Strategic Complementarities in a Dynamic Model of Technology Adoption: P2P Digital Payments*

Fernando Alvarez
University of Chicago

David Argente
Pennsylvania State University

Francesco Lippi
LUISS and EIEF

Esteban Méndez
Central Bank of Costa Rica

Diana Van Patten
Yale University

May 2023

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.

JEL Classification Numbers: E4, E5, O1, O2
Keywords: Technology Adoption, Strategic Complementarities, Means of Payment

*We thank Isaac Baley, Andres Blanco, Diego Comin, Nicolas Crouzet, Joe Hazell, Rishabh Kirpalani, John Leahy, Claudio Michelacci, Ben Moll, Guillermo Noguera, 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, 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.

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, equilibrium multiplicity, and the local equilibrium stability. 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 number of agents who have adopted the technology, $N(t)$, and on an idiosyncratic persistent

[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.

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 exogenously 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.[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 viceversa.

We obtain several theoretical results. First, we establish the monotonicity of the optimal decision rules and of the aggregation to study the set of equilibrium paths. In particular, from the optimal adoption decisions of agents facing a path N, we show that the optimal threshold path $\bar{x}$ is decreasing as a function of the path N. This is due to the strategic complementarities. Likewise, from the aggregation of decisions characterized by thresholds $\bar{x}$, we show that the path N is decreasing on the path $\bar{x}$. From these considerations, one can show that, for the same initial conditions, the equilibrium set is a lattice, i.e., the equilibrium paths can be ordered in terms of their intensity of adoption. This means that when there is more than one equilibrium, their paths do not cross. Additionally, all equilibrium path N increase as a function of the strength of strategic complementarity. 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. 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

[3] We also analyze 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 slow 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.

[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

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.

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 slowly within the network. The last two facts motivate the importance of a 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 provide evidence in support of a causal relation between the share of agents who have adopted (N) and usage of the app relies on arguably exogenous variations in the network size due to mass layoffs.[7]

We put together the theory and the data in a quantitative analysis. To match the initial slow 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

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

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. The intensive margin, in particular, allows us to tease out strategic complementarities from other channels. We calibrate other parameters using key moments from the data, including the half-life of the share of adopters. 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, we 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 and their stability (this is related to e.g., Cabral 1990, Matsuyama 1991). As a result, we can consider equilibria with low adoption rates due to coordination failures, a feature that is very relevant in low income countries. 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 adoption of electronic payment by *retailers* increases following an aggregate shock. Their

[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.

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 slow 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 slow 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 σ, 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 ν 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 a steady state value 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 τ 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:

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

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 :

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

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 G. 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:

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

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 θ_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 θ_0 and/or θ_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 θ 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 σ, 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 θ_n.

An immediate corollary of this proposition is that $m_0(x) = 1/U$ is a steady state 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 θ_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 σ is intuitive since, for a given U, a large σ makes the process to revert to the mean faster. Finally, if θ_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 steady state version of the model. We look for an initial condition m_0 such that the distribution is invariant, so that $\bar{x}(t) = \bar{x}_{ss}$ and $N(t) = N_{ss}$, both 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 a steady state, $D_t = 0$. 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 steady state 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) and imposing the $\sigma^2 = 0$ and the steady state $m_t = 0$ condition 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 steady states. We have the following result:

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

A key insight of the proposition is that multiple interior steady states occur when the complementarities are strong relative to the intrinsic value of the technology, i.e., when θ_0 is small and θ_n is large. We concentrate on the steady state of the deterministic model for two reasons. First, for small σ they provide a good benchmark for the steady state 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 steady state.[10]

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

Next we analyze the steady state of the stochastic version of the model. We will show that convergence to this equilibrium must be gradual, i.e., that 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 obtain $D(x,t) = \tilde{D}(x)$ and $\bar{x}(t) = \bar{x}_{ss}$. Given $\bar{x}_{ss}$ we obtain and $m(x,t) = \tilde{m}(x)$, from

[10] Indeed, in Appendix H 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 $N(t) = N_{ss}$.

which we derive N_{ss}. Given N_{ss}, we find $\tilde{D}$ 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 C.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* steady state threshold as a function of a given N_{ss}. The monotonicity properties of the function $\tilde{D}$ on the parameters N_{ss},θ_n,c and θ_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 θ_0 and θ_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:

$$
\begin{aligned}
\mathcal{X}_{ss}^{-1}(\bar{x}_{ss}) &= \frac{1}{\theta_n}\left[\frac{\rho c}{\left(\bar{x}_{ss}+\bar{A}_1 e^{\eta\bar{x}_{ss}}+\bar{A}_2 e^{-\eta\bar{x}_{ss}}\right)-\frac{\left(1+\eta\left(\bar{A}_1e^{\eta\bar{x}_{ss}}-\bar{A}_2e^{-\eta\bar{x}_{ss}}\right)\right)\left(e^{\eta\bar{x}_{ss}}+e^{-\eta\bar{x}_{ss}}\right)}{\eta\left(e^{\eta\bar{x}}-e^{-\eta\bar{x}_{ss}}\right)}}-\theta_0\right] \text{ where}\\
\bar{A}_1 &\equiv \frac{1}{\eta}\frac{\left(1-e^{-\eta U}\right)}{\left(e^{-\eta U}-e^{\eta U}\right)},\ \bar{A}_2\equiv\frac{1}{\eta}\frac{\left(1-e^{\eta U}\right)}{\left(e^{-\eta U}-e^{\eta U}\right)} \text{ and } \eta\equiv\sqrt{2\rho/\sigma^2}.
\end{aligned} \tag{19}
$$

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

σ 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 C.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 γ, and hence strictly decreasing in σ. 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 steady state, is described by the pair $\{\bar{x}_{ss}, N_{ss}\}$, which solves

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

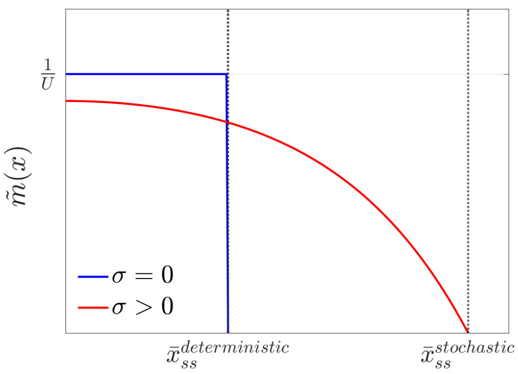
Next, we summarize the behavior of the steady states for small values of σ. We label the steady states with superscripts $\{H, L\}$ to hint at the associated High or Low level of adoption, so that $\bar{x}^H < \bar{x}^L$.

Proposition 6. Assume that $\nu > 0$ and that the parameters θ_0, θ_n, c and ρ are such that there are two interior steady states 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 steady states $\bar{x}_{ss}^H < \bar{x}_{ss}^L$. (ii) Each steady state is continuous with respect to σ at $\sigma = 0$. (iii) The sign of the comparative static differs across steady states, with

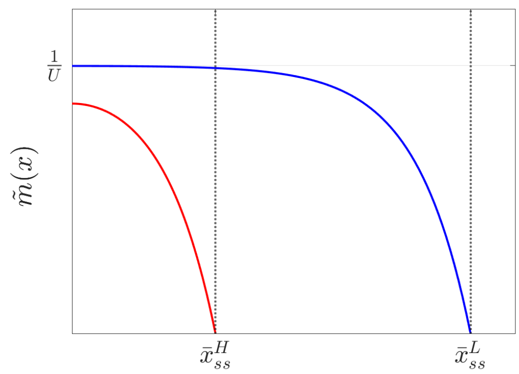
$$\frac{\partial \bar{x}_{ss}^H}{\partial c} > 0 > \frac{\partial \bar{x}_{ss}^L}{\partial c} \quad \text{and} \quad \frac{\partial \bar{x}_{ss}^L}{\partial \theta_0} > 0 > \frac{\partial \bar{x}_{ss}^H}{\partial \theta_0}$$

The proposition shows that the high adoption steady state 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 (θ_0). The comparative statics for the low adoption steady states are just the opposite.

Figure 1: Stochastic Steady State: Density of non-adopters: $m(x)$



(a) Deterministic and Stochastic SS



(b) High and Low Adoption SS

Panel (a) of Figure 1 compares the stationary density 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 steady state one, as agents adopt when $x(t) > \bar{x}(t)$ and it takes times for these $x's$ to crawl back below the steady state threshold. Panel (b) shows the invariant densities 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 steady state thresholds, although fewer of them in the low adoption 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 steady states 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 ω 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 ϵ 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 steady state N_{ss}. We will denote this Gateaux derivative by $\bar{y}$.

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

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

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 steady state 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 θ_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 ϵ 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 D.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 ω on the steady state $\tilde{m}$. In particular, we consider paths defined by $\bar{x}(t) = \bar{x}_{ss} + \epsilon\, \bar{y}(t)$ around the steady state x_{ss}, and $m_0(x) = \tilde{m}(x) + \epsilon\omega(x)$. We will denote this Gateaux derivative by n.

Proposition 8. Fix an interior steady state $\bar{x}_{ss}$, with its corresponding N_{ss}, and let $\tilde{m}$ be the corresponding steady state 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: ω 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 ω 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 ω to the initial distribution converges back to the steady state, 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, that implies a half-life $\mathbf{h}$ given by:

$$\mathbf{h} = \frac{\log(2)}{\nu + \frac{\sigma^2}{8}\left(\frac{\pi}{\bar{x}_{ss}}\right)^2}. \tag{26}$$

The proof's strategy is in Appendix D.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 ω.

Theorem 2. Fix an interior steady state $\bar{x}_{ss}$, with its corresponding N_{ss}, and let $\tilde{m}$ be the corresponding steady state 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 steady state. The linearized equilibrium must solve

$$n(t) = n_0(\omega)(t) + \Theta \int_0^T K(t,s)n(s)ds, \tag{27}$$

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

$$K(t,s) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} c_j e^{-\mu_i t - \psi_j s}\left[\frac{e^{(\mu_i+\psi_j)\min\{t,s\}} - 1}{\mu_i + \psi_j}\right] > 0. \tag{28}$$

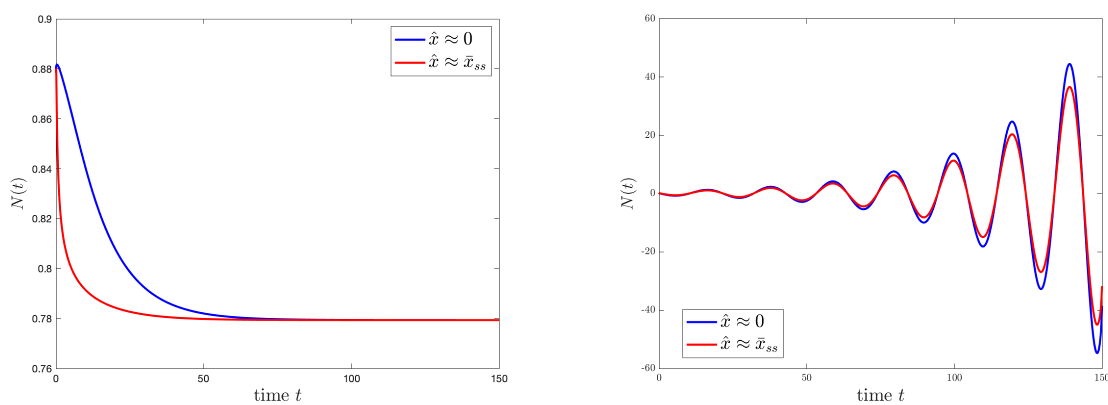
Moreover, $\mathrm{Lip}_K \equiv \sup_t \int |K(t,s)|ds \leq \left(\frac{\bar{x}_{ss}^2}{\sigma^2}\right)^2$. Furthermore, if $\Theta\,\mathrm{Lip}_K < 1$ there exists a unique bounded solution to equation (27) which is the limit of

$$n = \left[I + \Theta \mathcal{K} + \Theta^2 \mathcal{K}^2 + \dots \right] n_0(\omega) \quad \text{where} \quad \mathcal{K}(g)(t) \equiv \int_0^T K(t,s) g(s) ds, \tag{29}$$

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 θ_n because μ_j, ψ_j are a function of $\bar{x}_{ss}$, which is itself a function of θ_n. The coefficient Θ depends on θ_n directly and indirectly through $\bar{x}_{ss}$. Hence the equation (27) and its solution depend on which steady state we focus on. Second, if we discretize time so that $t \in \{\Delta_t(j-1) : j = 1, \dots, J\}$ 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 (27) becomes the linear equation $n = n_0 + \Theta \mathcal{K} n$. Third, the fact that $\Theta \mathcal{K} > 0$ implies that the terms $\Theta \mathcal{K} + \Theta^2 \mathcal{K}^2 + \dots$ in equation (29) give the amplification over and above n_0, due to the time-varying path of the barrier $\bar{x}$.

Figure 2: Perturbation of Stationary Equilbria



(a) High Adoption SS (b) Low Adoption SS

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 steady-state 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 the initial steady state. 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 (29) does not converge; i.e., that the system does not return to the initial steady state after being shocked.

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 E.1 characterizes the steady state 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:

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

subject to

$$N(t) = 1 - \int_0^{\bar{x}(t)} m(z,t)dz \quad \text{for all } t$$

$$m_t(x,t) = -\nu\,(m(x,t) - 1/U) + \frac{\sigma^2}{2} m_{xx}(x,t) \qquad \text{for } x \in (0, \bar{x}(t)) \text{ and all } t \geq 0 \quad \text{KFE}$$

$$m(x,t) = 0 \quad \text{for } x \in [\bar{x}(t), U] \text{ and all } t \geq 0 \qquad \text{Adoption}$$

$$m_x(0,t) = 0 \quad \text{for all } t \geq 0 \qquad \text{Reflecting}$$

$$m(x,0) = m_0(x) \qquad \text{initial condition}$$

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

[11] Appendix E.5 uses a linearized version of the problem to analyze dynamics around the steady state.

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 $\bar{x}(t)$, 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 E.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 λ, and the density of non-adopters m solve the p.d.e. for non-adopters:

$$\begin{aligned}
\rho\lambda(x,t) &= x\big(\theta_0 + \theta_n[1 - \int_0^{\bar{x}(t)} m(z,t)dz]\big) + \theta_n\big(\tfrac{U}{2} - \int_0^{\bar{x}(t)} m(z,t)z\,dz\big) \qquad (30)\\
&\quad + \tfrac{\sigma^2}{2}\lambda_{xx}(x,t) + \lambda_t(x,t) \text{ for } x \leq \bar{x}(t) \text{ and } t \geq 0\\
\lambda(x,t) &= c \text{ for } x \geq \bar{x}(t) \text{ and } t \geq 0\\
\lambda_x(\bar{x}(t),t) &= 0 \text{ for } t \geq 0 \qquad (31)\\
\lambda_x(0,t) &= 0 \text{ for } t \geq 0
\end{aligned}$$

and

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

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 λ in equation (30) 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

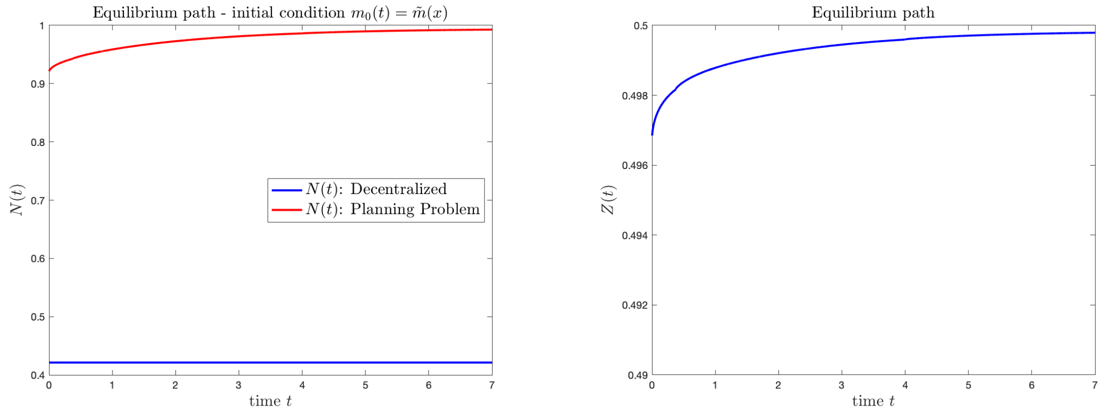
$$Z(t) \equiv \tfrac{U}{2} - \int_0^{\bar{x}(t)} m(x,t)\, x\, dx \quad \text{for all } t \geq 0 \tag{32}$$

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 (32). In particular, given $\bar{x}$ and m_0, using the KFE one solves for the path of m, and computing the integral in equation (32) 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 (31) 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 steady-state of the planning problem (see Appendix E.1 for a characterization of this steady state). 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 steady state.

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 I2.

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 (θ_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 I3, 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 I4.

[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 I1.

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 I2.

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 I1.

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 I7.

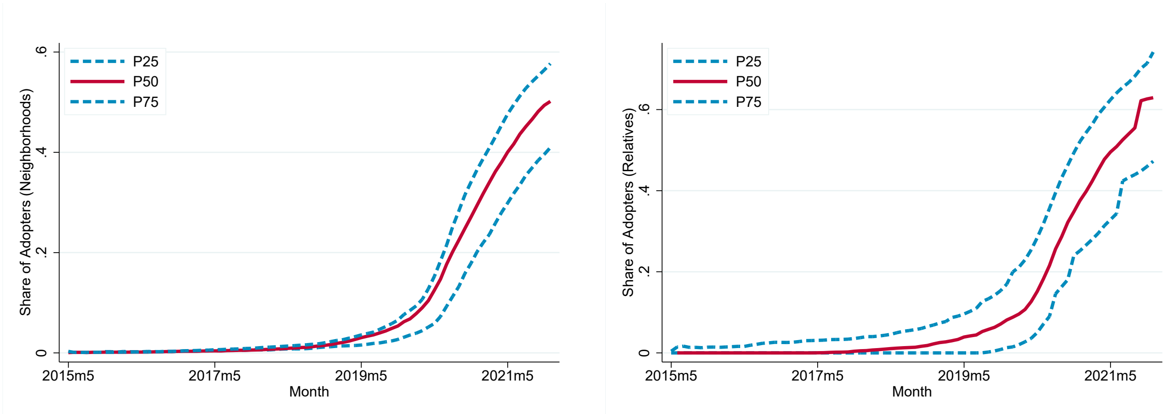
Fact 4: There is evidence of selection at entry. Through the lens of our model, early adopters—who started using the technology even when the network was small—should be more intense users (with higher x). Consistent with this notion, we document that early adopters have distinct characteristics as compared 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 I5.

