adopt the technology.

[22] Table G6 displays results considering instead the value of transactions per user.

Table 3: Changes in Number of Transactions and Network Changes

| *Dependent variable: IHS(Δ Number of Transactions)* | | | | |
|---|---|---|---|---|
| Δ Share Neighborhood Adopters | 1.008*** | | | 0.879*** |
| | (0.022) | | | (0.031) |
| Δ Share Coworkers Adopters | | 0.238*** | | 0.232*** |
| | | (0.007) | | (0.007) |
| Δ (Log) Wage | | 0.044*** | | 0.044*** |
| | | (0.001) | | (0.001) |
| Δ Share Relatives Adopters | | | 0.273*** | 0.308*** |
| | | | (0.003) | (0.004) |
| Observations | 32,391,602 | 16,232,003 | 30,633,379 | 15,355,945 |
| Time/Cohort FE | Yes | Yes | Yes | Yes |
| Adjusted R-squared | 0.014 | 0.019 | 0.015 | 0.020 |

*Notes:* The unit of observation is the individual. The dependent variable is transformed using the inverse hyperbolic sine function; Table 3 shows results with alternative transformations. All regressions control for network size (in levels) and for the number of COVID-19 cases (using an inverse hyperbolic sine transformation). We run regressions using data from May 2015, when the technology was introduced, to December 2021. Standard errors (clustered by individual) are in parentheses.

**Leave-One-Out Instrument and Balanced Panel** *Fact 5* above speaks to a *correlation* between the changes in the intensity with which someone uses the app and changes in the share of individuals in her network who have adopted it. We will proceed by refining our analysis of this relationship. First, we construct a leave-one-out instrument to address concerns related with mis-measurement and local common shocks. In particular, instead of focusing on the change in the degree of adoption in an individual's neighborhood, we instrument for it using the weighted average of adoption in neighborhoods which are in close proximity—but outside—an individual's own *district* (Costa Rica has 488 districts and 2,059 neighborhoods), where the weight is the inverse of distance divided by the total sum of distances.[23]

Table G7 shows the IV results. The first stage is in the table's top panel, and we estimate equation (35), but with our instrument as the main independent variable, in the bottom panel using the number of transactions as the dependent variable. Estimated coefficients are smaller than those of the OLS when implementing this IV strategy, as would be expected in the presence of common shocks, but remain positive and highly significant.[24]

Another possible refinement, this time to address concerns regarding selection in and out of the app, is to study the how changes in usage depend on changes in the share of adopters *following a balanced panel of adopters across time.* Thus, we repeat our estimations but now on a balanced panel of users who had already adopted by 2016 in Table G9. Again, our results are robust to the estimation using this subset of users; while coefficients are slightly

[23]Namely, for neighborhood $j$ in district $d$, we consider: $\sum_{k \neq d} \frac{(dist_{jk})^{-1}}{\sum_n (dist_{jn})^{-1}} N^n_{kt}$, that is, the (distance-weighted) total transactions in nearby neighborhoods indexed by $j$.

[24]Results considering the value of transactions as the dependent variable are reported in Table G8.

smaller—as would be expected in the case of positive selection into the app—they remain similar to the ones using the entire sample.

## 7.3 Changes in Networks of Coworkers After a Mass Layoff

We first documented a *correlation* between the intensity with which someone uses the app and the share of individuals in her network who have adopted it (*Fact 5*). The last section also performed some refinements addressing common shocks and mis-measurement (using a leave-one-out instrument) and selection (focusing on a balanced panel). In this section, we consider an alternative identification strategy to provide more evidence on the relationship between changes in usage and in the share of adopters being causal. This strategy focuses on the network of coworkers and implements a movers design, where we focus on the workers displaced during mass layoffs to examine the effect of network changes on the extensive and intensive margins of adoption.[25]

**Extensive Margin of Adoption** To analyze changes in the extensive margin of adoption, we consider the change in the probability of adoption for displaced workers *who had not adopted the app by the time they were rehired*, and how it depends on the change in the share of coworkers who had SINPE at their old and new firm. The main hypothesis of this exercise is that workers who were displaced during a mass layoff, and who ended up at firms where a larger share of colleagues had SINPE (larger $N$), have more incentives to adopt via the effect of strategic complementarities. We consider:

$$\begin{aligned} Adopt_i =& \alpha + \hat{\theta}\Delta N_i^{coworkers} + \hat{\gamma}\Delta \ln wage_i + \hat{\psi}\Delta \ln size_i + \hat{\lambda} date\, hired_i + \\ & \tilde{\nu} \ln \sum_{t=0}^{move} \left( \tilde{\xi}_{t,\text{ new firm}} - \tilde{\xi}_{t,\text{ old firm}} \right) + \hat{\omega}\Delta Covid_i + \epsilon_i, \end{aligned} \tag{36}$$

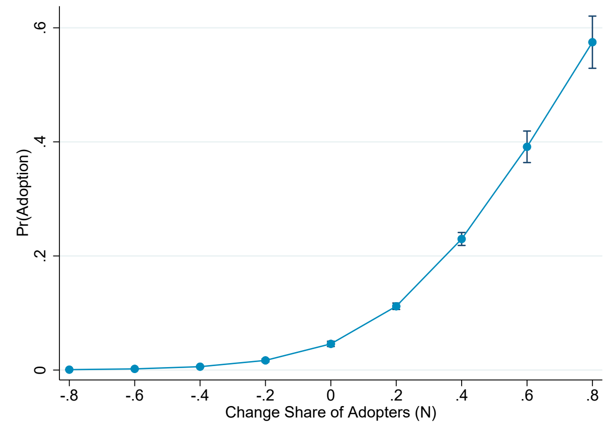
where $Adopt_i$ equals one if individual $i$ adopted SINPE within 6 months after arriving to her new firm, and zero otherwise; $\Delta N_i^{coworkers}$ is the change between the share of coworkers who had adopted at the old and the new employer; $\Delta \ln wage_i$ corresponds with the change in the average wage (in logs) across 6 months before the layoff and after the rehiring; $\Delta \ln size_i$ is the change in the number of workers in each firm, $date\, hired_i$ controls for the date in which individual $i$ was hired by the new firm; $\ln \sum_{t=0}^{move} \left( \tilde{\xi}_{t,\text{ new firm}} - \tilde{\xi}_{t,\text{ old firm}} \right)$ is the difference in the historical transactions made by workers at the new firm and the old firm up until the move occurred, which aims to control for factors, other than strategic complementarities,

[25] To define a mass layoff, we follow Davis and Von Wachter (2011) and identify establishments with at least 50 workers that contracted their monthly employment by at least 30% *and* had a stable workforce before this episode and did not recover in the following 12 months. More details are provided in Appendix G.4.1.

Figure 6: Adoption Probability and Changes in Coworkers' Network After a Mass Layoff

| | (1) | (2) | (3) |
|---|---|---|---|
| $\Delta N_i^{coworkers}$ | 8.298*** | 5.754*** | 5.420*** |
| | (0.119) | (0.158) | (0.163) |
| | [0.450] | [0.376] | [0.350] |
| $\Delta \ln wage_i$ | | -0.024 | -0.035 |
| | | (0.034) | (0.035) |
| $\Delta Covid_i$ | | | 0.112*** |
| | | | (0.016) |
| Observations | 22,249 | 17,658 | 17,658 |
| Time/Cohort FE | No | Yes | Yes |
| Pseudo R2 | 0.495 | 0.525 | 0.528 |

(a) Changes in Adoption Probability




(b) Marginal Effect of Network Changes

*Notes:* Panel (a): The unit of observation is the individual. We run regressions using data on mass layoffs that occurred between May 2015, when the technology was introduced, and December 2021. Standard errors are in parentheses. Marginal effects for the main variable of interest are reported in brackets. Panel (b): This figure plots the marginal effect of $\Delta N_i^{coworkers}$ in the specification described by Column (3) of Panel (a) in this figure. Vertical bars denote 95% confidence intervals.

which might facilitate adoption at the new vs. the old firm; and $\Delta Covid_i$ controls for the change in the cumulative COVID-19 cases (transformed using the inverse hyperbolic sine function) in the individual's neighborhood across the 6 months before the layoff and after the rehiring. Appendix G.4.1 provides more details on each of these variables.

Panel (a) of Figure 6 shows the results estimating equation (36) using a logit model. The marginal effects of changes in network adoption are reported in brackets. We consistently find that workers who, after a mass layoff, were hired by firms where the rate of SINPE adoption was higher than their previous employer's are more likely to adopt SINPE than their counterparts who moved to firms where the change in their coworkers' rate of adoption was smaller. The marginal effect of $\Delta N_i^{coworkers}$, under the specification described by Column (3) of Panel (a) is shown in Panel (b) of Figure 6. This marginal effect is monotonous and, as expected, is present only when the change in the share of adopters is positive.

**Intensive Margin of Adoption** It is also possible to estimate the relationship between share of adopters within one's network and intensity of usage. To do so, we again focus on workers who were fired during a mass layoff, but this time consider only displaced workers *who had already adopted and had used SINPE at least once by the time they were fired*. We then examine how the intensity with which they use the app changes depending on the change in the share of coworkers who had SINPE at their old and new firm. As explained in the former subsection, it is possible to derive the relationship in equation (35) from our theoretical model, which speaks to the technology's usage intensity. Thus, we consider:

$$\begin{aligned}\Delta \ln \tilde{\xi}_i =& \tilde{\alpha} + \tilde{\theta}\Delta N_i^{coworkers} + \tilde{\gamma}\Delta \ln wage_i + \hat{\psi}\Delta \ln size_i + \tilde{\lambda} date\, hired_i + \\ & \tilde{\omega}\Delta Covid_i + \tilde{\delta}\lambda_{ic} + \tilde{\nu} \ln \sum_{t=0}^{move} \tilde{\xi}_i + \tilde{\nu} \ln \sum_{t=0}^{move} \left( \tilde{\xi}_{t,\text{ new firm}} - \tilde{\xi}_{t,\text{ old firm}} \right) + \tilde{\epsilon}_i, \end{aligned} \tag{37}$$

where $\Delta \ln \tilde{\xi}_i$ refers to the change in monthly intensity with which individual $i$ used SINPE within 6 months *after* arriving to her new firm compared with 6 months *before* being fired, $\lambda_{ic}$ controls for cohort (i.e., the date when individual $i$ adopted SINPE), and $\ln \sum t = 0^{move} \tilde{\xi}_i$ is the sum of all historical transactions made by agent $i$ since she adopted the app. Other variables are defined in the same way as in equation (36).[26]