[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 I1 did

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 I6.

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, ξ_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:

$$\ln \xi_t^* = \frac{1}{1+p}\ln\left[(\theta_0 + \theta_n N_t)\right] + \frac{1}{1+p}\ln x_t. \tag{33}$$

not have a major impact on overall trends.

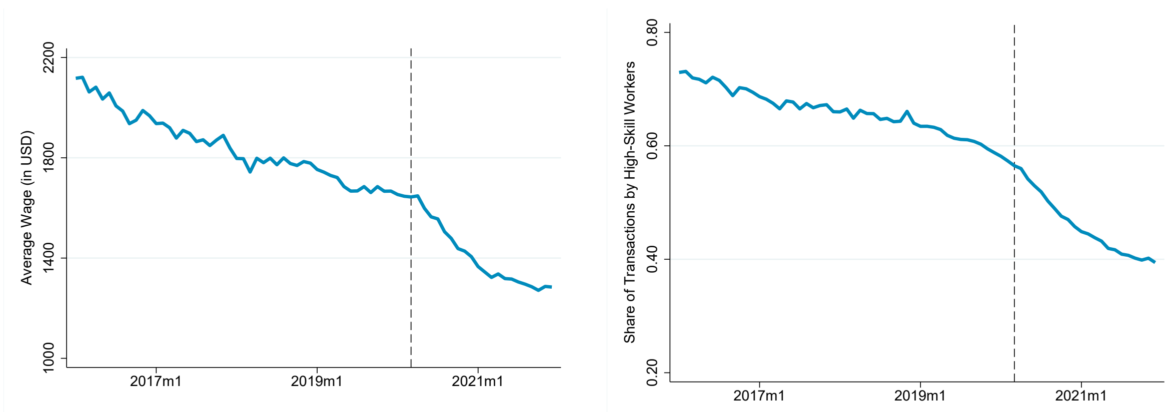
In this equation, note that if we were to remove the network×time variation, then $\ln \xi_{it}^{n}$ would proxy for $\ln x_t$, as through the lens of the model only the idiosyncratic variation would remain. Moreover, the model also predicts that users with a higher x would adopt the technology earlier. Thus, we can obtain a relation between intensity of usage (ξ_{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 ξ_{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, λ_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 I2.

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 (33). 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 (33), we obtain:

$$\ln \xi_t^* = \ln(\theta_0 + \theta_n N^*) + \frac{1}{1+p}\frac{\theta_n(N_t - N^*)}{\theta_0 + \theta_n N^*} + \frac{1}{1+p}\ln x_t. \tag{34}$$

Moreover, taking first differences, it follows that:

$$\Delta \ln \xi_t^* = \beta \Delta N_t + \nu_t, \tag{35}$$

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 β; 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 (35) 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 (35):

$$\Delta \ln \xi_{it}^n = \gamma + \beta \Delta N_t^n + \psi \Delta X_t^n + \lambda_t + \lambda_c + \epsilon_{it}^n, \tag{36}$$

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, ΔX_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, λ_t, which captures the fact that aggregate adoption is increasing over time along with cohort fixed-effects, λ_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 β 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 I3 and Table I4.[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 I5 shows the results of estimating equation (36) 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 I3 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 I6 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 I7 shows the IV results. The first stage is in the table's top panel, and we estimate equation (36), 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 I9. 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 I8.

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:

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

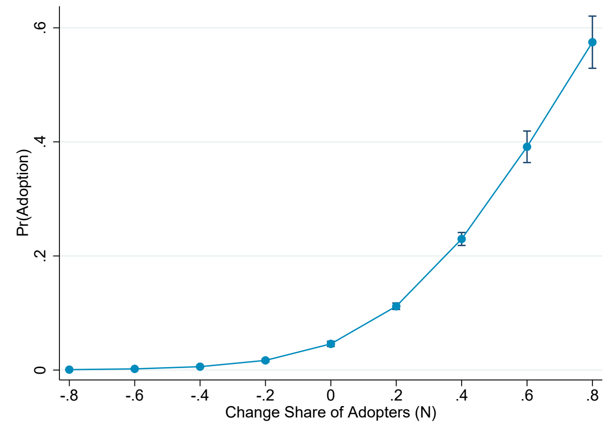
where $Adopt_i$ equals one if individual i adopted SINPE within 6 months after arriving to her new firm, and zero otherwise; $\Delta N_i^{coworkers}$ is the change between the share of coworkers who had adopted at the old and the new employer; $\Delta \ln wage_i$ corresponds with the change in the average 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 I.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 I.4.1 provides more details on each of these variables.

Panel (a) of Figure 6 shows the results estimating equation (37) 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 (36) from our theoretical model, which speaks to the technology's usage intensity. Thus, we consider:

$$\begin{aligned}\Delta \ln \tilde{\xi}_i =& \tilde{\alpha} + \tilde{\theta}\Delta N_i^{coworkers} + \tilde{\gamma}\Delta \ln wage_i + \hat{\psi}\Delta \ln size_i + \tilde{\lambda} date\, hired_i + \\ & \tilde{\omega}\Delta Covid_i + \tilde{\delta}\lambda_{ic} + \tilde{\nu}\ln \sum_{t=0}^{move} \tilde{\xi}_i + \tilde{\nu}\ln \sum_{t=0}^{move}\left(\tilde{\xi}_{t,\text{ new firm}} - \tilde{\xi}_{t,\text{ old firm}}\right) + \tilde{\epsilon}_i,\end{aligned} \tag{38}$$

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, λ_{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 (37).[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 (38) 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 I.4.1 provides more details on these variables and the choices made to conduct this exercise.
[27] Table I10 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 (38) 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]

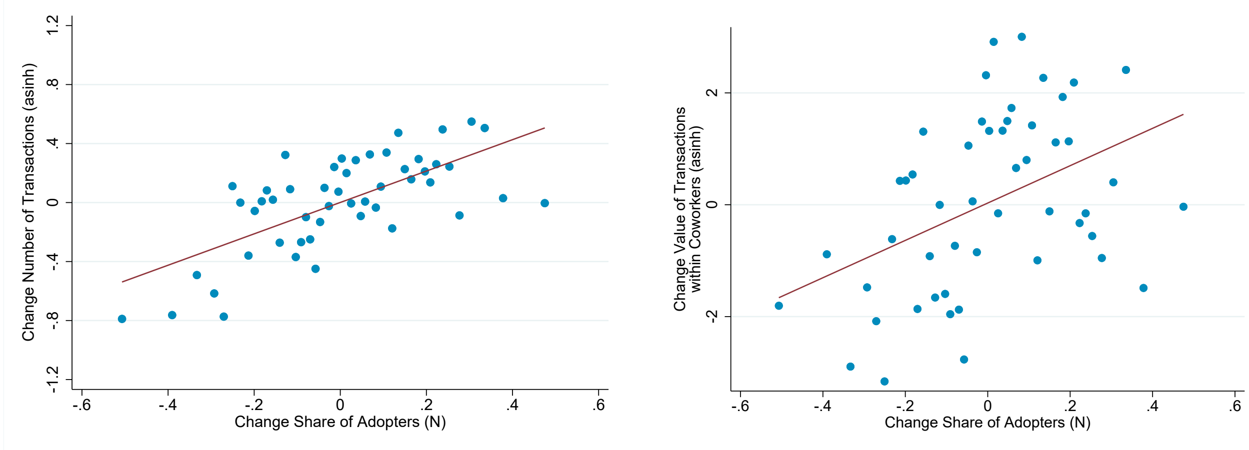
Column (4) controls for cohort, i.e., date of adoption, which aims to mitigate any effect of more experienced users behaving differently than beginners. Column (4) also controls for the total historical transactions made, which in a similar spirit as cohort, intends to mitigate any effect coming from 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 I8.

[29]The corresponding results using the value instead of the number of transactions are reported in Table I10.

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 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:

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

where $I(t)$ denotes the fraction of the population informed about the technology and β_0 indicates the number of meetings per unit of time between those informed and those uninformed (i.e., $1 - I(t)$), see Appendix F 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, β_0, and an initial condition for the informed population $I(0)$.

We calibrate ν 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 steady state. 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 ν 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 σ and θ_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 I9 imply $\widetilde{\theta} = \frac{2\beta}{1-2N^*\beta} \approx 2$ since $\beta = 0.76$ and $N^* = 0.17$. Table I7 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, σ is pinned down by the dispersion of the changes in transactions and θ_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.

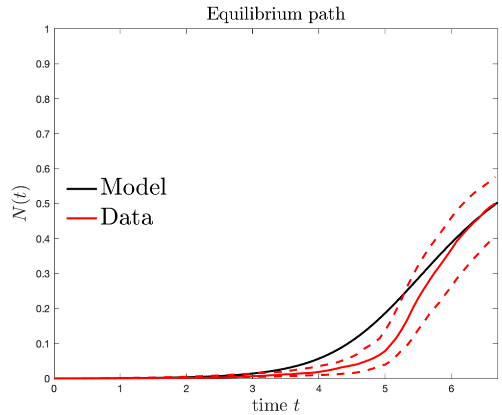
Results: Panel (a) of Figure 8 compares the path of adoption in the model and in the data. The solid red line indicates the diffusion of the technology in the median neighborhood and the dashed lines represent the 25^{th} and 75^{th} percentiles. The figure shows that both the speed and the level of 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 steady state, approximately 98% of the population knows about the application. Furthermore, the model predicts that in steady state 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 steady state level of adoption, we set all parameters to their estimated values and vary $\widetilde{\theta}$ and σ, while holding others constant. Panel (c) of Figure 8 shows how the steady state level of adoption changes with $\widetilde{\theta}$ and σ (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 σ is more subtle and results from two opposing forces. On the one hand, higher σ decreases N_{ss} since agents have a higher option value of waiting to adopt. On the other hand, higher σ 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 σ. 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 σ unambiguously increases $\bar{x}_{ss}$.

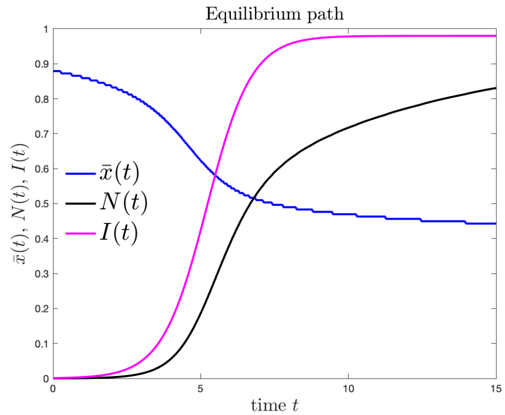
[32] A detailed description of the simulation and estimation procedure can be found in Section J.1.

[33] Section J.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.

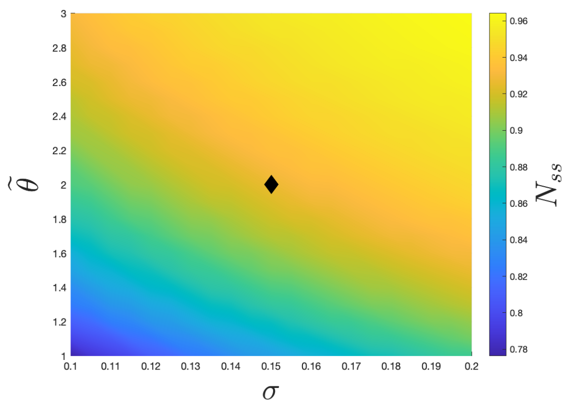
Figure 8: Path of Adopters (Short-Run and Long-Run)



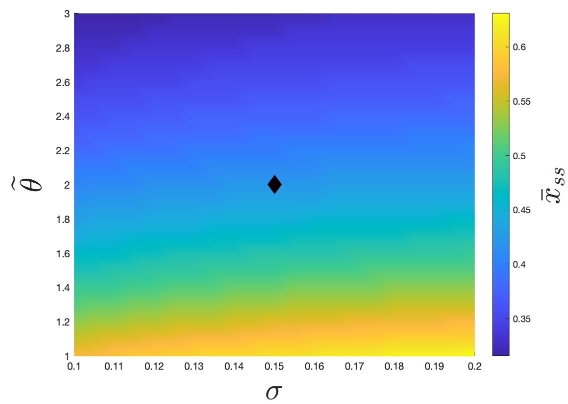
(a) Model vs Data



(b) Long-Run Path



(c) Comparative Statics: N_{ss}

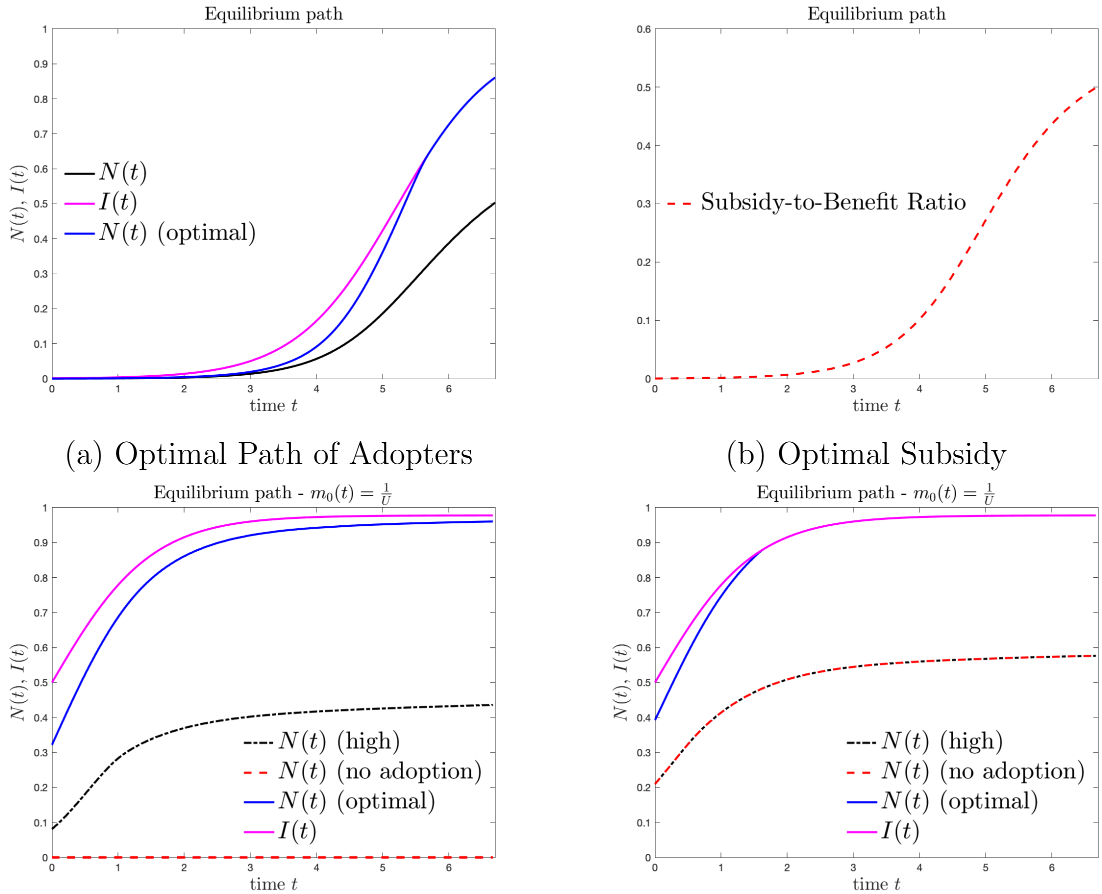


(d) Comparative Statics: $\bar{x}_{ss}$

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 σ, keeping the rest of the parameters constant. The black diamonds indicate the levels of $\widetilde{\theta}$ and σ 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 (32) 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 steady state, 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 J3 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.

Online Appendix for

Strategic Complementarities in a Dynamic Model of Techonology Adoption: P2P Digital Payments

May 4th, 2023

Fernando Alvarez
University of Chicago and NBER

David Argente
Penn State University and NBER

Francesco Lippi
LUISS and EIEF

Esteban Méndez
Central Bank of Costa

Diana Van Patten
Yale University and NBER

Table of Contents

Appendix A. Discretization and Computation of Equilibrium **1**

A.1 Finite Difference Computation of HJB for v, a Given N 1

A.2 Finite Difference Approximation of KFE for m Given $\bar{x}$ 3

A.3 Computing Equilibrium Set 5

Appendix B. Proofs for Propositions and Theorems **6**

Appendix C. Solution of the Steady State Problem **13**

C.1 Solution for $\tilde{a}(x)$ and $\tilde{v}(x)$ 13

C.2 Solution for $\tilde{m}(x)$ 15

Appendix D. Perturbation of the Stationary Equilibrium **16**

D.1 Linearization and Solution of the KB Equation 16

D.2 Linearization and Solution of the KF Equation 18

D.3 Equilibrium in the Perturbed MFG 19

Appendix E. Planning Problem **23**

E.1 Steady State: Planning Problem 23

E.2 Dynamics of N and Flow of Adoption Cost 24

E.3 Derivation of the PDE's for the Planner's Problem 25

E.4 Solution of the Steady State Planning Problem 31

E.5 Perturbation and Stability of Steady States 33

E.6 Perturbation of the Planning Problem 35

Appendix F. A "Pure" Learning Model **40**

F.1 Proofs for the Learning Model 46

Appendix G. HJB Equations for $a(x, t)$ and $v(x, t)$ **51**

Appendix H. The Dynamics of the Deterministic Model **52**

H.1 Proofs of the Deterministic Model 55

Appendix I. Empirical Appendix **57**

I.1 Descriptive Figures and Summary Statistics: SINPE 57
I.2 Evidence on Selection at Entry: Robustness 60
I.3 Evidence on Strategic Complementarities: Robustness 62
I.4 Mass Layoffs and Adoption Changes: Robustness and Details 66

Appendix J. Quantitative Exercises **69**
J.1 Calibration 69
J.2 Only Learning: $\widetilde{\theta} = 0$ 71
J.3 Comparative Statics 72

Appendix for Online Publication

A Discretization and Computation of Equilibrium

In this section, we describe an algorithm to compute the equilibrium. It is based on finding a fixed point of the finite difference approximation of the HBJ equation and the Kolmogorov forward equation.