This is our preferred specification for several reasons. First, while the results in Figure 6 could be partly driven by learning, this is less likely to happen under equation (37) as (i) workers had already adopted the app when they were fired—and we define "adoption" as making at least one transaction—so they were at least aware of the app's existence and had used it before; (ii) we control for tenure in the app (i.e., the cohort when the user adopted) and for the historical number of transactions in the app, which as shown before correlate with observables like age, skill, and wage. These variables aid in controlling for characteristics that are particularly relevant for intensity of usage and are also useful to address learning to better use the app after adopting. Second, of course, the choice of the new firm after a mass layoff is not exogenous, but this does not pose a measurement problem as long as sorting is not (both): (i) stronger after a mass layoff—note that there is no reason why this might be the case, especially as results hold even when we focus on job-to-job transitions, where workers had little time to find a new job after their being fired exogenously—and (ii) not controlled for by the cohort of the mover, which proxies for her idiosyncratic characteristics, and difference in the historical transactions at the new vs. the old firm. The latter control, in particular, helps us rule out stories where, for instance, workers select into firms where people use the app more intensively for reasons other than strategic complementarities (for instance, their demographics or the quality of internet at the firm).

Table 4 displays our results using the number of transactions per user as our dependent variable.[27] As with the extensive margin, changes in the intensity of usage depend positively and significantly on the change in the share of adopters at the old and new firm. Panel (a) of Figure 7 displays the marginal effect of these network changes following the specification described by Column (3) of Table 4. As this panel shows, not only is the relationship between usage and network changes positive, but also whenever a worker moves to a firm with a lower

[26] Appendix G.4.1 provides more details on these variables and the choices made to conduct this exercise.
[27] Table G10 reports the same results with the value of transactions as dependent variable.

Table 4: Intensity of Usage and Changes in Coworkers' Network After a Mass Layoff

*Dependent Variable: Δ Number of transactions (inverse hyperbolic sine)*

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| $\Delta N_i^{coworkers}$ | 2.460*** | 1.500*** | 1.008*** | 0.943*** |
| | (0.153) | (0.191) | (0.196) | (0.197) |
| $\Delta \ln wage_i$ | | 0.400*** | 0.349*** | 0.363*** |
| | | (0.046) | (0.044) | (0.048) |
| $\Delta Covid_i$ | | | 0.165*** | 0.155*** |
| | | | (0.020) | (0.023) |
| Observations | 1,554 | 1,554 | 1,554 | 1,554 |
| Time FE | No | Yes | Yes | Yes |
| Cohort FE | No | No | No | Yes |
| Adjusted R-squared | 0.141 | 0.222 | 0.257 | 0.280 |

*Notes:* The unit of observation is the individual. We run regressions using data on mass layoffs which occurred between May 2015, when the technology was introduced, until December 2021. While time and cohort fixed-effects' inclusion varies across columns, all other independent variables in equation (37) are present across columns. Standard errors are in parentheses.

adoption rate, her usage decreases (i.e., the change on the vertical axis is negative).[28]

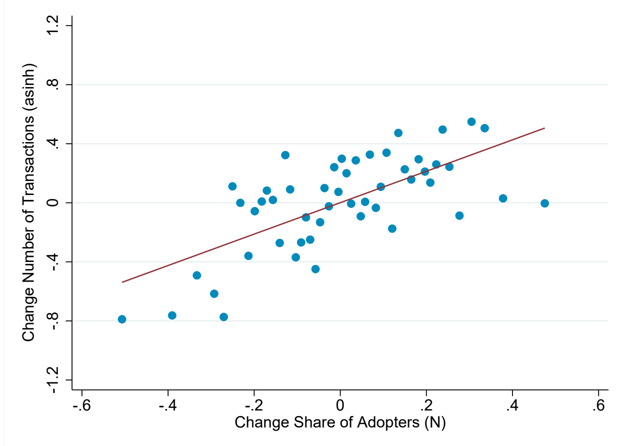
Column (4) controls for cohort, i.e., date of adoption, which aims to mitigate any effect of more experienced users behaving differently than beginners. Column (4) also controls for the total historical transactions made, which in a similar spirit as cohort, intends to mitigate any effect coming from learning how to use the app from others. Interestingly, as compared with Column (3), adding these controls does not change the coefficient of interest. This result aligns with the following intuition: while at the extensive margin it is hard to disentangle between strategic complementarities and "learning from others" about the technology, at the intensive margin—once users have already adopted and used the app—a learning story is less plausible, as reflected by $\tilde{\theta}$ not changing after controlling for cohort and historical usage.

The analysis can be taken to an even more detailed level if, instead of considering all transactions in the left-hand-side variable, we focus only on those which had a coworker as a counterpart. This subsample allows us to better identify changes in usage intensity which are a direct consequence of the arguably exogenous changes in the network of coworkers. Reassuringly, results are remarkably similar to those using all transactions, as shown in Panel (b) of Figure 7 and Table 4.[29]

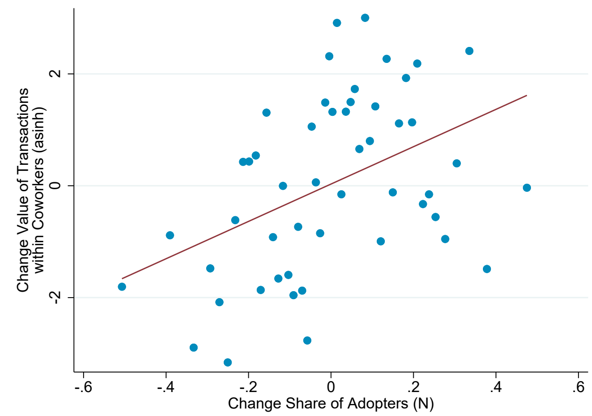
[28] The marginal effect considering the value of transactions as dependent variable, as opposed to the number of transactions, is reported in Figure G8.

[29] The corresponding results using the value instead of the number of transactions are reported in Table G10.

Figure 7: Marginal Effect of Network Changes on Usage Intensity




(a) All transactions




(b) Transactions with coworkers only

*Notes:* Panel (a) plots the marginal effect of $\Delta N_i^{coworkers}$ in the specification described by Column (4) of Table 4. Bars denote 95% confidence intervals. The dependent variable in this estimation is the number of transactions (transformed using the inverse hyperbolic sine function) on each period for each user. Panel (b) is similar, but differs as the dependent variable in this estimation is the number of transactions *which have a coworker as a counterpart* (transformed using the inverse hyperbolic sine function) on each period for each user.

# 8 Quantitative Performance and Optimal Subsidy

In this section, we calibrate our model and evaluate its performance relative to SINPE data. We begin by describing an extension of the model that combines the model of strategic complementarities with a learning model. This hybrid version is helpful to make the model consistent with the evidence presented in the panels of Figure 4 where the path for $N(t)$ is smooth and relatively flat at the beginning. The presence of a learning element seems consistent with the fact that only about 5% of the adults report knowing about SINPE Movil as reported by the 2017 Survey of Payment Methods conducted by the Central Bank of Costa Rica. We next describe our calibration procedure in detail.

**A Learning Model with Strategic Complementarities:** It is straightforward to extend our benchmark model of strategic complementarities to include random diffusion of the technology across agents. We adapt Bass's (1969) model of information diffusion for new products to our setup. We assume newborn agents are initially uninformed, and become informed by randomly matching with informed agents. Thus, the variational inequality of the adoption decision (i.e., net value of adoption $a(x,t) - c$ and the net optimal value $v(x,t)$) are the same as in the model with strategic complementarities, since this decision to adopt can only be made after agents are aware of the technology. However, the law of motion of $m$

needs to be modified to include the inflow of informed agents as in a random diffusion model:

$$m_t(x,t) = \frac{\sigma^2}{2} m_{xx}(x,t) + \frac{\beta_0}{U} I(t)(1 - I(t)) - \nu m(x,t) \text{ all } t \geq 0 \text{ and } x \in [0, \bar{x}]$$
$$m(x,t) = 0 \text{ all } t \geq 0 \text{ and } x \in [\bar{x}, U]$$
$$m_x(0,t) = 0 \text{ all } t \geq 0$$

where $I(t)$ denotes the fraction of the population informed about the technology and $\beta_0$ indicates the number of meetings per unit of time between those informed and those uninformed (i.e., $1 - I(t)$), see Appendix E for details.[30]

**Calibration:** We calibrate the hybrid model to match the adoption path of the average neighborhood in Figure 4. We interpret the flow benefit of agents who adopt the technology as being proportional to how many transactions they conduct as described in Section 7.2 and assuming a quadratic adjustment cost (i.e., $p = 1$). By Lemma 1, $U$ can be normalized without loss of generality, so the problem features 6 independent parameters: $\nu, \rho, \theta_n, \theta_0, \sigma$, and $c$ (we use the normalization $U = 1$). In addition, the model that includes learning requires us to calibrate an additional parameter, $\beta_0$, and an initial condition for the informed population $I(0)$.

We calibrate $\nu$ to 0.0278 to match the rate at which agents stop using SINPE; namely, the average fraction of agents in 2019-2021 which had adopted SINPE but did not conduct a single transaction in the app within a year. We use the last three years of the data, when the adoption rate is higher, to focus on periods closer to a stationary equilibrium. We set the discount factor $r$ to be consistent with a 5 percent annual interest rate. This value is a lower bound for $r$, which can admit higher values if we assume agents expect new technologies to arrive in the future and replace SINPE. The values of $\nu$ and $r$ imply $\rho = r + \nu = 0.0778$.

Since we are targeting the path of the mean neighborhood, we set $\widetilde{\theta} \equiv \frac{\theta_n}{\theta_0}$ equal to 2 based on our reduced form regressions for neighborhoods.[31] We calibrate $\sigma = 0.15$ and $\theta_0 = 86$ using simulated method of moments to match various micro-data moments. Both $\sigma$ and $\theta_0$ are

[30] The appendix develops a model of pure learning featuring random diffusion of the technology across agents. In the model, agents can be either uninformed about the technology, or informed about it. If they are informed, they can decide to pay a cost $c$ and adopt it. Once an agent adopts the technology her flow benefit depends on the idiosyncratic value of the random variable $x$, but not on the size of the network, i.e., $\theta_n = 0$. The model has four main conclusions: i) it has a unique equilibrium ii) it has a logistic S shape adoption profile if the initial share of informed agents is small, iii) the use of the technology for those that adopt depends only on the cohort, and iv) the equilibrium is constrained efficient.