We define the discretization of the model as follows:

Definition 3. A discretized version of the model is defined by positive integers I, J which determine the time and space step sizes: $\Delta_t = \frac{T}{J-1}$ and $\Delta_x = \frac{U}{I-1}$. Thus $t \in \mathbb{T} \equiv \{\Delta_t(j-1) : j = 1, \ldots, J\}$ and $x(t) \in \mathbb{X} \equiv \{\Delta_x(i-1) : i = 1, \ldots, I\}$. The reflecting BM is replaced by a process with: $x(t+\Delta_t) = x(t) \pm \Delta_x$ each with probability $q = \frac{1}{2}\frac{\sigma^2 \Delta_t}{(\Delta_x)^2}/(1-\nu\Delta_t)$, and $x(t+\Delta_t) = x(t)$ with probability $1-2q$ for $0 < x(t) < U$. If $x(t) = 0$ or $x(t) = U$, then $x(t+\Delta_t) = x(t)$, with prob. $1-q$, and $x(t+\Delta_t) = \Delta_x$, or $x(t+\Delta_t) = U - \Delta_x$ with probability q. Agents die with probability $\nu\Delta_t$, and use a discount factor $(1-\Delta_t r)$. The period flow of those that adopted the technology is $[\theta_0 + \theta_n N(t)]\, x(t)\Delta_t$. Agents that die are replaced by other whose x is drawn from a uniform discrete distribution with probabilities Δ_x/U for each x. For any $0 < \Delta_t < 1/(r+\nu)$, the value of J, and hence Δ_x must be chosen so that $0 < q \leq 1/2$. In this case the value functions v and a can be represented as a vector on $v \in \mathbb{R}^{I\times J}$, the distribution of non-adopters $m \in \mathbb{R}_+^{I\times J}$, threshold path $\bar{x} : \mathbb{T} \to \mathbb{X}$, and the path of the measure of adopters $N : \mathbb{T} \to [0,1]^J$. The initial condition is given by $m_0 \in \mathbb{R}_+^I$ and the terminal value by $v_T \in \mathbb{R}_+^I$.

Next we derive and describe the decision problem in discrete time using HBJ, and later derive and describe the discrete time version of the Kolmogorov forward equation.

A.1 Finite Difference Computation of HJB for v, a Given N

In this section we derive the finite difference approximation for $a(x,t)$ given the path $N = \{N_j\}_{j=1}^J$.

$$\rho a_{ij} = x_i\left(\theta_0 + \theta_n N_j\right) + \frac{\sigma^2}{2}\left[\frac{a_{i+1,j} - 2a_{i,j} + a_{i-1,j}}{(\Delta_x)^2}\right] + \frac{a_{i,j} - a_{i,j-1}}{\Delta_t}$$

for $i = 2, 3, \ldots, I-1$ and $j = 2, 3, \ldots, J-1$, which can be rearranged to give:

$$a_{i,j-1} = \Delta_t \, x_i \, (\theta_0 + \theta_n N_j) + \frac{\sigma^2 \Delta_t}{2(\Delta_x)^2} \left[a_{i+1,j} - 2a_{i,j} + a_{i-1,j} \right] + a_{i,j} - \rho \Delta_t a_{i,j}$$

Thus we define:

$$p = \frac{\sigma^2}{2} \frac{\Delta_t}{(\Delta_x)^2} \frac{1}{(1 - \rho \Delta_t)} \tag{39}$$

and write:

$$a_{i,j-1} = \Delta_t \, x_i \, (\theta_0 + \theta_n N_j) + (1 - \rho \Delta_t) \left[p a_{i-1,j} + (1-2p) a_{i,j} + p a_{i+1,j} \right] \tag{40}$$

for $i = 2, 3, \ldots, I-1$, and $j = 2, 1, J-1$, and

$$a_{1,j-1} = \Delta_t \, x_1 \, (\theta_0 + \theta_n N_j) + (1 - \rho \Delta_t) \left[(1-p) a_{1,j} + p a_{2,j} \right] \tag{41}$$

$$a_{I,j-1} = \Delta_t \, x_I \, (\theta_0 + \theta_n N_j) + (1 - \rho \Delta_t) \left[p a_{I-1,j} + (1-p) a_{I,j} \right] \tag{42}$$

for $j = 2, \ldots, J-1$ and at the terminal time we impose:

$$a_{i,J} = a_{i,T} \text{ for } i = 1, 2, \ldots, I \tag{43}$$

If we require that $p \in (0,1)$ and $1 - 2p \in (0,1)$ then

$$\frac{1}{\Delta_t} = \frac{J-1}{T} > \rho \text{ and}$$

$$\sigma \frac{\sqrt{\Delta_t}}{\sqrt{1 - \rho \Delta_t}} = \sigma \frac{\sqrt{T}}{\sqrt{J - 1 - \rho T}} < \Delta_x = \frac{U}{I-1}$$

We will use $a_T = \tilde{a}$, i.e., the steady state $\tilde{a}$ given N_{ss} as:

$$\tilde{a}_i = \Delta_t \, x_i \, (\theta_0 + \theta_n N_{ss}) + (1 - \rho \Delta_t) \left[p \tilde{a}_{i-1} + (1-2p) \tilde{a}_i + p \tilde{a}_{i+1} \right] \tag{44}$$

for $i = 2, 3, \ldots, I-1$ and

$$\tilde{a}_1 = \Delta_t \, x_1 \, (\theta_0 + \theta_n N_{ss}) + (1 - \rho \Delta_t) \left[(1-p) \tilde{a}_1 + p \tilde{a}_2 \right] \tag{45}$$

$$\tilde{a}_I = \Delta_t \, x_I \, (\theta_0 + \theta_n N_{ss}) + (1 - \rho \Delta_t) \left[p \tilde{a}_{I-1} + (1-p) \tilde{a}_I \right] \tag{46}$$

Now we set the equations for v using a. Following a similar derivation we get:

$$v_{i,j-1} = \max\left\{-c + a_{i,j}\,,\,(1-\rho\Delta_t)\left[pv_{i-1,j} + (1-2p)v_{i,j} + pv_{i+1,j}\right]\right\} \tag{47}$$

for $i = 2, 3, \ldots, I-1$, and $j = 2, 1, J-1$, and

$$v_{1,j-1} = \max\left\{-c + a_{1,j}\,,\,(1-\rho\Delta_t)\left[(1-p)v_{1,j} + pv_{2,j}\right]\right\} \tag{48}$$

$$v_{I,j-1} = \max\left\{-c + a_{I,j}\,,(1-\rho\Delta_t)\left[pv_{I-1,j} + (1-p)v_{I,j}\right]\right\} \tag{49}$$

for $j = 2, \ldots, J-1$ and at the terminal time we impose:

$$v_{i,J} = v_{i,T} \text{ for } i = 1, 2, \ldots, I$$

Given v and a we can compute $\bar{x}$, which correspond to an J dimensional array as:

$$\begin{aligned}
\bar{x}_j &= \min_{\{i=1,\ldots,I\}} \{x_i : v_{i,j} = -c + a_{i,j}\} \text{ for all } j = 1, 2, \ldots, J \\
\bar{i}_j &= \min_{\{i=1,\ldots,I\}} \{i : v_{i,j} = -c + a_{i,j}\} \text{ for all } j = 1, 2, \ldots, J \text{ so that} \\
\bar{x}_j &= x_{\bar{i}_j} \text{ for all } j = 1, 2, \ldots, J
\end{aligned}$$

We let $\mathbb{X}$ be the set:

$$\mathbb{X} = \left\{\{x_j\}_{j=1}^J : x_j = (i-1)\Delta_x \text{ each } i = 1, 2, \ldots I \text{ and } j = 1, 2, \ldots, J\right\}$$

We will use $v_T = \tilde{v}$, the steady state $\tilde{v}$ given $\tilde{a}$ as:

$$\tilde{v}_i = \max\left\{-c + \tilde{a}_i\,,\,(1-\rho\Delta_t)\left[p\tilde{v}_{i-1} + (1-2p)\tilde{v}_i + p\tilde{v}_{i+1}\right]\right\} \tag{50}$$

for $i = 2, 3, \ldots, I-1$ and

$$\tilde{v}_1 = \max\left\{-c + \tilde{a}_1\,,\,(1-\rho\Delta_t)\left[(1-p)\tilde{v}_1 + p\tilde{v}_2\right]\right\} \tag{51}$$

$$\tilde{v}_I = \max\left\{-c + \tilde{a}_I\,,\,(1-\rho\Delta_t)\left[p\tilde{v}_{I-1} + (1-p)\tilde{v}_I\right]\right\} \tag{52}$$

A.2 Finite Difference Approximation of KFE for m Given $\bar{x}$

In this section we derive the finite difference approximation for $m(x,t)$ given the path $\bar{x} = \{\bar{x}_j\}_{j=1}^J$. We let $\bar{i}_j$ the index for which $\bar{x}_j = x_{\bar{i}_j}$ for all j.

$$\frac{m_{i,j+1}-m_{i,j}}{\Delta_t} = \frac{\sigma^2}{2}\left[\frac{m_{i+1,j}-2m_{i,j}+m_{i-1,j}}{(\Delta_x)^2}\right] - \nu\left(m_{i,j}-\frac{1}{U}\right) \text{ for } i=2,3,\ldots,\bar{i}_j-1$$
$$m_{i,j+1} = 0 \text{ for } i=\bar{i}_j,\ldots,I$$

and $j = 1,2,\ldots,J$. We can rewrite the first equation as:

$$m_{i,j+1} = \frac{\sigma^2}{2}\frac{\Delta_t}{(\Delta_x)^2}\left[m_{i+1,j}-2m_{i,j}+m_{i-1,j}\right] - \nu\Delta_t\left(m_{i,j}-\frac{1}{U}\right) + m_{i,j} \text{ for } i=2,3,\ldots,\bar{i}_j-1$$
$$m_{i,j+1} = 0 \text{ for } i=\bar{i}_j,\ldots,I$$

Defining q as

$$q = \frac{\sigma^2}{2}\frac{\Delta_t}{(\Delta_x)^2}\frac{1}{(1-\nu\Delta_t)} \tag{53}$$

we can write it as:

$$m_{1,j+1} = (1-\nu\Delta_t)\left(qm_{2,j}+(1-q)m_{1,j}\right)+\nu\Delta_t\frac{1}{U} \tag{54}$$
$$m_{i,j+1} = (1-\nu\Delta_t)\left(qm_{i+1,j}+(1-2q)m_{i,j}+qm_{i-1,j}\right)+\nu\Delta_t\frac{1}{U} \text{ for } i=2,3,\ldots,\bar{i}_j-1 \tag{55}$$
$$m_{i,j+1} = 0 \text{ for } i=\bar{i}_j,\ldots,I \tag{56}$$

and $j = 1,2,\ldots,J$,

$$m_{i,1} = m_0(x_i) \text{ and } i=1,2,\ldots,I \tag{57}$$

Given m we can compute the corresponding N, i.e.:

$$N_j = 1 - \left(\sum_{i=1}^{I} m_{i,j}\Delta_x - m_{1,j}\Delta_x/2 - m_{\bar{i}_j-1,j}\Delta_x/2\right) \text{ for } j=1,2,\ldots,J \tag{58}$$

This gives $\mathcal{N}(\bar{x};m_0)$.

There is also the corresponding steady state version for $\tilde{m}$, given the index $\bar{i}^{ss}$:

$$\tilde{m}_1 = (1-\nu\Delta_t)\left(q\tilde{m}_2+(1-q)\tilde{m}_1\right)+\nu\Delta_t\frac{1}{U}$$
$$\tilde{m}_i = (1-\nu\Delta_t)\left(q\tilde{m}_{i+1}+(1-2q)\tilde{m}_i+q\tilde{m}_{i-1}\right)+\nu\Delta_t\frac{1}{U} \text{ for } i=2,3,\ldots,\bar{i}^{ss}$$
$$\tilde{m}_i = 0 \text{ for } i=\bar{i}^{ss},\ldots,I$$

and

$$N_{ss} = 1 - \left(\sum_{i=1}^{I} \tilde{m}_i \Delta_x - \tilde{m}_1 \Delta_x / 2 - \tilde{m}_{\bar{i}^{ss}-1} \Delta_x / 2 \right)$$

A.3 Computing Equilibrium Set

In this section we set up the fixed point given an initial condition m_0 and terminal value functions $v_T = \tilde{v}$, $a_T = \tilde{a}$ and $D_T = a_T - v_T$ for some steady state. Recall that $\mathcal{F} : [0,1]^J \to [0,1]^J$ is defined as in equation (6). Thus, successive paths for N are indexed by k and computed as

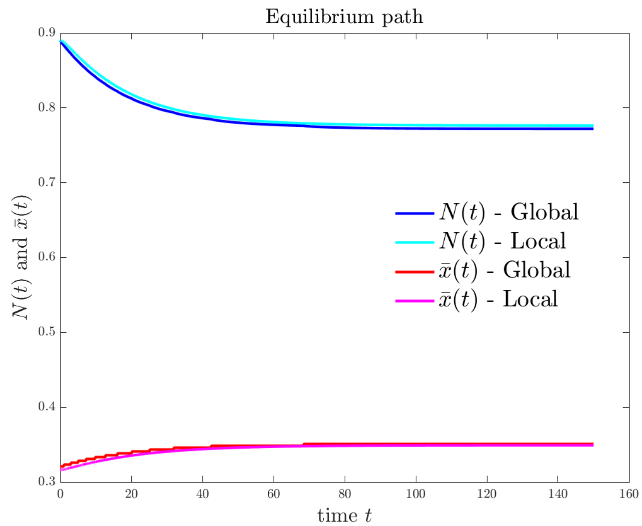
$$N^{k+1} = \mathcal{F}\left(N^k; m_0, D_T\right) \equiv \mathcal{N}\left(\mathcal{X}\left(N^k; D_T\right); m_0\right) \text{ for } k = 0, 1, 2, \ldots$$

for some initial condition N^0. To compute the equilibrium with the lowest path for N we start with the initial condition $N^0 = \{0, 0, \ldots, 0\}$. To compute the equilibrium with the highest path for N we start with the initial condition $N^0 = \{1, 1, \ldots, 1\}$. The convergence of N^k for large k is ensured by Tarski's theorem.

In Figure A1 we compare the computation that follows from discretizing time and state space with the one that comes from linearizing the model, i.e., our perturbation. Both computations start with the same initial conditions. For this figure we take as terminal value function the steady states values corresponding to the high adoption equilibrium, i.e., high value of N_{ss} and low value of $\bar{x}_{ss}$. The common initial condition is one where $m_0(x) = \tilde{m}(x)/2$. We make two remarks about the initial condition. First, it amounts to starting the economy with more agents with the technology than in the steady state (recall that $\tilde{m}$ is the steady state density of agents without the technology). Second, the shock (deviation from the steady state) is not a small one, hence the local perturbation might lose accuracy in principle.

The figure contains four lines. The two top lines display the computation of the path of N based on discretization (label as Global) with the one based on the perturbation (label as local). The two bottom lines display the computation of the path of $\bar{x}$ based on discretization (label as Global) with the one based on the perturbation (label as local). It is apparent that both methods gives very similar answer, i.e that the linearization is accurate for initial conditions far away from the steady states. The other feature apparent with these computations is that the steady state is stable even starting far away from the steady state.

Figure A1: Global vs Local Solutions



B Proofs

Proof. (of Proposition 1).

As a preliminary step we establish a correspondence and inequality between sample paths of a Brownian Motion with reflected barriers 0 and U but with different initial conditions. In particular, we can write $x(t, \alpha)$ for each sample path α:

$$x(t, \alpha) = x(0, \alpha) + \sigma \left[W(\omega, t) - W(\omega, 0)\right] + u(t, \alpha) - d(t, \alpha)$$

where ω are the sample path of the standard Brownian Motion denoted by W, where $u(\cdot, \alpha)$ and $d(\cdot, \alpha)$ are increasing processes in each sample path, where $u(s, \alpha)$ only increases when $x(s, \alpha) = 0$, and where $d(s, \alpha)$ only increases when $x(s, \alpha) = U$ for $s \in [0, t]$. Consider any sample path α for which $x(0, \alpha) = x_1$ with a corresponding sample path ω for the standard Brownian Motion W. Then there is a corresponding sample path α' where $x(0, \alpha') = x_2$, and with $\omega = \omega'$ for W, i.e., the two sample paths correspond to the same path of W. Thus, these two sample paths occur with the same probability. From the last observation it follows that we can represent the sample path α by the pair $\omega, x(0)$, where $x(0) = x(0, \alpha)$. Finally, if $x_1 < x_2$, comparing these two sample paths we obtain $x(t, \alpha') \geq x(t, \alpha)$, i.e., we can pair the sample paths that start with different initial conditions and that occur with the same probability, and obtain that the one that starts at a higher value is (weakly) higher for all future times, and strictly higher for t small enough.