[31] We aim to be conservative, as this parameter admits values between 2 and 3. The estimates using a balance panel (cohort of 2016) in Table G9 imply $\widetilde{\theta} = \frac{2\beta}{1-2N^*\beta} \approx 2$ since $\beta = 0.76$ and $N^* = 0.17$. Table G7 shows that in the leave-one-out specification, $\beta = 0.69$, and using the average adoption $N^* = 0.41$ we find $\widetilde{\theta} = \frac{2\beta}{1-2N^*\beta} = 3.1$. Figure 8 shows the performance of the model for different values of $\widetilde{\theta}$.

estimated jointly so that they are consistent with the empirical distribution of transactions. Intuitively, $\sigma$ is pinned down by the dispersion of the changes in transactions and $\theta_0$ by the level of transactions in the data.[32] We set $\frac{c}{U\theta_0} = 8.5$ to match the fraction of the population which had adopted the technology by the end of 2021 and $\beta_0 = 1.35$ to match the fraction of people informed about the technology mentioned above (approximately 5% two years after the launch). In Figure 8 We display the path of adopters starting at $I(0) = 0.001$, that is, 0.1 percent of population in the average neighborhood is informed about SINPE at the time it was launched.

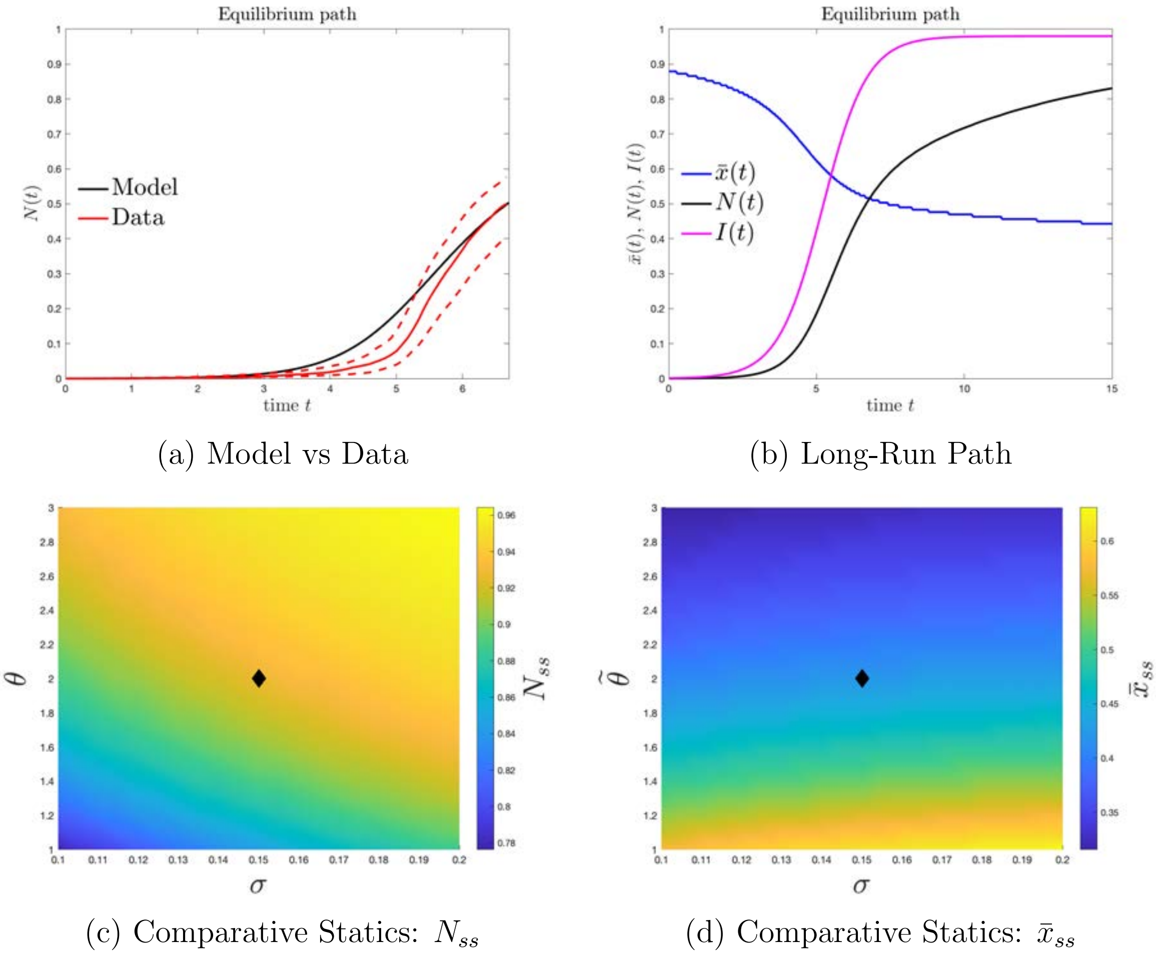
**Results:** Panel (a) of Figure 8 compares the path of adoption in the model and in the data. The solid red line indicates the diffusion of the technology in the median neighborhood and the dashed lines represent the $25^{th}$ and $75^{th}$ percentiles. The figure shows that both the speed and the level of adoption generated by the model are consistent with those in the data. Panel (b) shows the path of $I(t)$, $N(t)$ and $\bar{x}(t)$. The path of $I(t)$ shows that most people are informed about the technology within the first 7 years; in the stationary distribution, approximately 98% of the population knows about the application. Furthermore, the model predicts that in stationary distribution 93% of the population living in the median neighborhood will adopt the application, as shown by the path of $N(t)$. Importantly, the declining path of $\bar{x}(t)$ indicates that, consistent with our empirical evidence, the model features selection: agents that benefit the most from the technology adopt first.[33]