Now we turn to the main result. We proceed by contradiction, assuming that while it is optimal to adopt at (x_1, t), it is not optimal to adopt for (x_2, t) with $x_2 > x_1$. Without loss of generality we assume that $t = 0$. Our hypothesis imply that for all stopping times with $\tau_1 > 0$ it is not convenient to wait if $x(0) = x_1$, and thus

$$-c + \mathbb{E}\left[\int_0^\infty e^{-\rho t} x(t)\left(\theta_0 + \theta_n N(t)\right) dt \,|\, x(0) = x_1\right] \geq \\ \mathbb{E}\left[-ce^{-\rho\tau_1} + \int_{\tau_1}^\infty e^{-\rho t} x(t)\left(\theta_0 + \theta_n N(t)\right) dt \,|\, x(0) = x_1\right]. \tag{59}$$

or equivalently that

$$-c + \mathbb{E}\left[\int_0^{\tau_1} e^{-\rho t} x(t)\left(\theta_0 + \theta_n N(t)\right) dt \,|\, x(0) = x_1\right] + c\,\mathbb{E}\left[e^{-\rho\tau_1} \,|\, x(0) = x_1\right] \geq 0\,.$$

Likewise, for $x(0) = x_2$ there exists a $\tau^* > 0$ for which it is optimal to wait:

$$-c + \mathbb{E}\left[\int_0^{\tau^*} e^{-\rho t} x(t)\left(\theta_0 + \theta_n N(t)\right) dt \,|\, x(0) = x_2\right] + c\,\mathbb{E}\left[e^{-\rho\tau^*} \,|\, x(0) = x_2\right] \leq 0\,.$$

We use the characterization for the sample paths described above, to construct a stopping time that only depends on the path ω as: $\tau_1(\omega, x_1) = \tau^*(\omega, x_2)$ for all ω. Using this equality, we immediately obtain $\mathbb{E}\left[e^{-\rho\tau_1} \,|\, x(0) = x_1\right] = \mathbb{E}\left[e^{-\rho\tau^*} \,|\, x(0) = x_2\right]$. Furthermore, using our characterization above for each path ω, we obtain:

$$\mathbb{E}\left[\int_0^{\tau_1} e^{-\rho t} x(t)\left(\theta_0 + \theta_n N(t)\right) dt \,|\, x(0) = x_1\right] < \mathbb{E}\left[\int_0^{\tau_1} e^{-\rho t} x(t)\left(\theta_0 + \theta_n N(t)\right) dt \,|\, x(0) = x_2\right] \\ = \mathbb{E}\left[\int_0^{\tau^*} e^{-\rho t} x(t)\left(\theta_0 + \theta_n N(t)\right) dt \,|\, x(0) = x_2\right]$$

Using this strict inequality we get a contradiction with equation (59), and hence we establish the desired result. □

Proof. (of Lemma 1).

The proof is readily obtained by using the definitions $\hat{a}(z,t) \equiv \theta_0 a(zU,t)$ and $\hat{v}(z,t) \equiv \theta_0 v(zU,t)$. It is straightforward to verify that these functions satisfy the partial differential equations for $\hat{a}(z)$ and $\hat{v}(z)$ for $z \in (0,1)$, including smooth pasting, value matching and boundary conditions. □

Proof. (of Proposition 2).

For this proof we set up the problem as a stopping time problem. We first prove a useful result in Lemma 4, showing that $\tau(N') \leq \tau(N)$ if $N' \geq N$. To convert the result on the monotonicity of the stopping times, into a result of the threshold $\bar{x}$, we note that the optimal decision rule is of the threshold type, as established in Proposition 1. We also show that exactly the same argument holds for the monotonicity with respect to θ. These results allow us to apply Topkis's (1978) theorem, which immediately establishes the proposition's result.

Next we set up the problem in terms of stopping times, and then state and prove Lemma 4.

□

Decision Problem as Stopping Times. Fix $x_0 \in [0, U]$ and $t_0 \in [0, T]$. Let $N \in C([t_0, T]) = \{N : [t_0, T] \to [0, 1]\}$ and τ denote a stopping time. Let Ω denote the sample paths that start at time t_0 with $x(t_0) = x_0$. A set $\mathbb{L}^{t_0,x_0} = \{\tau : \Omega \to [t_0, T]\}$ is a lattice since $\min\{\tau_1, \tau_2\}$ and $\max\{\tau_1, \tau_2\}$ are stopping times.

Let $\omega \in \Omega$ be a sample path that corresponds to a continuation of (x_0, t_0) with measure $\mu(\cdot|x_0, t_0)$. We denote by $x(\cdot, \omega) : [t_0, T] \to [0, U]$ the sample path of the process for x that starts at $x(t) = x_0$. Then the objective function can be written as

$$F(\tau, N; x_0, t_0) = \int f(\tau(\omega), x(\cdot, \omega), N)\mu(d\omega|x_0, t_0)$$

where

$$f(\tau, x(\cdot, \omega), N; x_0, t_0) = \left[\int_{\tau}^{T} e^{-\rho t} x(t, \omega)\left[\theta_0 + \theta_n N(t)\right] dt - e^{-\rho\tau} c\right]$$

where $F : \mathbb{L}^{t_0,x_0} \times C([t_0, T]) \to \mathbb{R}$. We have the following important lemma:

LEMMA 4. Let $\theta \equiv (\theta_0, \theta_n) \geq 0$ and fix (x_0, t_0). We establish three properties of $F(\tau, N; x_0, t_0)$: (i) it is submodular in τ; (ii) it has decreasing differences in (τ, N); (iii) it has decreasing differences in (τ, θ).

Proof. (of Lemma 4). Result (i): Submodularity in τ follows because F is additive across sample paths for all τ and τ'. We omit x_0, t_0 to simplify the notation. Fixing N we want to show:

$$F(\max\{\tau, \tau'\}, N) - F(\tau, N) \leq F(\tau', N) - F(\min\{\tau, \tau'\}, N)$$

which follows because for each sample path ω we have:

$$f(\max\{\tau, \tau'\}, N) - f(\tau, N) \leq f(\tau', N) - f(\min\{\tau, \tau'\}, N).$$

which holds since: $0 = f(\max\{\tau, \tau'\}, N) - f(\tau, N) - f(\tau', N) + f(\min\{\tau, \tau'\}, N)$.

Result (ii): We prove the submodularity of F, namely that given $\tau' > \tau$ and $N' > N$ we have

$$F(\tau', N') - F(\tau, N') \leq F(\tau', N) - F(\tau, N)$$

To this end consider $\tau'(\omega) \geq \tau(\omega)$ and compute:

$$F(\tau', N) - F(\tau, N) = \int \left(f(\tau', N) - f(\tau, N)\right) \mu(d\omega)$$

and for each ω

$$\begin{aligned} f(\tau', N, \omega) - f(\tau, N, \omega) &= \int_{\tau'}^{T} e^{-\rho t} \left[\theta_0 + \theta_n N(t)\right] x(t, \omega) dt - e^{-\rho\tau'} c \\ &- \left(\int_{\tau}^{T} e^{-\rho t} \left[\theta_0 + \theta_n N(t)\right] x(t, \omega) dt - e^{-\rho\tau} c\right) \\ &= -\int_{\tau}^{\tau'} e^{-\rho t} \left[\theta_0 + \theta_n N(t)\right] x(t, \omega) dt - e^{-\rho\tau'} c + e^{-\rho\tau} c. \end{aligned}$$

Thus, for all $N'(t) \geq N(t)$ and all t

$$\begin{aligned} &\left(f(\tau', N', \omega) - f(\tau, N', \omega)\right) - \left(f(\tau', N, \omega) - f(\tau, N, \omega)\right) \\ &= -\int_{\tau}^{\tau'} e^{-\rho t} \left[\theta_0 + \theta_n N'(t)\right] x(t, \omega) dt + \int_{\tau}^{\tau'} e^{-\rho t} \left[\theta_0 + \theta_n N(t)\right] x(t, \omega) dt \\ &= -\theta_n \int_{\tau}^{\tau'} e^{-\rho t} \left[N'(t) - N(t)\right] x(t, \omega) dt \leq 0 \end{aligned}$$

Thus

$$F(\tau', N') - F(\tau, N') - (F(\tau', N) - F(\tau, N)) = -\theta_n \int \left(\int_{\tau(\omega)}^{\tau'(\omega)} e^{-\rho t} \left[N'(t) - N(t)\right] x(t, \omega) dt\right) \mu(d\omega) \leq 0$$

Result (iii): Following the same steps followed in (ii) assuming $\theta' > \theta$ gives:

$$F(\tau', \theta') - F(\tau, \theta') - (F(\tau', \theta) - F(\tau, \theta)) = -\int \left(\int_{\tau(\omega)}^{\tau'(\omega)} e^{-\rho t} \left[(\theta_0' - \theta_0) + (\theta_n' - \theta_n) N(t)\right] x(t, \omega) dt\right) \mu(d\omega) \leq 0$$

□

Proof. (of Proposition 3) The fraction of agents that have not adopted at time t can be written as

$$M(t) \equiv \int_0^{\bar{x}(t)} m(z,t)dz = \int_0^U m_0(x)P(x,0,t)dx + \int_0^U \frac{\nu}{U} \int_0^t P(x,s,t)ds\, dx$$

where

$$P(x,s,t) = Pr\left[X(r) \leq \bar{x}(r), \text{ for all } r \in [s,t] \mid X(s) = x\right] e^{-\nu(t-s)} \tag{60}$$

where $X(\cdot)$ is a Brownian motion with reflecting barriers in $[0, U]$. Thus $P(x,s,t)$ is the fraction of agents that at time s have $X(s) = x$, survive until t, and also have had $X(r)$ below the threshold $\bar{x}(r)$ at all times $r \in [s,t]$. The first term in equation (60) is the fraction of those that have not adopted at in the initial distribution, and still have not adopted, and survive, at time t. The second term keeps tract of those cohort that have died at time s, and replaced by new agents, and themselves survive and not adopt up to time t.

Consider two paths $\bar{x}' \geq \bar{x}$ and the corresponding probabilities and measure of non-adopters $P'(x,s,t)$ and $M'(t)$ computed with $\bar{x}'$, and $P(x,s,t)$ and $M(t)$ computed with $\bar{x}$. The set of events $\{X(r) \leq \bar{x}(r), \text{ for all } r \in [s,t]\}$ is included in the set of events $\{X(r) \leq \bar{x}'(r), \text{ for all } r \in [s,t]\}$, since $\bar{x}(r) \leq \bar{x}'(r)$, and hence $P'(x,s,t) \geq P(x,s,t)$. Thus $M'(t) \geq M(t)$. Since $N'(t) = 1 - M'(t)$ and $N(t) = 1 - M(t)$, obtaining the desired result that $N'(t) \leq N(t)$.

The monotonicity with respect to m_0 follows immediately, since $\int_0^U m_0(x)P(x,0,t)dx$ is increasing in m_0 because $P(x,0,t)$ is non-negative.

□

Proof. (of Theorem 1) The proof uses Tarski's fixed point theorem for the function $\mathcal{F}$ as defined in equation (6). We restrict attention to the discrete time, discrete state version of the model so that we can we apply Tarski in a complete lattice.

We note that $\{N : \{0, \Delta_t, \ldots, T\} \to [0,1]\} = [0,1]^J$ where J is the integer that defines Δ_t. This set is a complete lattice. This function is monotone by virtue of Proposition 2 and Proposition 3. Then, Tarski's fixed point theorem implies that the set of fixed points is a lattice.

The comparative static result follows from the properties of the mapping $\mathcal{X}$ and $\mathcal{N}$ established in Proposition 2 and Proposition 3. □

Proof. (of Proposition 4) If an equilibrium without adoption exists, then $N(t) = N(0)e^{-\nu t}$, and hence if someone will adopt, it will adopt at time $t = 0$. Moreover, if someone will adopt it will be the one with $x = U$. Thus, we compute the value of $\underline{N}$ such that:

$$\begin{aligned}
c &= \mathbb{E}\left[\int_0^\infty e^{-\rho t} x(t)\left[\theta_0 + \theta_n N(t)\right] dt | x(0) = U\right] \\
&= \theta_0 \mathbb{E}\left[\int_0^\infty x(t) e^{-\rho t} dt | x(0) = U\right] + \theta_n N(0) \mathbb{E}\left[\int_0^\infty x(t) e^{-(\rho+\nu)t} dt | x(0) = U\right]
\end{aligned}$$

We note that $\tilde{a}(x;q) = \mathbb{E}\left[\int_0^\infty x(t) e^{-qt} dt | x(0) = x\right]$ solves the o.d.e. $q\tilde{a}(x) = 1 + \tilde{a}''(x)$ with boundary conditions $\tilde{a}'(0) = \tilde{a}'(U) = 0$. The solution of this o.d.e. is:

$$\begin{aligned}
\tilde{a}(x;q) &= \frac{1}{q}\left[x + \bar{A}_1 e^{\eta x} + \bar{A}_2 e^{-\eta x}\right] \\
\bar{A}_1 &\equiv \frac{1}{\eta} \frac{\left(1 - e^{-\eta U}\right)}{\left(e^{-\eta U} - e^{\eta U}\right)},\ \bar{A}_2 \equiv \frac{1}{\eta} \frac{\left(1 - e^{\eta U}\right)}{\left(e^{-\eta U} - e^{\eta U}\right)} \text{ and } \eta \equiv \sqrt{2q/\sigma^2}
\end{aligned}$$

Evaluating $\tilde{a}(x;q)$ at $x = U$ we get:

$$\tilde{a}(U;q) = \frac{1}{q}\left[U - \frac{\coth(\eta U)}{\eta} + \frac{\operatorname{csch}(\eta U)}{\eta}\right]$$

Using this in the expression for $\underline{N}$ we obtain the desired expression. $\square$

Proof. (of Proposition 5) First note that $x = U$ is a (non-interior) steady state if, in case nobody adopts ($N = 0$), then those with $x = U$ find it optimal not to adopt, which is equivalent to $\theta_0 U < \rho c$.

An interior steady state is the zero of $q(x) \equiv (\theta_0 + \theta_n)x - (\rho c + x^2\theta_n/U)$ which belongs to $(0, U)$. Note that $q(0) = -\rho c < 0$. In case (i), we have $q(U) = \theta_0 U - \rho c > 0$. Thus there is only one interior solution belonging to $(0, U)$. In case (ii), we have $q(U) = \theta_0 U - \rho c < 0$. In this case, since $q(x)$ is quadratic it can have zero, one, or two solutions. Note that fixing an x we have three properties: (1) $\partial q(x)/\partial \theta_n = x(1 - x/U) > 0$ if $x \in (0, U)$, (2) $\theta_n = 0$ then, $q(x) = x\theta_0 - \rho c = U(\theta_0 x/U - \rho c/U) < U(\theta_0 - \rho c/U) < 0$, where the last inequality holds in case (ii), and (3) that for large enough θ_n then $q(x) = \theta_0 x - \rho c + x\theta_n(1 - x/U) > 0$ for $x \in (0, U)$. Hence, we can find a θ_n^* such that for $\theta_n \in [0, \theta_n^*)$ there is no interior root, for $\theta_n = \theta_n^*$ there is exactly one interior root, and for $\theta_n > \theta_n^*$ there are two interior roots. $\square$

Proof. (of Lemma 2) The monotonicity of $\mathcal{X}_{ss}$ with respect to the parameters $\bar{\theta}_{ss} \equiv (\theta_0 + \theta_n N)/\rho$ is established in Appendix C.1. It is obtained by solving the o.d.e. for the value functions, and using the boundary conditions. It is clear that the optimal threshold, fixing η, solves an implicit equation $\psi(\gamma \bar{x}_{ss}) = \eta c/\bar{\theta}_{ss}$, where the function ψ is derived in Appendix C.1. This function is strictly increasing, and satisfies $\psi(0) = 0$. Thus $\mathcal{X}_{ss}$ is strictly decreasing in

$\bar{\theta}_{ss}$ and strictly increasing in c. A first order approximation of ψ gives the expansion used in the lemma. $\square$

Proof. (of Lemma 3) That $\mathcal{N}_{ss}$ is decreasing in $\bar{x}$ follows immediately since $\tanh(z)$ is, for positive z, concave and has $\tanh'(0) = 1$. Thus $\mathcal{N}_{ss}(\bar{x}) = \frac{1}{U}(-1 + \tanh(\bar{x}\gamma)) < 0$ if $\bar{x} > 0$.

That $\mathcal{N}_{ss}$ is strictly decreasing in γ follows from differentiating $\tanh(\bar{x}\gamma)/\gamma$ with respect to γ. This derivative is proportional to $-(\tanh(\bar{x}\gamma) - \bar{x}\gamma \text{sech}^2(\bar{x}\gamma)) = -(\tanh(\bar{x}\gamma) - \bar{x}\gamma \tanh'(\bar{x}\gamma)) < 0$, where we used that $\tanh(z)$ is strictly concave for $z > 0$. $\square$

Proof. (of Proposition 6). In the deterministic case, i.e., when $\sigma = 0$, there are at most two interior steady states (the case we focus on). To simplify the notation let $N^o(\bar{x}_{ss}) \equiv \mathcal{X}_{ss}^{-1}(\bar{x}_{ss})$ and $N^a(\bar{x}_{ss}) \equiv \mathcal{N}_{ss}(\bar{x}_{ss})$. In each of the steady states we write

$$N^a\left(\bar{x}^j(c)\right) = N^o\left(\bar{x}^j(c), c\right) \tag{61}$$

where $j = \{H, L\}$ (for high and low adoption, with $\bar{x}^H < \bar{x}^L$).

The functions N^a and N^o and their derivatives are continuous functions of $\bar{x}_{ss}, \sigma, c, \theta_0$. In each of the steady states the functions N^a and N^o have strictly different slopes. Some analysis shows that the functions N^a, N^o intersect twice, and the derivative of $N^a - N^o$ with respect to $\bar{x}_{ss}$ is positive when the curves intersect at $\bar{x}_{ss}^H$ and negative when the curves intersect at the $\bar{x}_{ss}^L$. We summarize this by writing $N_{\bar{x}}^a(\bar{x}_{ss}^H) - N_{\bar{x}}^o(\bar{x}_{ss}^H) > 0$ while the derivative is negative at $\bar{x}_{ss}^L$.