To gain further intuition on the determinants of the level of adoption in a stationary distribution, we set all parameters to their estimated values and vary $\widetilde{\theta}$ and $\sigma$, while holding others constant. Panel (c) of Figure 8 shows how the stationary level of adoption changes with $\widetilde{\theta}$ and $\sigma$ (a black diamond denotes $N_{ss}$'s level in the baseline calibration). As $\widetilde{\theta}$ increases, so does the strength of the strategic complementarities, and not surprisingly, $N_{ss}$ increases as $\widetilde{\theta}$ rises. The effect of $\sigma$ is more subtle and results from two opposing forces. On the one hand, higher $\sigma$ decreases $N_{ss}$ since agents have a higher option value of waiting to adopt. On the other hand, higher $\sigma$ increases $N_{ss}$, since it implies a smaller density of non-adopters below $\bar{x}_{ss}$. In our calibration the latter effect dominates and $N_{ss}$ increases with $\sigma$. Panel (d) displays a similar exercise for $\bar{x}_{ss}$. It shows that strategic complementarities play an important role in decreasing the adoption threshold. Moreover, given $\widetilde{\theta}$, a higher $\sigma$ unambiguously increases $\bar{x}_{ss}$.

[32] A detailed description of the simulation and estimation procedure can be found in Section H.1.

[33] Section H.2 presents a version of the model without strategic complementarities and only learning (i.e., $\widetilde{\theta} = 0$). In this case, the path of $\bar{x}(t)$ is completely flat. In contrast to what is observed in the data, a pure learning model features no selection in the adoption of the technology and is constrained efficient: the optimal subsidy to adopt the technology is zero.

Figure 8: Path of Adopters (Short-Run and Long-Run)




(a) Model vs Data

(b) Long-Run Path

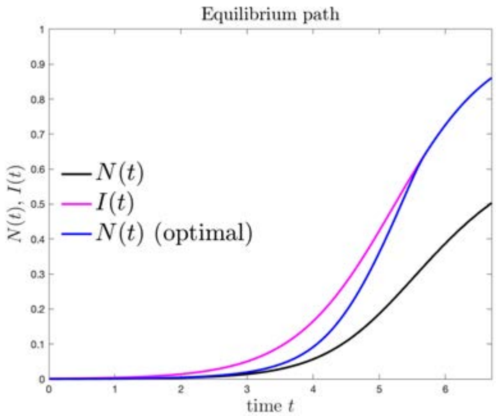
(c) Comparative Statics: $N_{ss}$

(d) Comparative Statics: $\bar{x}_{ss}$

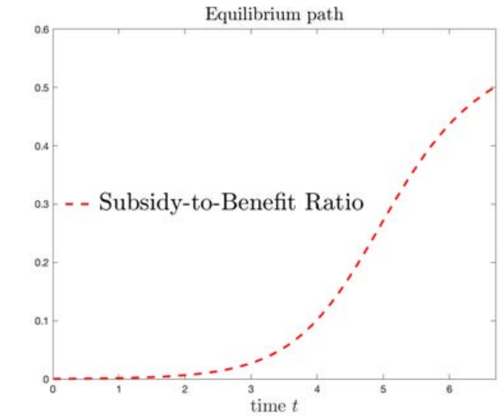
*Notes*: Panel (a) compares the path of adopters in the model and in the data. The solid red line shows the patterns of diffusion of the technology in the median neighborhood, where the percentile is calculated in the last period of the sample using the share of individuals that had adopted the technology. The dashed red lines show the $25^{th}$ and $75^{th}$ percentiles. Panel (b) shows the share of informed agents, $I(t)$, the share of adopters, $N(t)$, and the levels of $\bar{x}(t)$ predicted by the model under our baseline calibration. Panel (c) and (d) show how $N_{ss}$ and $\bar{x}_{ss}$ change with $\widetilde{\theta}$ and $\sigma$, keeping the rest of the parameters constant. The black diamonds indicate the levels of $\widetilde{\theta}$ and $\sigma$ in our baseline calibration.

**Optimal Subsidy:** Panel (a) of Figure 9 shows the optimal adoption path in the model with complementarities (blue line) relative to the path of adopters from the decentralized equilibrium (black line). During the first three years after the launch of the technology, the optimal level of adoption is similar to that of the equilibrium without subsidy. After that, the optimal path of adopters from the planning problem is higher than that of the decentralized equilibrium. In fact, by the beginning of 2021, it is equal to the total number of informed agents in the economy, more than 30 percentage points higher than the levels of adoptions observed in the data.