Note that c does not enter in N^a. Differentiating equation (61) with respect to c:

$$\left[N_{\bar{x}}^a\left(\bar{x}(c)\right) - N_{\bar{x}}^o\left(\bar{x}(c), c\right)\right] \frac{\partial \bar{x}(c)}{\partial c} = N_c^o\left(\bar{x}(c), c\right) > 0$$

and again using the properties of each steady state:

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

Following exactly the same steps we get:

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

$\square$

C Solution of the Steady State Problem

C.1 Solution for $\tilde{a}(x)$ and $\tilde{v}(x)$

The solution to $\tilde{a}$ is of the form:

$$\tilde{a}(x) = x\frac{\theta_0 + \theta_n N_{ss}}{\rho} + A_1 e^{\eta x} + A_2 e^{-\eta x}$$

for $\eta = \sqrt{2\rho/\sigma^2}$, and

$$0 = \frac{\theta_0 + \theta_n N_{ss}}{\rho} + \eta(A_1 - A_2) = \frac{\theta_0 + \theta_n N_{ss}}{\rho} + \eta(A_1 e^{\eta U} - A_2 e^{-\eta U})$$

Thus, given $\theta_0 + \theta_n N_{ss}$, the constants (A_1, A_2) are the solution of two linear equations. Moreover, the values of A_1, A_2 are proportional to $\bar{\theta}_{ss}$ given by

$$\bar{\theta}_{ss} \equiv \frac{\theta_0 + \theta_n N_{ss}}{\rho} = \eta(A_2 - A_1) = \eta(A_2 e^{-\eta U} - A_1 e^{\eta U})$$

Let $\bar{A}_i \equiv A_i/\bar{\theta}_{ss}$, we can write:

$$1 = \eta(\bar{A}_2 - \bar{A}_1) = \eta(\bar{A}_2 e^{-\eta U} - \bar{A}_1 e^{\eta U})$$

which has solution:

$$\bar{A}_1 = \frac{1}{\eta}\frac{\left(1 - e^{-\eta U}\right)}{\left(e^{-\eta U} - e^{\eta U}\right)} \quad , \quad \bar{A}_2 = \frac{1}{\eta}\frac{\left(1 - e^{\eta U}\right)}{\left(e^{-\eta U} - e^{\eta U}\right)}$$

The solution for $\tilde{v}$ for $x \in [0, \bar{x}_{ss}]$ is of the form

$$\tilde{v}(x) = B_1 e^{\eta x} + B_2 e^{-\eta x}$$

Given the solution for $\tilde{a}$, then $B_1, B_2, \bar{x}_{ss}$ solve:

$$\begin{aligned} 0 &= \eta(B_1 - B_2) \\ \tilde{a}_x(\bar{x}_{ss}) &= \eta(B_1 e^{\eta \bar{x}_{ss}} - B_2 e^{-\eta \bar{x}_{ss}}) \\ \tilde{a}(\bar{x}_{ss}) - c &= B_1 e^{\eta \bar{x}_{ss}} + B_2 e^{-\eta \bar{x}_{ss}} \end{aligned}$$

Thus, using the first equation $B_1 = B_2 = B$ and taking the ratio of these equations:

$$\frac{\tilde{a}(\bar{x}_{ss}) - c}{\tilde{a}_x(\bar{x}_{ss})} = \frac{1}{\eta} \frac{e^{\eta \bar{x}_{ss}} + e^{-\eta \bar{x}_{ss}}}{\left(e^{\eta \bar{x}_{ss}} - e^{-\eta \bar{x}_{ss}}\right)}$$

Replacing the expressions for $\tilde{a}(\bar{x}_{ss})$ and $\tilde{a}'(\bar{x}_{ss})$, we obtain:

$$\frac{\bar{x}_{ss} + \bar{A}_1 e^{\eta \bar{x}_{ss}} + \bar{A}_2 e^{-\eta \bar{x}_{ss}} - c/\bar{\theta}_{ss}}{1 + \eta\left(\bar{A}_1 e^{\eta \bar{x}_{ss}} - \bar{A}_2 e^{-\eta \bar{x}_{ss}}\right)} = \frac{1}{\eta} \frac{e^{\eta \bar{x}_{ss}} + e^{-\eta \bar{x}_{ss}}}{\left(e^{\eta \bar{x}_{ss}} - e^{-\eta \bar{x}_{ss}}\right)}$$

Note that this is one equation for $\bar{x}_{ss}$ as a function of $\bar{\theta}_{ss}$ (recall that $\bar{A}_1, \bar{A}_2$ are known constants). The last expression can be written as

$$\eta \bar{x}_{ss} + \eta \bar{A}_1 e^{\eta \bar{x}_{ss}} + \eta \bar{A}_2 e^{-\eta \bar{x}_{ss}} - \frac{e^{\eta \bar{x}_{ss}} + e^{-\eta \bar{x}_{ss}}}{\left(e^{\eta \bar{x}_{ss}} - e^{-\eta \bar{x}_{ss}}\right)} \left(1 + \eta\left(\bar{A}_1 e^{\eta \bar{x}_{ss}} - \bar{A}_2 e^{-\eta \bar{x}_{ss}}\right)\right) = \frac{\eta}{\bar{\theta}_{ss}} c$$

which gives equation (19) in the main text.

Letting $y \equiv \eta \bar{x}_{ss}$ and defining $\psi(y)$ we can write

$$\begin{aligned}\psi(y) &\equiv y + \eta\left(\bar{A}_1 e^{y} + \bar{A}_2 e^{-y}\right) - \frac{e^{y} + e^{-y}}{\left(e^{y} - e^{-y}\right)} \left(1 + \eta\left(\bar{A}_1 e^{y} - \bar{A}_2 e^{-y}\right)\right) \\ &= \frac{\eta}{\bar{\theta}_{ss}} c\end{aligned}$$

We can approximate the left hand side around $\bar{x}_{ss} = 0$, which corresponds to $c = 0$. Using that $\eta \bar{A}_2 = \eta \bar{A}_1 + 1$, we have the following properties.

1. $\psi(0) = 0$, $\psi(y) > 0$ if $y > 0$
2. $\psi'(y) = \frac{e^{2y}+1}{(e^{y}+1)^2}$ so $\psi'(0) = \frac{1}{2}, \psi'(\infty) = 1$, and $\psi''(y) > 0$,
3. $\psi(y) = \frac{y}{2} + \frac{y^3}{24} + o(y^4)$ and $\lim_{y \to \infty} \frac{\psi(y) - y}{y} = 0$

Now we use ψ to solve for $\bar{x}_{ss} = \chi(\eta, c/\bar{\theta}_{ss})$ i.e. $\frac{\eta c}{\bar{\theta}_{ss}} = \psi(\eta \chi(\eta, c/\bar{\theta}_{ss}))$. $\bar{x}_{ss}$ is the unique solution of $\frac{\psi(\eta \bar{x}_{ss})}{\eta} = \frac{c}{\bar{\theta}_{ss}}$, which always exists. For fixed $0 < \eta < \infty$ and small c using the first order approximation:

$$y = \eta \bar{x}_{ss} = 2 \frac{\eta c}{\bar{\theta}_{ss}} \text{ or } \bar{x}_{ss} = 2 \frac{c}{\bar{\theta}_{ss}}$$

since $\eta = \frac{\sqrt{2\rho}}{\sigma}$ the option value for a fixed $\bar{\theta}$ is given by:

$$\lim_{c \to 0} \frac{\chi(\eta, c/\bar{\theta}_{ss})}{\chi(\infty, c/\bar{\theta}_{ss})} = 2$$

For fixed $0 < \eta < \infty$ and small c, using the third order approximation $y^3 + 12y = \hat{\kappa} \equiv \frac{24\eta c}{\bar{\theta}_{ss}}$ or:

$$\begin{aligned}
\bar{x}_{ss} &= \frac{1}{\eta}\left(\frac{1}{2}\hat{\kappa} + \sqrt{\frac{1}{4}\hat{\kappa}^2 + \frac{12^3}{27}}\right)^{1/3} + \frac{1}{\eta}\left(\frac{1}{2}\hat{\kappa} - \sqrt{\frac{1}{4}\hat{\kappa}^2 + \frac{12^3}{27}}\right)^{1/3} \\
&= \frac{1}{\eta}\left(\frac{1}{2}\right)^{1/3}\left[\left(\hat{\kappa} + \sqrt{\hat{\kappa}^2 + 16}\right)^{1/3} + \left(\hat{\kappa} - \sqrt{\hat{\kappa}^2 + 16}\right)^{1/3}\right]
\end{aligned}$$

For the case when σ is small (i.e., η is large), let $S(y) \equiv y - \psi(y) + 1$ and recall that $\lim_{y\to\infty} S(y) = 0$. Then, using the definitions of y and $\psi(y)$, this implies

$$\lim_{\sigma \to 0} \frac{\sqrt{2\rho}}{\sigma}\left(\chi\left(\infty, \frac{c}{\bar{\theta}_{ss}}\right) - \frac{c}{\bar{\theta}_{ss}} - \frac{\sigma}{\sqrt{2\rho}}\right) = 0$$

Thus, for σ small we can use:

$$\bar{x}_{ss} = \frac{c}{\bar{\theta}_{ss}} + \frac{\sigma}{\sqrt{2\rho}} + o(\sigma)$$

Alternatively, note that $\bar{x}_{ss} - \frac{\sigma}{\sqrt{2\rho}}$ is the derivative of $\frac{\psi(\eta \bar{x}_{ss})}{\eta}$ with respect to σ evaluated at $\sigma = 0$.

C.2 Solution for $\tilde{m}(x)$

We can write the solution of the KFE as the sum the two homogeneous and the particular solution m_p, given $\bar{x}_{ss}$, i.e.

$$\tilde{m}(x) = C_1 e^{\gamma x} + C_2 e^{-\gamma x} + m_p(x)$$

where $\gamma = \sqrt{2\nu/\sigma^2}$. The solution is

$$\tilde{m}(x) = \frac{1}{U}\left[1 - \frac{(e^{\gamma x} + e^{-\gamma x})}{(e^{\gamma \bar{x}_{ss}} + e^{-\gamma \bar{x}_{ss}})}\right] \text{ for } x \in [0, \bar{x}_{ss}]$$

Finally, we want to compute:

$$1 - N_{ss} = \int_0^{\bar{x}_{ss}} \tilde{m}(x)dx = \int_0^{\bar{x}_{ss}} \frac{1}{U}\left[1 - \frac{(e^{\gamma x} + e^{-\gamma x})}{(e^{\gamma \bar{x}_{ss}} + e^{-\gamma \bar{x}_{ss}})}\right] dx$$

This gives another equation for $\bar{x}_{ss}$ as function of $\bar{\theta}$.

D Perturbation of the Stationary Equilibrium

We study the evolution of the MFG where the initial condition is given by a small perturbation ϵ of the steady state distribution:

$$m_0(x) = \tilde{m}(x) + \epsilon\,\omega(x)\ . \tag{62}$$

We consider an equilibrium with $\{\bar{x}(t,\epsilon), N(t,\epsilon), D(x,t,\epsilon), m(x,t,\epsilon)\}$. We will linearize this equilibrium with respect to ϵ and evaluate it at $\epsilon = 0$. For all $t \in [0,T]$, we denote these derivatives as follows:

$$\begin{aligned}
p(x,t) &\equiv \frac{\partial}{\partial \epsilon} m(x,t,\epsilon)\bigg|_{\epsilon=0} \\
d(x,t) &\equiv \frac{\partial}{\partial \epsilon} D(x,t,\epsilon)\bigg|_{\epsilon=0} \\
n(t) &\equiv \frac{\partial}{\partial \epsilon} N(t,\epsilon)\bigg|_{\epsilon=0} \\
\bar{y}(t) &\equiv \frac{\partial}{\partial \epsilon} \bar{x}(t,\epsilon)\bigg|_{\epsilon=0}
\end{aligned}$$

D.1 Linearization and Solution of the KB Equation

We differentiate $D(x,t,\epsilon)$ with respect to ϵ at each (x,t) to obtain $d(x,t)$ which solves the following p.d.e

$$\rho d(x,t) = x\theta_n n(t) + \frac{\sigma^2}{2} d_{xx}(x,t) + d_t(x,t) \tag{63}$$

for $x \in [0, \bar{x}_{ss}]$ and $t \in [0, T]$. The boundary conditions are obtained by differentiating the boundaries in equation (10) with respect to ϵ. This gives:

$$\begin{aligned} d(\bar{x}_{ss}, t) &= 0 \\ \tilde{D}_{xx}(\bar{x}_{ss})\bar{y}(t) + d_x(\bar{x}_{ss}, t) &= 0 \\ d_x(0, t) &= 0 \end{aligned} \tag{64}$$

for $t \in [0, T]$ and $d(x, T) = 0$ for $x \in [0, \bar{x}_{ss}]$. Note that equation (64) defines $\bar{y}(t)$ and that $\tilde{D}_{xx}(\bar{x}_{ss}) = \tilde{a}_{xx}(\bar{x}_{ss}) - \tilde{v}_{xx}(\bar{x}_{ss}) < 0$.

Taking the derivative of the solution for $d(x, t)$ in equation (63) with respect to x and combining it with equation (64) we find

$$\bar{y}(t) = \frac{\theta_n}{\tilde{D}_{xx}(\bar{x}_{ss})} \int_t^T G(\tau - t) n(\tau) d\tau \tag{65}$$

where $G(s) \equiv \sum_{j=0}^{\infty} c_j e^{-\psi_j s} \geq 0$ for $s \geq 0$, $\psi_j \equiv \rho + \frac{\sigma^2}{2}\left(\frac{\pi(\frac{1}{2}+j)}{\bar{x}_{ss}}\right)^2$, and $c_j \equiv 2\left(1 - \frac{\cos(\pi j)}{\pi(j+\frac{1}{2})}\right)$. An important property of this is that, since $G(s) \geq 0$ and $\tilde{D}_{xx}(\bar{x}_{ss}) < 0$, an increase in future adoption of the technology (i.e., future values of $n(\tau) > 0$ for $\tau > t$), then the threshold for adoption is smaller (i.e., more people will adopt today). Next we provide details of the solution of the p.d.e. for d. We have

Lemma 5. *The solution for the KBE equation for d, satisfying the p.d.e. in equation (63), and the boundary conditions in equation (64), is given by*

$$d(x, t) = \sum_{j=0}^{\infty} \varphi_j(x) \hat{d}_j(t) \qquad \text{for } x \in [0, \bar{x}_{ss}] \text{ and } t \in [0, T]$$

where for all $j = 1, 2, ...$ we have:

$$\begin{aligned} \varphi_j(x) &\equiv \sin\left(\left(\frac{1}{2} + j\right)\pi\left(1 - \frac{x}{\bar{x}_{ss}}\right)\right) && \text{for } x \in [0, \bar{x}_{ss}] \\ \hat{d}_j(t) &\equiv \int_t^T e^{-\psi_j(\tau - t)} \hat{z}_j(\tau) d\tau && \text{for } t \in [0, T] \\ \hat{z}_j(t) &\equiv \theta_n n(t) \frac{\langle \varphi_j, x \rangle}{\langle \varphi_j, \varphi_j \rangle} = \theta_n n(t) \frac{2\bar{x}_{ss}}{\left(\frac{1}{2} + j\right)\pi}\left(1 - \frac{\cos(\pi j)}{\pi(j + \frac{1}{2})}\right) && \text{for } t \in [0, T] \\ \text{where } \psi_j &\equiv \rho + \frac{\sigma^2}{2}\left(\frac{\pi(\frac{1}{2} + j)}{\bar{x}_{ss}}\right)^2 \quad \text{and} \quad \hat{d}_j(T) = 0 \end{aligned}$$

where $\langle \varphi_j, h \rangle \equiv \int_0^{\bar{x}_{ss}} h(x)\varphi_j(x) dx$. The proof can be done by verifying that the equation holds at the boundaries, and that for $t > 0$ the p.d.e in equation (63) holds in the interior since $\partial_{xx}\varphi_j(x) = -\left(\frac{\pi(\frac{1}{2}+j)}{\bar{x}_{ss}}\right)^2 \varphi_j(x)$, and $\partial_t \hat{d}_j(t) = \psi_j \hat{d}_j(t) - \hat{z}_j(t)$ for $t \in [0, T]$ and $j = 1, 2, ...,$ and since the $\{\varphi_j(x)\}$ form an orthogonal basis for functions. Note finally that the boundary holds at $t = 0$ for $x \in [0, \bar{x}_{ss}]$, and that the derivative of the solution for d, used to solve for $\bar{y}$ in equation (64), is

$$d_x(\bar{x}_{ss}, t) = -\theta_n \int_t^T \sum_{j=0}^{\infty} c_j e^{-\psi_j(s-t)} n(s) ds \quad \text{where} \quad c_j \equiv 2\left(1 - \frac{\cos(\pi j)}{\pi(j+\frac{1}{2})}\right) \quad .$$

D.2 Linearization and Solution of the KF Equation

We differentiate the KFE for $m(x, t, \epsilon)$ with respect to ϵ at each (x, t) to obtain:

$$p_t(x, t) = \frac{\sigma^2}{2} p_{xx}(x, t) - \nu p(x, t) \tag{66}$$

for $x \in [0, \bar{x}_{ss}]$ and $t \in [0, T]$.

Differentiating the boundary conditions $m(\bar{x}(t, \epsilon), t, \epsilon) = 0$ and $m_x(0, t, \epsilon) = 0$ with respect to ϵ we get

$$\begin{aligned} \tilde{m}_x(\bar{x}_{ss})\bar{y}(t) + p(\bar{x}_{ss}, t) &= 0 \\ p_x(0, t) &= 0 \end{aligned} \tag{67}$$

The initial condition comes from differentiating $m_0(x)$ with respect to ϵ

$$p(0, x) = \omega(x) \tag{68}$$

The solution for p satisfies the p.d.e given in equation (66), its boundary conditions in equation (67), and the initial condition in equation (68). We have

Lemma 6. The solution for the KFE equation for p, satisfying the p.d.e given in equation (66), the boundary conditions in equation (67), and the initial condition in equation (68), is given by

$$\begin{aligned} p(x, t) &= \sum_{j=0}^{\infty} \varphi_j(x) \hat{p}_j(t) + r(t) && \text{for } x \in [0, \bar{x}_{ss}] \text{ and } t \in [0, T] \\ r(t) &\equiv -\tilde{m}_x(\bar{x}_{ss})\, \bar{y}(t) && \text{for } t \in [0, T] \end{aligned}$$

where for all $j = 1, 2, ...$ we have:

$$\hat{p}_j(t) \equiv \hat{p}_j(0)e^{-\mu_j t} + \int_0^t e^{-\mu_j(t-\tau)}\hat{q}_j(\tau)d\tau \qquad \text{for } t \in [0,T]$$

$$\hat{q}_j(t) \equiv -(r'(t) + \nu r(t))\frac{\langle 1, \varphi_j\rangle}{\langle \varphi_j, \varphi_j\rangle} \qquad \text{for } t \in [0,T]$$

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

$$\text{where } \hat{p}_j(0) = \frac{\langle \varphi_j, \omega - r(0)\rangle}{\langle \varphi_j, \varphi_j\rangle} \quad \text{and} \quad \mu_j \equiv \nu + \frac{\sigma^2}{2}\left(\frac{\pi(\frac{1}{2}+j)}{\bar{x}_{ss}}\right)^2$$

where $\langle \varphi_j, h\rangle \equiv \int_0^{\bar{x}_{ss}} h(x)\varphi_j(x)dx$. The proof can be done by verifying that the equations hold at the boundaries, that for $t > 0$ the p.d.e holds in the interior since

$$\hat{p}_j'(t) = -\mu_j\hat{p}_j(t) + \hat{q}_j(t) \qquad \text{for } t \in [0,T] \text{ and } j = 1, 2, ...$$

and since $\{\varphi_j(x)\}$ form an orthogonal bases for functions, and finally that the boundary holds at $t = 0$ for $x \in [0, \bar{x}_{ss}]$, and it holds at $x = \bar{x}_{ss}$ for every $0 < t < T$

Given $p(x,t)$ we can compute $n(t)$ as:

$$\begin{aligned} n(t) &= -\int_0^{\bar{x}_{ss}} p(x,t)dx \\ &= n_0(t) + \frac{\tilde{m}_x(\bar{x}_{ss})\sigma^2}{\bar{x}_{ss}}\int_0^t J(t-\tau)\bar{y}(\tau)d\tau \end{aligned} \tag{69}$$

where $J(s) = \sum_{j=0}^{\infty} e^{-\mu_j s}$ with $\mu_j = \nu + \frac{1}{2}\sigma^2\left(\frac{\pi(\frac{1}{2}+j)}{\bar{x}_{ss}}\right)^2$ and $n_0(t) \equiv -\sum_{j=0}^{\infty}\frac{\bar{x}_{ss}}{\pi(\frac{1}{2}+j)}\frac{\langle\varphi_j,\omega\rangle}{\langle\varphi_j,\varphi_j\rangle}e^{-\mu_j t}$.

D.3 Equilibrium in the Perturbed MFG

Recall that from equation (65), $\bar{y}(t)$ is equal to

$$\bar{y}(t) = \frac{\theta_n}{\tilde{D}_{xx}(\bar{x}_{ss})}\int_t^T G(\tau - t)n(\tau)d\tau$$

where $G(s) \equiv \sum_{j=0}^{\infty} c_j e^{-\psi_j s}$ for $s \geq 0$. From equation (69) we also know that $n(t)$ is

$$n(t) = n_0(t) + \frac{\tilde{m}_x(\bar{x}_{ss})\sigma^2}{\bar{x}_{ss}} \int_0^t J(t-\tau)\bar{y}(\tau)d\tau$$

where $J(s) = \sum_{j=0}^{\infty} e^{-\mu_j s}$ and $n_0(t) \equiv -\sum_{j=0}^{\infty} \frac{\bar{x}_{ss}}{\pi(\frac{1}{2}+j)} \frac{\langle \varphi_j, \epsilon \rangle}{\langle \varphi_j, \varphi_j \rangle} e^{-\mu_j t}$. Combining equation (65) and equation (69) we get

$$\begin{aligned}
n(t) &= n_0(t) + \Theta(\bar{x}_{ss}) \int_0^t \int_\tau^T J(t-\tau)\bar{G}(s-\tau)n(s)dsd\tau \\
&= n_0(t) + \Theta(\bar{x}_{ss}) \int_0^T \int_0^{\min\{s,t\}} J(t-\tau)G(s-\tau)n(s)dsd\tau \\
&= n_0(t) + \Theta(\bar{x}_{ss}) \int_0^T K(t,s)n(s)ds
\end{aligned}$$

where $K(t,s) = \int_0^{\min\{s,t\}} J(t-\tau)\bar{G}(s-\tau)d\tau$ and $\Theta(\bar{x}_{ss}) \equiv \frac{\tilde{m}_x(\bar{x}_{ss})\sigma^2\theta_n}{\bar{x}_{ss}\tilde{D}_{xx}(\bar{x}_{ss})}$. Using the definitions of $J(s)$ and $G(s)$ we find

$$\begin{aligned}
K(t,s) &= \int_0^{\min\{s,t\}} J(t-\tau)G(s-\tau)d\tau \\
&= \int_0^{\min\{s,t\}} \left(\sum_{j=0}^{\infty} e^{-\mu_j(t-\tau)}\right)\left(\sum_{j=0}^{\infty} c_j e^{-\psi_j(s-\tau)}\right) d\tau \\
&= \sum_{j=0}^{\infty}\sum_{j=0}^{\infty} c_j e^{-\mu_i t - \psi_j s} \int_0^{\min\{s,t\}} e^{(\mu_i+\psi_j)\tau} d\tau \\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} c_j e^{-\mu_i t - \psi_j s} \left[\frac{e^{(\mu_i+\psi_j)\min\{t,s\}} - 1}{\mu_i + \psi_j}\right].
\end{aligned}$$

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

To calculate the Lipschitz bound $\mathrm{Lip}_K \equiv \sup_{t\in[0,T]} \int_0^T |K(t,s)|ds$, let

$$\kappa_{ij}(t) \equiv \int_0^T e^{-\mu_i t - \psi_j s}(e^{(\mu_i+\psi_j)\min\{t,s\}} - 1)$$

so that

$$\int_0^T K(t,s)ds = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} c_j \frac{\kappa_{ij}(t)}{\mu_i + \psi_j}.$$

Computing the integrals in $\kappa_{ij}(t)$ we get

$$\begin{aligned}
\kappa_{ij}(t) &= \int_0^t e^{-\mu_i t-\mu_i s}ds + \int_t^T e^{-\psi_j t-\psi_j s}ds - \int_0^T e^{-\mu_i t-\psi_j s}ds \\
&= \frac{e^{-\mu_i t}(e^{\mu_i t}-1)}{\mu_i} + \frac{e^{\psi_j t}(e^{-\psi_j T}-e^{-\psi_j t})}{-\psi_j} - \frac{e^{-\mu_i t}(e^{-\psi_j T}-1)}{-\psi_j} \\
&= \left(\frac{\psi_j+\mu_i}{\psi_j\mu_j}\right)(1-e^{-\mu_i t}) + e^{-\psi_j T}(e^{-\mu_i t}-e^{\psi_j t})
\end{aligned}$$

and as $T\to\infty$

$$\begin{aligned}
\kappa_{ij}(t) &= \left(\frac{\psi_j+\mu_i}{\psi_j\mu_j}\right)(1-e^{-\mu_i t}) \\
&\le \frac{\psi_j+\mu_i}{\psi_j\mu_i}.
\end{aligned}$$

Using that $\int_0^T |K(t,s)|ds \le \int_0^\infty |K(t,s)|ds$ we get

$$\begin{aligned}
\int_0^T K(t,s)ds &= \sum_{i=0}^\infty \sum_{j=0}^\infty c_j \frac{\kappa_{ij}(t)}{\mu_i+\psi_j} \\
&\le \sum_{i=0}^\infty \sum_{j=0}^\infty c_j \frac{1}{\mu_i\psi_j} \\
&= \left(\sum_{i=0}^\infty \frac{1}{\mu_i}\right)\left(\sum_{j=0}^\infty \frac{c_j}{\psi_j}\right).
\end{aligned}$$

We can use the definitions of μ_j, ψ_j, and c_j to further simplify this expression. First note that

$$\begin{aligned}
\sum_{i=0}^\infty \frac{1}{\mu_i} &= \sum_{i=0}^\infty \frac{1}{\nu+\frac{1}{2}\sigma^2\left(\frac{\pi(\frac{1}{2}+j)}{\bar{x}_{ss}}\right)^2} \\
&\le \frac{2\bar{x}_{ss}^2}{\sigma^2}\sum_{i=0}^\infty \frac{1}{\left(\pi(\frac{1}{2}+j)\right)^2} \\
&= \frac{\bar{x}_{ss}^2}{\sigma^2}
\end{aligned}$$

where we obtain the bound for $\nu = 0$. Notice also that

$$
\begin{aligned}
\sum_{j=0}^{\infty} \frac{c_j}{\psi_j} &= \sum_{j=0}^{\infty} \frac{2\left(1 - \frac{\cos(\pi j)}{\pi(j+\frac{1}{2})}\right)}{\rho + \frac{1}{2}\sigma^2 \left(\frac{\pi(\frac{1}{2}+j)}{\bar{x}_{ss}}\right)^2} \\
&\leq \frac{4\bar{x}_{ss}^2}{\sigma^2} \sum_{j=0}^{\infty} \frac{\left(1 - \frac{\cos(\pi j)}{\pi(j+\frac{1}{2})}\right)}{\left(\pi(\frac{1}{2}+j)\right)^2} \\
&= \frac{4\bar{x}_{ss}^2}{\sigma^2} \sum_{j=0}^{\infty} \left(\frac{1}{\left(\pi(\frac{1}{2}+j)\right)^2} - \frac{(-1)^j}{\left(\pi(\frac{1}{2}+j)\right)^3}\right) \\
&= \frac{4\bar{x}_{ss}^2}{\sigma^2} \sum_{j=0}^{\infty} \left(\frac{1}{2} - \frac{1}{4}\right) \\
&= \frac{\bar{x}_{ss}^2}{\sigma^2}
\end{aligned}
$$

where the bound is obtained for $\rho = 0$. Putting these together we find the Lipschitz bound

$$
\begin{aligned}
\mathrm{Lip}_K \equiv \sup_{t\in[0,T]} \int_0^T K(t,s)ds &\leq \left(\sum_{i=0}^{\infty} \frac{1}{\mu_i}\right) \left(\sum_{j=0}^{\infty} \frac{c_j}{\psi_j}\right) \\
&= \left(\frac{\bar{x}_{ss}^2}{\sigma^2}\right)^2.
\end{aligned}
$$

A sufficient condition for the existence and uniqueness of the equilibrium IRF, i.e., of the uniqueness and existence of a solution to equation (27) is that $|\Theta(\bar{x}_{ss})|\mathrm{Lip}_K < 1$. To establish a bound for $\Theta(\bar{x}_{ss})$, in terms of the fundamental model parameters, that ensures existence and uniqueness, we use the definition of $\Theta(\bar{x}_{ss})$ and the Lipschitz bound as follows:

$$
\begin{aligned}
\Theta(\bar{x}_{ss}) \left(\frac{\bar{x}_{ss}}{\sigma^2}\right)^2 &= \frac{\tilde{m}_x(\bar{x}_{ss})\sigma^2\theta_n}{\bar{x}_{ss}\tilde{D}_{xx}(\bar{x}_{ss})} \left(\frac{\bar{x}_{ss}^2}{\sigma^2}\right)^2 \\
&= \frac{\tilde{m}_x(\bar{x}_{ss})\theta_n\bar{x}_{ss}^3}{\tilde{D}_{xx}(\bar{x}_{ss})\sigma^2} \\
&= \frac{\theta_n(\gamma\bar{x}_{ss})^2}{2U} \frac{\tanh(\gamma\bar{x}_{ss})}{\left(\theta_0 + \theta_n\left(1 - \frac{\gamma\bar{x}_{ss}}{\gamma U} + \frac{\tanh(\gamma\bar{x}_{ss})}{\gamma U}\right)\right)\gamma\bar{x}_{ss} - \rho c\gamma}
\end{aligned}
$$

where we obtained $D_{xx}(\bar{x}_{ss})$ evaluating equation (9) at $\bar{x}_{ss}$ and using equation (20), and we calculate $\tilde{m}_x(\bar{x}_{ss})$ from $\tilde{m}(x) = \frac{1}{U}\left(1 - \frac{\cosh(\gamma x)}{\cosh(\gamma\bar{x}_{ss})}\right)$.

E Planning Problem

This section collects several results used to analyze the planning problem.

E.1 Steady State: Planning Problem

A steady state is given by two constants N_{ss} and $\bar{x}_{ss}$ that solve the time invariant version of the p.d.e. stated in Section 6. The p.d.e. for non-adopters in steady state is

$$\begin{aligned}
\rho\tilde{\lambda}(x) &= x\big(\theta_0 + \theta_n N_{ss}\big) + \theta_n Z_{ss} + \tfrac{\sigma^2}{2}\tilde{\lambda}_{xx}(x) \text{ if } x \le \bar{x}_{ss} && \text{KBE}\\
\tilde{\lambda}(\bar{x}_{ss}) &= c && \text{FOC}\\
\tilde{\lambda}_x(\bar{x}_{ss}) &= 0 && \text{Smooth Pasting}\\
\tilde{\lambda}_x(0) &= 0 && \text{Reflecting}\\
0 &= -\nu\tilde{m}(x) + \nu f(x) + \frac{\sigma^2}{2}\tilde{m}_{xx}(xx) \text{ if } x \le \bar{x}_{ss} && \text{KFE}\\
\tilde{m}(\bar{x}_{ss}) &= 0 \text{ and } \tilde{m}_x(0) = 0
\end{aligned}$$

and given $\tilde{m}$ and $\bar{x}_{ss}$, N_{ss} and Z_{ss} are defined as:

$$N_{ss} = 1 - \int_0^{\bar{x}_{ss}} \tilde{m}(x)dx$$
$$Z_{ss} = U/2 - \int_0^{\bar{x}_{ss}} x\tilde{m}(x)dx$$

Recall that $\tilde{\lambda}(\bar{x}_{ss})$ is the Lagrange multiplier of the law of motion of the density of agents that have not adopted in steady state. The details of the solution can be found in Appendix E.4. The following proposition summarizes the solution of stochastic steady state of the planning problem.

Proposition 11. Let $\tilde{\theta}_{ss} \equiv \frac{1}{\rho}(\theta_0 + \theta_n N_{ss})$ and $\eta \equiv \sqrt{2\rho/\sigma^2}$. For fixed $0 < \eta < \infty$ and small c, $\bar{x}_{ss} = 2\left(\frac{c}{\tilde{\theta}_{ss}} - \frac{\theta_n Z_{ss}}{\rho\tilde{\theta}_{ss}}\right)$. For the case when σ is small (i.e., η is large), $\bar{x}_{ss} = \frac{c}{\tilde{\theta}_{ss}} - \frac{\theta_n Z_{ss}}{\rho\tilde{\theta}_{ss}} + \frac{\sigma}{\sqrt{2\rho}}$

Proposition 11 indicates that the solution of the stochastic version of the planning problem also has the option value present in the equilibrium. This proposition can be used to show that the steady state level of adoption in the planning problem is higher than the adoption level of the the high-activity steady-state equilibrium.

E.2 Dynamics of N and Flow of Adoption Cost

Recall that

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

Taking the derivative with respect to time

$$\begin{aligned} N_t(t) &= -\frac{d}{dt}\int_0^{\bar{x}(t)} m(x,t)dx \\ &= \underbrace{-m(\bar{x}(t),t)}_{=0}\frac{d\bar{x}(t)}{dt} - \int_0^{\bar{x}(t)} m_t(x,t)dx \end{aligned}$$

where the first term is zero from the exit point of the distribution of non-adopters. Using the law of motion of m

$$\begin{aligned} N_t(t) &= -\int_0^{\bar{x}(t)} \left(-\nu m(x,t) + \nu f(x) + \frac{\sigma^2}{2} m_{xx}(x,t)\right) dx \\ &= \nu \int_0^{\bar{x}(t)} m(x,t) - \frac{\nu\bar{x}(t)}{U} - \frac{\sigma^2}{2}\int_0^{\bar{x}(t)} m_{xx}(x,t)dx \\ &= \nu\left(1 - N(t)\right) - \frac{\nu\bar{x}(t)}{U} - \frac{\sigma^2}{2}\left(\underbrace{m_x(\bar{x}(t),t)}_{<0} - \underbrace{m_x(0,t)}_{=0}\right) \end{aligned}$$

where the last term is zero from our assumption of reflecting barriers. Let the adoption cost per unit of time $A(t)$ be defined as

$$\begin{aligned} A(t) &\equiv c\left(N_t(t) + \nu N(t)\right) \\ &= c\left(\nu\left(1 - N(t)\right) - \frac{\nu\bar{x}(t)}{U} - \frac{\sigma^2}{2} m_x(\bar{x}(t),t) + \nu N(t)\right) \\ &= c\left(\nu\left(1 - \frac{\bar{x}(t)}{U}\right) - \frac{\sigma^2}{2} m_x(\bar{x}(t),t)\right) \end{aligned}$$

where the first term are the agents that are replaced with $x \geq \bar{x}(t)$. The second term are the agents that hit $\bar{x}(t)$ from below per unit of time so they pay c and adopt the technology.

E.3 Derivation of the PDE's for the Planner's Problem

To derive the problem in continuous time, we write the adoption problem in a discrete-time discrete state setup. We do so by using finite-difference approximation and then we consider the planning problem in that set-up. We obtain the first order conditions for a problem in finite dimensions. Lastly, we take the limit to develop the corresponding p.d.e's.