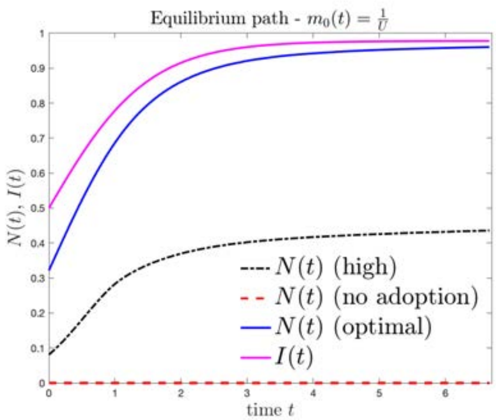
Figure 9: Planning Problem: Solution and Optimal Subsidy



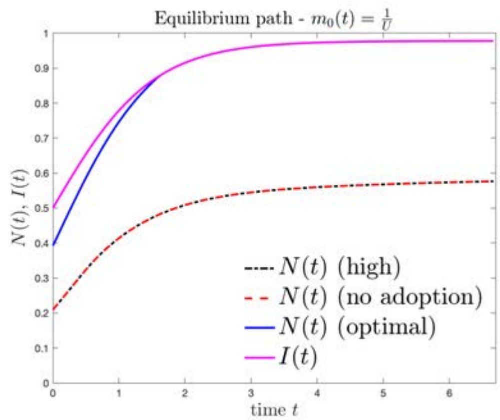

(a) Optimal Path of Adopters




(b) Optimal Subsidy




(c) Multiplicity: High Cost




(d) No Multiplicity: Low Cost

*Notes*: Panel (a) shows the share of informed agents, $I(t)$, the share of adopters in the decentralized model, $N(t)$, and the optimal levels of adoption, $N(t)$ (optimal), according to the solution of the planning problem. Panel (b) shows the path of the optimal subsidy $\theta_n Z(t)$ and the flow benefit of the average adopter, $Z(t)(\theta_0 + \theta_n N(t))$. Panel (c) shows the share of informed agents, $I(t)$, the share of adopters in the decentralized model, $N(t)$, and the optimal levels of adoption, $N(t)$ (optimal), according to the solution of the planning problem for a high adoption cost and 70% of the population informed 7 months after the launch of the technology. Panel (b) shows the same variables for a lower adoption cost and 70% of the population informed 7 months after the initial launch.

Panel (b) shows the path of the optimal subsidy.[34] As the share of adopters increases, so does the externality. Thus, the optimal subsidy, which is the same across agents, increases over time. To see why, notice the optimal subsidy in equation (31) can be written as

$$\theta_n Z(t) = \theta_n N \times \mathbb{E}(x|\text{adopted}),$$

where the expectation over $x$ is taken over the set of agents that have adopted the technology. The first term $\theta_n N$ captures the size of the adoption externality, i.e., the additional

[34] Figure 9 shows the subsidy $\theta_n Z(t)$ as a ratio of the net flow benefits (i.e., $(\theta_0 + \theta_n N(t))\mathbb{E}(x|adopted)$. In the invariant distribution, the subsidy-to-benefit ratio is approximately 0.5.

benefits for agents that adopt the technology. Thus, the subsidy increases as more agents adopt. Conversely, $\mathbb{E}(x|adopted)$ decreases as more agents adopt, since the marginal adopter has lower idiosyncratic benefits from adopting the technology. Intuitively, the planner internalizes that subsidizing agents with low $x$ also benefits the rest of the agents, even if the subsidy to incentivize these agents to adopt is large. The first component of the optimal subsidy dominates and eventually pushes the economy to universal adoption.[35] Importantly, *the planner is also constrained by the share of people who are informed*; otherwise, while the subsidy would still be increasing and the same across agents, there would be a "jump" in the subsidy's level as soon as the application is launched, as depicted in panel (b) of Figure 3.

**High Adoption Cost:** Our model can be used to study economies with higher adoption costs featuring multiple equilibria. We consider an economy with higher adoption cost $c$ and higher fraction of the population informed about the technology at launch. This example is motivated by a recent experience in El Salvador, where 70% of the population knew about a payment app introduced by the government (i.e., Chivo Wallet) 7 months after its initial launch.[36] Panel (c) shows the possible paths of adopters $N(t)$ for this economy. It shows that, when the adoption cost is larger, the decentralized equilibrium where nobody adopts the technology is not ruled out; for the same initial conditions there is an equilibrium with high adoption and one with no adoption. Panel (d) shows the same paths for a lower adoption cost. Our model allows for the study and quantification of policies that eliminate the no adoption equilibrium even if the optimal subsidy is not implemented. In this case, a large enough permanent subsidy can lower the adoption cost, solve the coordination failure, and send the decentralized economy to the high adoption equilibrium, i.e., from Panel (c) to (d).[37]

# 9 Conclusion

Understanding the adoption process of a technology and the transition from low to high adoption is challenging, especially in the presence of strategic complementarities. This paper develops a new dynamic model of technology adoption which allows us to model this transition. The model provides a framework to generate gradual adoption through a novel mechanism—waiting for others to adopt—and allows us to derive predictions that can be

[35] Figure H3 shows that only for lower levels of $\widetilde{\theta}$ the planner prescribes lower adoption levels.

[36] The app allows users to digitally trade both bitcoin and dollars.