First we derive the finite difference approximation for a Brownian motion reflected between two barriers. The time step Δ so that times are between $t = 0, \Delta, 2\Delta, \ldots$. The space step is Δ_x so that $x \in \{x_1, x_2, \ldots, x_I\}$, where $x_1 = 0, x_J = U$ and $x_{i+1} - x_i = \Delta_x$. The p.d.e. inside the barriers is

$$m_t(x,t) = -\nu m(x,t) + \nu f(x) + \frac{\sigma^2}{2} m_{xx}(x,t)$$

Its finite difference approximation is:

$$\frac{m_{i,t+\Delta} - m_{i,t}}{\Delta} = -\nu m_{i,t} + \nu f_i + \frac{\sigma^2}{2} \frac{(m_{i+1,t} - 2m_{i,t} + m_{i-1,t})}{(\Delta_x)^2}$$

for $i = 2, \ldots, I-1$. We can write the finite difference approximation as:

$$\begin{aligned} m_{i,t+\Delta} = m_{i,t} &\left(1 - \nu\Delta - \sigma^2 \frac{\Delta}{(\Delta_x)^2}\right) + f_i \nu \Delta \\ &+ \frac{\sigma^2}{2} \frac{\Delta}{(\Delta_x)^2} m_{i+1,t} + \frac{\sigma^2}{2} \frac{\Delta}{(\Delta_x)^2} \Delta m_{i-1,t} \end{aligned}$$

For the finite approximation, we have that since the law of motion must be local, and mean preserving:

$$\begin{aligned} m_{1,t+\Delta} = m_{1,t} &\left(1 - \nu\Delta - \sigma^2 \frac{\Delta}{(\Delta_x)^2}\right) + f_1 \nu \Delta \\ &+ \frac{\sigma^2}{2} \frac{\Delta}{(\Delta_x)^2} m_{2,t} + \frac{\sigma^2}{2} \frac{\Delta}{(\Delta_x)^2} m_{1,t} \\ m_{I,t+\Delta} = m_{I,t} &\left(1 - \nu\Delta - \sigma^2 \frac{\Delta}{(\Delta_x)^2}\right) + f_I \nu \Delta \\ &+ \frac{\sigma^2}{2} \frac{\Delta}{(\Delta_x)^2} m_{I-1,t} + \frac{\sigma^2}{2} \frac{\Delta}{(\Delta_x)^2} m_{I,t} \end{aligned}$$

We can write the l.o.m. at the boundaries as:

$$m_{1,t+\Delta} = m_{1,t}\,(1-\nu\Delta) + f_1\nu\Delta + \frac{\sigma^2}{2}\frac{\Delta}{\Delta_x}\frac{(m_{2,t}-m_{1,t})}{\Delta_x}$$
$$m_{I,t+\Delta} = m_{I,t}\,(1-\nu\Delta) + f_I\nu\Delta + \frac{\sigma^2}{2}\frac{\Delta}{\Delta_x}\frac{(m_{I-1,t}-m_{I,t})}{\Delta_x}$$

At the reflecting boundaries $x = 0$ and $x = U$, the boundary conditions is $m_x(x,t) = 0$. Note that as $\Delta_x \to 0$ we require that

$$\frac{(m_{I-1,t}-m_{I,t})}{\Delta_x} = \frac{(m_{2,t}-m_{1,t})}{\Delta_x} \to 0$$

Now we get back to the planning problem. We will have two measures, $\{m_{i,t}\}$ and $\{g_{i,t}\}$. $m_{i,t}$ is the measures of those that have not adopted and $g_{i,t}$ the measure of those that have adopted. Let $\alpha_{it} \geq 0$ be the measure of adopting at t with $x = x_i$ at t. Thus at time t, the measure $\alpha_{i,t}$ is transferred from measure $m_{i,t}$ to measure $g_{i,t}$ Note that $m_{i,t} + g_{i,t} = \frac{1}{I}$ since the sum of the two is the invariant distribution. The initial condition are $g_{i,0} = 0\ \forall i$ and $m_{i,0} = \frac{1}{I}$ all non-adopters. The law of motion of the state is then:

$$\begin{aligned}
0 \leq m_{1,t+\Delta} &= m_{1,t}\left(1-\nu\Delta-\sigma^2\frac{\Delta}{(\Delta_x)^2}\right) + f\nu\Delta \\
&+ \frac{\sigma^2}{2}\frac{\Delta}{(\Delta_x)^2}m_{2,t} + \frac{\sigma^2}{2}\frac{\Delta}{(\Delta_x)^2}m_{1,t} - \alpha_{1,t} \\
0 \leq m_{i,t+\Delta} &= m_{i,t}\left(1-\nu\Delta-\sigma^2\frac{\Delta}{(\Delta_x)^2}\right) + f\nu\Delta \\
&+ \frac{\sigma^2}{2}\frac{\Delta}{(\Delta_x)^2}m_{i+1,t} + \frac{\sigma^2}{2}\frac{\Delta}{(\Delta_x)^2}\Delta m_{i-1,t} - \alpha_{i,t} \text{ for } i = 2,\dots,I-1 \\
0 \leq m_{I,t+\Delta} &= m_{I,t}\left(1-\nu\Delta-\sigma^2\frac{\Delta}{(\Delta_x)^2}\right) + f\nu\Delta \\
&+ \frac{\sigma^2}{2}\frac{\Delta}{(\Delta_x)^2}m_{I-1,t} + \frac{\sigma^2}{2}\frac{\Delta}{(\Delta_x)^2}m_{I,t} - \alpha_{I,t}
\end{aligned}$$

which can be written in vector notation as:

$$m_{t+1} = L\,m_t - \alpha_t \geq 0$$

where L is an $I \times I$ stochastic matrix which depends on I, ν, σ^2, Δ and Δ_x. We assume that

$\Delta\left(\nu + (\sigma/\Delta_x)^2\right) < 1$ so that all implied probabilities are positive.

$$\max_{\{\alpha_t, m_{t+\Delta}\}_{t=0}^{\infty}} \sum_{\{t=0,\Delta,2\Delta,\dots\}} \left(\frac{1}{1+\Delta r}\right)^t \left\{\mathcal{U}(m_t)\,\Delta\, - \sum_{i=1}^{I} \alpha_{it} c\right\}$$

where

$$\mathcal{U}(m_t) \equiv \sum_{i=1}^{I} \left(\frac{1}{I} - m_{it}\right) \left(\theta_0 + \theta_n \left[1 - \sum_{j=1}^{I} m_{j,t}\right]\right) x_i$$

subject to the law of motion:

$$m_{t+1} = L\, m_t - \alpha_t \text{ for all } t = 0, \Delta, 2\Delta, \dots$$

and subject to non-negativity:

$$m_{j,t+1} \geq 0 \text{ and } \alpha_{j,t} \geq 0 \text{ for all } j = 1, \dots, I, \text{ and for all } t = 0, \Delta, 2\Delta, \dots$$

Let $\left(\frac{1}{1+\Delta r}\right)^t \lambda_{it}$ be Lagrange multiplier of the law of motion for m_{it}. Let L_i be the i^{th} row vector of the matrix L . The Lagrangian $\mathcal{L}$ becomes:

$$\begin{aligned}\mathcal{L} = &\sum_{\{t=0,\Delta,\dots\}} \left(\frac{1}{1+\Delta r}\right)^t \left\{\mathcal{U}\left(m_t\right)\Delta - \sum_{i=1}^{I} \alpha_{it} c\right\} \\ &+ \sum_{\{t=0,\Delta,\dots\}} \left(\frac{1}{1+\Delta r}\right)^t \left\{\sum_{i=1}^{I} \lambda_{it}\left(m_{i,t+\Delta} - L_i \cdot m_t + \alpha_{it}\right)\right\}\end{aligned}$$

Derivative of Lagrangian with respect to α_{it}:

$$\frac{\partial \mathcal{L}}{\partial \alpha_{jt}} = \left(\frac{1}{1+\Delta r}\right)^t \left[\lambda_{j,t} - c\right]$$

Derivative of Lagrangian with respect to m_{jt} for $2 \leq j \leq I-1$:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial m_{j,t}} = &\left(\frac{1}{1+\Delta r}\right)^t \frac{\partial \mathcal{U}\left(m_t\right)}{\partial m_{j,t}} \Delta \\ &+ \left(\frac{1}{1+\Delta r}\right)^t \left[\lambda_{j,t-\Delta}(1+\Delta r) - \lambda_{j,t}\left(1 - \nu\Delta - \sigma^2 \frac{\Delta}{(\Delta_x)^2}\right)\right] \\ &- \left(\frac{1}{1+\Delta r}\right)^t \frac{\sigma^2}{2} \frac{\Delta}{(\Delta_x)^2} \left[\lambda_{j+1,t} + \lambda_{j-1,t}\right]\end{aligned}$$

where

$$\frac{\partial \mathcal{U}\left(m_t\right)}{\partial m_{jt}} = - x_j \left(\theta_0 + \theta_n \left(1 - \sum_{i=1}^{I} m_{i,t}\right)\right) - \theta_n \sum_{i=1}^{I} \left(\frac{1}{I} - m_{it}\right) x_i$$
$$= -x_j \left(\theta_0 + \theta_n N_t\right) - \theta_n \left(\frac{U}{2} - \sum_{i=1}^{I} m_{it} x_i\right)$$

We can write m_{jt} for $2 \leq j \leq I-1$:

$$\left(\frac{1}{1+\Delta r}\right)^{-t} \frac{\partial \mathcal{L}}{\partial m_{jt}} = \frac{\partial \mathcal{U}\left(m_t\right)}{\partial m_{jt}} \Delta + \lambda_{j,t-\Delta}(1+\Delta r)$$
$$- \lambda_{j,t} \left(1 - \nu\Delta - \sigma^2 \frac{\Delta}{(\Delta_x)^2}\right) - \frac{\sigma^2}{2} \frac{\Delta}{(\Delta_x)^2} \left[\lambda_{j+1,t} + \lambda_{j-1,t}\right]$$

and rearranging:

$$(1+\Delta r)\lambda_{j,t-\Delta} = \left(\frac{1}{1+\Delta r}\right)^{-t} \frac{\partial \mathcal{L}}{\partial m_{jt}} - \frac{\partial \mathcal{U}\left(m_t\right)}{\partial m_{jt}} \Delta$$
$$+ \lambda_{j,t} \left(1 - \nu\Delta - \sigma^2 \frac{\Delta}{(\Delta_x)^2}\right) + \frac{\sigma^2}{2} \frac{\Delta}{(\Delta_x)^2} \left[\lambda_{j+1,t} + \lambda_{j-1,t}\right]$$

dividing by Δ and further rearranging the expressions:

$$(r+\nu)\lambda_{j,t-\Delta} = \left(\frac{1}{1+\Delta r}\right)^{-t} \frac{1}{\Delta} \frac{\partial \mathcal{L}}{\partial m_{jt}} - \frac{\partial \mathcal{U}\left(m_t\right)}{\partial m_{jt}} - \nu\left(\lambda_{j,t} - \lambda_{j,t-\Delta}\right)$$
$$+ \left(\frac{\lambda_{j,t} - \lambda_{j,t-\Delta}}{\Delta}\right) + \frac{\sigma^2}{2} \left(\frac{\lambda_{j+1,t} - 2\lambda_{j,t} + \lambda_{j-1,t}}{(\Delta_x)^2}\right)$$

For the bottom boundary $j = 1$ we have:

$$\left(\frac{1}{1+\Delta r}\right)^{-t} \frac{\partial \mathcal{L}}{\partial m_{1t}} = \frac{\partial \mathcal{U}\left(m_t\right)}{\partial m_{1t}} \Delta + \lambda_{1,t-\Delta}(1+\Delta r)$$
$$- \lambda_{1,t} \left(1 - \nu\Delta - \sigma^2 \frac{\Delta}{(\Delta_x)^2}\right) - \frac{\sigma^2}{2} \frac{\Delta}{(\Delta_x)^2} \left[\lambda_{1,t} + \lambda_{2,t}\right]$$

$$(r+\nu)\lambda_{1,t-\Delta} = \left(\frac{1}{1+\Delta r}\right)^{-t} \frac{1}{\Delta} \frac{\partial \mathcal{L}}{\partial m_{1t}} - \frac{\partial \mathcal{U}\left(m_t\right)}{\partial m_{1t}} - \nu\left(\lambda_{1,t} - \lambda_{1,t-\Delta}\right)$$
$$+ \left(\frac{\lambda_{1,t} - \lambda_{1,t-\Delta}}{\Delta}\right) + \frac{\sigma^2}{2} \frac{1}{\Delta_x} \left(\frac{\lambda_{2,t} - \lambda_{1,t}}{\Delta_x}\right)$$

For the top boundary $j = I$:

$$
\begin{aligned}
(r+\nu)\lambda_{I,t-\Delta} = & \left(\frac{1}{1+\Delta r}\right)^{-t} \frac{1}{\Delta}\frac{\partial \mathcal{L}}{\partial m_{It}} - \frac{\partial \mathcal{U}(m_t)}{\partial m_{It}} - \nu\left(\lambda_{I,t} - \lambda_{I,t-\Delta}\right) \\
& + \left(\frac{\lambda_{I,t} - \lambda_{I,t-\Delta}}{\Delta}\right) + \frac{\sigma^2}{2}\frac{1}{\Delta_x}\left(\frac{\lambda_{I-1,t} - \lambda_{I,t}}{\Delta_x}\right)
\end{aligned}
$$

Thus the limit as $\Delta \downarrow 0$ and $\Delta_x \downarrow 0$ is that

$$
\lambda_x(0,t) = \lambda_x(U,t) = 0
$$

First order condition with respect to α_{it} for $t = 0, \Delta, \ldots$ and $j = 1, \ldots, I$::

$$
\begin{aligned}
\lambda_{j,t} - c \leq 0\,,\ \alpha_{jt} \geq 0 \text{ and} \\
\alpha_{j,t}\left[\lambda_{j,t} - c\right] = 0
\end{aligned}
$$

First order condition with respect to m_{jt} for $t = \Delta, 2\Delta, \ldots$ and $j = 1, \ldots, I$:

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial m_{jt}} \leq 0\,,\ m_{jt} \geq 0 \text{ and} \\
m_{jt}\frac{\partial \mathcal{L}}{\partial m_{jt}} = 0
\end{aligned}
$$

Note that as $\Delta \downarrow 0$ and $\Delta_x \downarrow 0$ and $x = x_j$ we have

$$
\frac{\partial \mathcal{U}(m_t)}{\partial m_{jt}} \to x(\theta_0 + \theta_n N(t)) + \theta_n\left(\frac{U}{2} - \int_0^U m(z,t)z\,dz\right)
$$

Consider a $x_j = x$ for $j = 2, \ldots, I-1$ or $0 < x < U$. Take the f.o.c. for $m_{j,t}$ derived above and assume that $\frac{\partial \mathcal{L}}{\partial m_{jt}} = 0$. Take the limit as $\Delta \downarrow 0$ and $\Delta_x \downarrow 0$:

$$
\begin{aligned}
(r+\nu)\lambda(x,t) = & \, x(\theta_0 + \theta_n N(t)) + \theta_n\left(\frac{U}{2} - \int_0^U m(z,t)z\,dz\right) \\
& + \lambda_t(x,t) + \frac{\sigma^2}{2}\lambda_{xx}(x,t)
\end{aligned}
$$

If instead $\frac{\partial \mathcal{L}}{\partial m_{jt}} \leq 0$, then

$$(r+\nu)\lambda(x,t) \leq x(\theta_0 + \theta_n N(t)) + \theta_n \left(\frac{U}{2} - \int_0^U m(z,t) z \, dz \right) \\ + \lambda_t(x,t) + \frac{\sigma^2}{2} \lambda_{xx}(x,t)$$

We derive **smooth pasting** here. Suppose that at t we have $\lambda_{i,t} = c$ for all $i \geq j$, i.e., for all $x \geq \bar{x}(t)$, or $\lambda(x,t) < c$ for $x < \bar{x}(t)$ and $\lambda(x,t) = c$ for $x \geq \bar{x}(t)$. Assume also $m_{j,t} > 0$ and $m_{j-1,t} > 0$, so that $\partial \mathcal{L}/\partial m = 0$ for both. Then we can write the f. o.c. as:

$$(r+\nu)c = -\frac{\partial \mathcal{U}(m_t)}{\partial m_{jt}} - \nu \left(c - \lambda_{j,t-\Delta} \right) \\ + \left(\frac{c - \lambda_{j,t-\Delta}}{\Delta} \right) + \frac{\sigma^2}{2} \frac{1}{\Delta_x} \left(\frac{c - 2c + \lambda_{j-1,t}}{\Delta_x} \right)$$

Taking the limit as $\Delta_x \downarrow 0$ we have: $\lambda_x(\bar{x}(t), t) = 0$.

In summary, a planner problem is given by $\{\bar{x}(t), \lambda(x,t), m(x,t)\}$ the path of optimal threshold so that adoption occurs for $x \geq \bar{x}(t)$, the Lagrange multiplier V, and the density of non-adopters m, respectively, such that the p.d.e. for the non-adopters is:

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

The p.d.e. for the non-adopters:

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

The conditions for $\bar{x}$ are:

- We look for $\bar{x}(\cdot)$ to be continuous $t \geq 0$.

Conditions for m:

- We look for $m(\cdot, t)$ to be continuous for all $x \in [0, U]$ and $t \geq 0$.

- We look for $m(\cdot, t)$ to be C^2 for all $x \in [0, \bar{x}(t)]$, and $t \geq 0$.
- We look for $m(x, \cdot)$ to be C^1 for all $x \in [0, \bar{x}(t)]$, and $t \geq 0$.
- The initial boundary condition for m is $m(x, 0) = 0$ for all $x \in [0, U]$

Conditions for λ:

- We look for $\lambda(\cdot, t)$ to be C^1 for all $x \in [0, U]$.
- We look for $\lambda(\cdot, t)$ to be C^2 for all $x \in [0, \bar{x}(t)]$, and $t \geq 0$.
- We look for $\lambda(x, \cdot)$ to be C^1 for all $x \in [0, \bar{x}(t)]$, and $t \geq 0$.
- The final boundary for λ is $\lambda(x, T) = 0$ for all $x \in [0, U]$ (T may be $+\infty$).