[37] The Salvadorean government did in fact implement a similar subsidy. As an incentive to adopt, citizens who downloaded Chivo Wallet received a $30 bitcoin bonus from the government. Our model suggests that the subsidy was not large enough to rule-out the no adoption equilibrium given the low levels of adoption of Chivo Wallet reported by Alvarez, Argente and Van Patten (2022a).

tested empirically. We solve for the social planner's problem. The planner in our setup controls the entire distribution of adopters across time. The presence of strategic complementarities enrich the problem and allow us to link our results to the "big push" literature, as they imply that small subsidies can lead to large changes in adoption given the multiplicity of equilibria. In our framework, the optimal subsidy increases over time but it is flat across people, thus, easily implementable. The methodology we develop can be useful for a wide set of multidimensional dynamic problems, and can be applied to studying any technology that features strategic complementarities, learning, or both.

Our application analyzes new electronic methods of payment, which are particularly relevant today and are undertaking a digital transformation. This revolution has been echoed by a growing interest from monetary authorities to promote and develop digital payment platforms, both in developed and developing countries. Using individual- and transaction-level data on SINPE, a national electronic payment system adopted by a large fraction of the adult population in Costa Rica, along with extensive data on the networks of each user, we document that strategic complementarities play an important role in the adoption of this technology. SINPE also provides a rich environment to calibrate the model, which allows us to estimate the optimal time-varying adoption subsidy and the degree of selection into adoption across time. These results have implications for the launch and implementation of payment technologies with similar features such as CBDCs.

# References


Alvarez, F.E., Argente, D., Van Patten, D., 2022a. Are Cryptocurrencies Currencies? Bitcoin as Legal Tender in El Salvador. Technical Report 29968. National Bureau of Economic Research.

Alvarez, F.E., Lippi, F., Souganidis, T., 2022b. Price Setting with Strategic Complementarities as a Mean Field Game. Working Paper 30193. National Bureau of Economic Research.

Aron, J., 2018. Mobile Money and the Economy: A Review of the Evidence. World Bank Research Observer 33, 135–188.

Auer, R., Cornelli, G., Frost, J., 2020. Rise of the central bank digital currencies: drivers, approaches and technologies. BIS Working Papers 880. Board of Governors of the Federal Reserve System.

Bass, F.M., 1969. A New Product Growth for Model Consumer Durables. Management Science 15, 215–227.

Benhabib, J., Perla, J., Tonetti, C., 2021. Reconciling Models of Diffusion and Innovation: A Theory of the Productivity Distribution and Technology Frontier. Econometrica 89, 2261–2301. doi:https://doi.org/10.3982/ECTA15020.

Björkegren, D., 2018. The Adoption of Network Goods: Evidence from the Spread of Mobile Phones in Rwanda. The Review of Economic Studies 86, 1033–1060.

Buera, F.J., Hopenhayn, H., Shin, Y., Trachter, N., 2021. Big Push in Distorted Economies. Working Paper 28561. National Bureau of Economic Research.

Cabral, L.M., 1990. On the adoption of innovations with 'network' externalities. Mathematical Social Sciences 19, 299–308.

Calvo, G.A., 1983. Staggered prices in a utility-maximizing framework. Journal of Monetary Economics 12, 383–398.

Carapella, F., Flemming, J., 2020. Central Bank Digital Currency: A Literature Review. FEDS Notes. Board of Governors of the Federal Reserve System.

Comin, D., Hobijn, B., 2010. An Exploration of Technology Diffusion. The American Economic Review 100, 2031–2059. URL: http://www.jstor.org/stable/41038754.

Crouzet, N., Gupta, A., Mezzanotti, F., 2023. Shocks and Technology Adoption: Evidence from Electronic Payment Systems.

Davis, S.J., Haltiwanger, J., 1992. Gross Job Creation, Gross Job Destruction, and Employment Reallocation. The Quarterly Journal of Economics 107, 819–863.

Davis, S.J., Von Wachter, T.M., 2011. Recessions and the Cost of Job Loss. Technical Report. National Bureau of Economic Research.

Economides, N., Jeziorski, P., 2017. Mobile Money in Tanzania. Marketing Science 36, 815–837.

Griliches, Z., 1957. Hybrid Corn: An Exploration in the Economics of Technological Change. Econometrica 25, 501–522. URL: http://www.jstor.org/stable/1905380.

ILO, 2002. Women and men in the informal economy: A statistical picture.

Mansfield, E., 1961. Technical Change and the Rate of Imitation. Econometrica 29, 741–766. URL: http://www.jstor.org/stable/1911817.

Matsuyama, K., 1991. Increasing Returns, Industrialization, and Indeterminacy of Equilibrium. The Quarterly Journal of Economics 106, 617–650.

Milgrom, P., Shannon, C., 1994. Monotone Comparative Statics. Econometrica 62, 157–180.

Parente, S.L., Prescott, E.C., 1994. Barriers to Technology Adoption and Development. Journal of Political Economy 102, 298–321.

Reinganum, J.F., 1981. On the Diffusion of New Technology: A Game Theoretic Approach. The Review of Economic Studies 48, 395–405.

Stokey, N.L., 2020. Technology diffusion. Review of Economic Dynamics .

Suri, T., 2017. Mobile Money. Annual Review of Economics 9, 497–520. doi:10.1146/annurev-economics-063016-103638.

Topkis, D.M., 1978. Minimizing a Submodular Function on a Lattice. Operations research 26, 305–321.