E.4 Solution of the Steady State Planning Problem

The solution for $\tilde{\lambda}$ of the form

$$\tilde{\lambda}(x) = x\frac{\theta_0 + \theta_n N_{ss}}{\rho} + \frac{\theta_n}{\rho} Z_{ss} x + C_1 e^{\eta x} + C_2 e^{-\eta x}$$

for $\eta = \sqrt{2\rho/\sigma^2}$, and

$$\begin{aligned} \frac{\theta_0 + \theta_n N_{ss}}{\rho} + \eta(C_1 e^{\eta \bar{x}_{ss}} - C_2 e^{-\eta \bar{x}_{ss}}) = 0 \\ \frac{\theta_0 + \theta_n N_{ss}}{\rho} + \eta(C_1 - C_2) = 0 \end{aligned}$$

Thus, given $\theta_0 + \theta_n N_{ss}$, and $\bar{x}_{ss}$, the constants (C_1, C_2) are the solution of two linear equations. Moreover, the values of A_1, A_2 are proportional to $\tilde{\theta}_{ss}$ given by

$$\tilde{\theta}_{ss} \equiv \frac{\theta_0 + \theta_n N_{ss}}{\rho} = \eta(C_2 - C_1) = \eta(C_2 e^{-\eta \bar{x}_{ss}} - C_1 e^{\eta \bar{x}_{ss}})$$

Let $\tilde{C}_i \equiv C_i / \tilde{\theta}_{ss}$. We can write:

$$1 = \eta(\tilde{C}_2 - \tilde{C}_1) = \eta(\tilde{C}_2 e^{-\eta \bar{x}_{ss}} - \tilde{C}_1 e^{\eta \bar{x}_{ss}})$$

which has solution:

$$\tilde{C}_1 = \frac{1}{\eta}\frac{(1-e^{-\eta\bar{x}_{ss}})}{(e^{-\eta\bar{x}_{ss}}-e^{\eta\bar{x}_{ss}})}$$
$$\tilde{C}_2 = \frac{1}{\eta}\frac{(1-e^{\eta\bar{x}_{ss}})}{(e^{-\eta\bar{x}_{ss}}-e^{\eta\bar{x}_{ss}})}$$

Using value matching we get:

$$\eta\bar{x}_{ss} + \frac{\eta\theta_n}{\rho\tilde{\theta}_{ss}}Z_{ss} + \eta(\tilde{C}_1 e^{\eta\bar{x}_{ss}} + \tilde{C}_2 e^{-\eta\bar{x}_{ss}}) = \frac{\eta}{\tilde{\theta}_{ss}}c$$

Letting $y \equiv \eta\bar{x}_{ss}$ we can write

$$\tilde{\psi}(y) \equiv y + \eta(\tilde{C}_1 e^{y} + \tilde{C}_2 e^{-y}) + \eta\frac{\theta_n}{\rho\tilde{\theta}_{ss}}Z_{ss}$$

Using $\eta\tilde{C}_2 = 1 + \eta\tilde{C}_1$ and the definition of $\tilde{C}_1$ we get

$$\tilde{\psi}(y) \equiv y + e^{-y} - \frac{(1-e^{-y})}{(e^{y}-e^{-y})}(e^{y}+e^{-y}) + \eta\frac{\theta_n}{\rho\tilde{\theta}_{ss}}Z_{ss}$$

We have the following properties:

1. $\tilde{\psi}(0) = \eta\frac{\theta_n}{\rho\tilde{\theta}_{ss}}Z_{ss}$
2. $\tilde{\psi}'(y) = \frac{e^{2y}+1}{(e^{y}+1)^2}$ so $\tilde{\psi}'(0) = \frac{1}{2}, \tilde{\psi}'(\infty) = 1$, and $\tilde{\psi}''(y) > 0$,
3. $\tilde{\psi}(y) = \frac{y}{2} + \frac{y^3}{24} + o(y^4) + \eta\frac{\theta_n}{\rho\tilde{\theta}_{ss}}Z_{ss}$ and $\lim_{y\to\infty} \frac{\tilde{\psi}(y)-y-\eta\frac{\theta_n}{\rho\tilde{\theta}_{ss}}Z_{ss}}{y} = 0$

For fixed $0 < \eta < \infty$ and small c using the first order approximation:

$$\bar{x}_{ss} = 2\left(\frac{c}{\tilde{\theta}_{ss}} - \frac{\theta_n}{\rho\tilde{\theta}_{ss}}Z_{ss}\right)$$

For the case when σ is small (i.e., η is large) we find:

$$\bar{x}_{ss} = \frac{c}{\tilde{\theta}_{ss}} + \frac{\sigma}{\sqrt{2\rho}} - \frac{\theta_n}{\rho\tilde{\theta}_{ss}}Z_{ss}$$

Defining $\gamma = \sqrt{2\nu/\sigma^2}$, for the uniform case we have:

$$\begin{aligned}
N_{ss} &= 1 - \int_0^{\bar{x}_{ss}(N_{ss})} \tilde{m}(s; N_{ss}) dx \\
&= 1 - \int_0^{\bar{x}_{ss}} \frac{1}{U} \left[1 - \frac{\left(e^{\gamma x} + e^{-\gamma x}\right)}{\left(e^{\gamma \bar{x}_{ss}} + e^{-\gamma \bar{x}_{ss}}\right)} \right] dx \\
&= 1 - \frac{\bar{x}_{ss}}{U} + \frac{\left(e^{\gamma \bar{x}_{ss}} - e^{-\gamma \bar{x}_{ss}}\right)}{\gamma U \left(e^{\gamma \bar{x}_{ss}} + e^{-\gamma \bar{x}_{ss}}\right)}
\end{aligned}$$

and

$$\begin{aligned}
Z_{ss} &= U/2 - \int_0^{\bar{x}_{ss}(N_{ss})} x \tilde{m}(s; N_{ss}) dx \\
&= U/2 - \int_0^{\bar{x}_{ss}} \frac{x}{U} \left[1 - \frac{\left(e^{\gamma x} + e^{-\gamma x}\right)}{\left(e^{\gamma \bar{x}_{ss}} + e^{-\gamma \bar{x}_{ss}}\right)} \right] dx \\
&= U/2 - \frac{\bar{x}_{ss}^2}{2U} + \frac{1}{U \left(e^{\gamma \bar{x}_{ss}} + e^{-\gamma \bar{x}_{ss}}\right)} \int_0^{\bar{x}_{ss}} \left(x e^{\gamma x} + x e^{-\gamma x} \right) dx \\
&= U/2 - \frac{\bar{x}_{ss}^2}{2U} + \frac{\bar{x}}{\gamma U} \frac{\left(e^{\gamma \bar{x}_{ss}} - e^{-\gamma \bar{x}_{ss}}\right)}{\left(e^{\gamma \bar{x}_{ss}} + e^{-\gamma \bar{x}_{ss}}\right)} + \frac{1}{\gamma^2 U} \frac{2}{\left(e^{\gamma \bar{x}_{ss}} + e^{-\gamma \bar{x}_{ss}}\right)} - \frac{1}{\gamma^2 U}
\end{aligned}$$

E.5 Perturbation and Stability of Steady States

In this section we analyze the linearization of the planning problem around the steady state. This linearization is analogous to the one for the equilibrium in Section 5.

We approximate $\bar{x}(t) = \mathcal{X}^P(N, Z)(t)$ by taking the directional derivative (Gateaux) with respect to arbitrary perturbations n of a constant path N, and z of a constant path Z. In particular, we consider paths defined by $N(t) = N_{ss} + \epsilon\, n(t)$ and $Z(t) = Z_{ss} + \epsilon\, z(t)$ around the steady state N_{ss} and Z_{ss}. We will denote this Gateaux derivative by $\bar{y}$.

Proposition 12. Let λ_T be equal to the steady state value function $\tilde{\lambda}$ corresponding to that steady state. Let $n : [0, T] \to \mathbb{R}$ and $z : [0, T] \to \mathbb{R}$ be two arbitrary perturbations. Then

$$\begin{aligned}
\bar{y}(t) &\equiv \lim_{\epsilon \downarrow 0} \frac{\mathcal{X}^{\mathcal{P}}(N_{ss} + \epsilon n, Z_{ss} + \epsilon z; \tilde{\lambda})(t) - \mathcal{X}^P(N_{ss}, Z_{ss}; \tilde{\lambda})(t)}{\epsilon} \\
&= \int_t^T G_{yn}(\tau - t) n(\tau) d\tau + \int_t^T G_{yz}(\tau - t) z(\tau) d\tau \qquad (70)
\end{aligned}$$

where

$$G_{yn}(\tau - t) = \frac{\theta_n}{\tilde{\lambda}_{xx}(\bar{x}_{ss})} \sum_{j=0}^{\infty} c_j e^{-\psi_j(\tau-t)} n(\tau) d\tau$$

$$G_{yz}(\tau - t) = \frac{2\theta_n}{\tilde{\lambda}_{xx}(\bar{x}_{ss})\bar{x}_{ss}} \sum_{j=0}^{\infty} c_j e^{-\psi_j(\tau-t)} z(\tau) d\tau$$

and ψ_j, c_j, and γ are defined as in Proposition 7.

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

Proposition 13. Let $\tilde{m}$ be the corresponding steady state distribution of non-adopters for the planner. 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

$$\begin{aligned} z(t) &\equiv \lim_{\epsilon \downarrow 0} \frac{\mathcal{Z}(\bar{x}_{ss} + \epsilon y; \tilde{m} + \epsilon w)(t) - \mathcal{Z}(\bar{x}_{ss}; \tilde{m})(t)}{\epsilon} \\ &= z_0(\omega)(t) + \int_0^t H_{zy}(t-s)\bar{y}(s) ds \end{aligned} \tag{71}$$

where

$$z_0(\omega)(t) \equiv -\sum_{j=0}^{\infty} \frac{\bar{x}_{ss}^2(\pi j + \frac{1}{2} - \cos(j\pi))}{\pi(\frac{1}{2} + j)} \frac{\langle \varphi_j, \omega \rangle}{\langle \varphi_j, \varphi_j \rangle} e^{-\mu_j t} \text{ and} \tag{72}$$

$$H_{zy}(q) = \tilde{m}_x(\bar{x}_{ss})\sigma^2 \sum_{j=0}^{\infty} \eta_j e^{-\mu_j q} \tag{73}$$

where $\varphi_j, \tilde{m}_x, \mu_j$ and γ are defined as in Proposition 8.

Thus we can write $Z(t) = Z_{ss} + \epsilon z(t) + o(\epsilon)$. This formula has the effect of two perturbations. One is the perturbation on the initial condition m_0 given by ω, whose effect is in the term $z_0(\omega)(t)$. Alternatively, $z_0(\omega)(t)$ is the effect at time t on the path $Z(t)$ of a perturbation of the initial condition keeping the threshold rule $\bar{x}$ fixed. As in the case of n_0 we can specialize ω by Dirac-delta function $\delta_{\hat{x}}$, so that we concentrate the perturbation around a value $x = \hat{x}$. The proof of this can be found in Appendix E.6.1.

Theorem 3. Let $\bar{x}_{ss}$ be the steady state of the planner problem, with its corresponding N_{ss}, Z_{ss}, and let $\tilde{m}$ be the corresponding steady state distribution of non-adopters. Let $m_0(x) = \tilde{m}(x) + \epsilon\omega(x)$. Let λ_T be equal to the value function $\tilde{\lambda}$ corresponding to that steady state. The linearized equilibrium must solve

$$\bar{y}(t) = \bar{y}_0(t) + \tilde{\Theta} \int_0^T \tilde{K}(t,s)\bar{y}(s)ds \text{ where} \tag{74}$$

$$\bar{y}_0(\omega)(t) \equiv \int_t^T G_{yn}(\tau - t)n_0(\omega)(\tau)d\tau + \int_t^T G_{yz}(\tau - t)z_0(\omega)(\tau)d\tau \tag{75}$$

where n_0 is derived in Proposition 8, z_0 is derived in Proposition 13, $\tilde{\Theta} \equiv \frac{\theta_n \tilde{m}_x(\bar{x}_{ss})\sigma^2}{\tilde{\lambda}_{xx}(\bar{x}_{ss})\bar{x}_{ss}}$ and where the kernel $\tilde{K}$ is given by

$$\tilde{K}(t,s) = \sum_{j=0}^{\infty}\sum_{i=0}^{\infty}(c_j + c_i)e^{\psi_j t + \mu_i s}\left(\frac{e^{-(\psi_j+\mu_i)\max\{t,s\}} - e^{-(\psi_j+\mu_i)T}}{\psi_j + \mu_i}\right) > 0 \tag{76}$$

We have that $\text{Lip}_{\tilde{K}} \leq \left(\frac{\bar{x}_{ss}^2}{\sigma^2}\right)^2$. Furthermore, if $\tilde{\Theta}\,\text{Lip}_{\tilde{K}} < 1$ there exists a unique bounded solution to equation (74) which is the limit of

$$\bar{y}(t) = \left[I + \tilde{\Theta}\tilde{\mathcal{K}} + \tilde{\Theta}^2\tilde{\mathcal{K}}^2 + \dots\right]\bar{y}_0(\omega) \quad \text{where} \quad \tilde{\mathcal{K}}(g)(t) \equiv \int_0^T \tilde{K}(t,s)g(s)ds$$

and where $\tilde{\mathcal{K}}^{j+1}(g)(t) \equiv \int_0^T \tilde{K}(t,s)\,\tilde{\mathcal{K}}^j(g)(s)\,ds$ for any bounded $g : [0,T] \to \mathbb{R}$. The operator $\tilde{\mathcal{K}}$ is self-adjoint, and positive definite.

We again consider a perturbation to the steady state density of non-adopters. In this case, we let $m_0(x)$ equal the steady state distribution of no-adopters of the decentralized problem, so that the shock resembles a starting equilibrium with lower adoption than that prescribed by the planning solution.

E.6 Perturbation of the Planning Problem

We consider the planning problem with $\{\bar{x}(t,\epsilon), N(t,\epsilon), \lambda(x,t,\epsilon), m(x,t,\epsilon)\}$. We again linearize this equilibrium with respect to ϵ and evaluate it at $\epsilon = 0$. We differentiate $\lambda(x,t,\epsilon)$ with respect to ϵ at each (x,t) to obtain $\ell(x,t) \equiv \frac{\partial}{\partial\epsilon}\lambda(x,t,\epsilon)\big|_{\epsilon=0}$ which solves the following p.d.e

$$\rho\ell(x,t) = x\theta_n n(t) + \theta_n z(t) + \frac{\sigma^2}{2}\ell_{xx}(x,t) + \ell_t(x,t) \tag{77}$$

for $x \in [0, \bar{x}_{ss}]$ and $t \in [0, T]$ and where $z(t) \equiv \frac{\partial}{\partial \epsilon} Z(t, \epsilon)\big|_{\epsilon=0}$ and $n(t) \equiv \frac{\partial}{\partial \epsilon} N(t, \epsilon)\big|_{\epsilon=0}$. The boundary conditions are:

$$
\begin{aligned}
\ell(x, T) &= 0 \\
\ell_x(0, t) &= 0 \\
\ell(\bar{x}_{ss}, t) &= 0 \\
\tilde{\lambda}_{xx}(\bar{x}_{ss}) \bar{y}(t) + \ell_x(\bar{x}_{ss}, t) &= 0
\end{aligned} \tag{78}
$$

PROPOSITION 14. The solution for the KBE equation for ℓ is given by

$$
\ell(x, t) = \sum_{j=0}^{\infty} \varphi_j(x) \hat{\ell}(t) \qquad \text{for } x \in [0, \bar{x}_{ss}] \text{ and } t \in [0, T]
$$

where for all $j = 1, 2, \ldots$ we have:

$$
\begin{aligned}
\hat{\ell}(t) &= \int_t^T e^{-\psi_j(\tau - t))} \hat{s}_j(\tau) d\tau && \text{for } t \in [0, T] \\
\hat{s}_j(t) &= -\theta_n n(t) \frac{\langle \varphi_j, x \rangle}{\langle \varphi_j, \varphi_j \rangle} - \theta_n z(t) \frac{\langle \varphi_j, 1 \rangle}{\langle \varphi_j, \varphi_j \rangle} && \text{for } t \in [0, T] \\
\varphi_j(x) &= \sin\left(\left(\frac{1}{2} + j \right) \pi \left(1 - \frac{x}{\bar{x}_{ss}} \right) \right) && \text{for } x \in [0, \bar{x}_{ss}] \\
\langle \varphi_j, h \rangle &\equiv \int_0^1 h(x) \varphi_j(x) dx \\
\hat{\ell}(T) &= 0 \\
\psi_j &= \rho + \frac{1}{2} \sigma^2 \left(\frac{\pi(\frac{1}{2} + j)}{\bar{x}_{ss}} \right)^2
\end{aligned}
$$

The proof can be done by verifying that the equation hold at the boundaries, that for $t > 0$ the p.d.e holds in the interior since

$$
\hat{\ell}'_j(t) = \psi_j \hat{\ell}(t) + \hat{s}_j(t) \qquad \text{for } t \in [0, T] \text{ and } j = 1, 2, \ldots
$$

and since $\{\varphi_j(x)\}$ form an orthogonal bases for functions, and finally that the boundary holds at $t = 0$ for $x \in [0, \bar{x}_{ss}]$.

Note that the derivative of the solution for λ is

$$\ell_x(\bar{x}_{ss}, t) = -\theta_n \int_t^T \sum_{j=0}^{\infty} c_j e^{-\psi_j(\tau - t)} n(\tau) d\tau - \theta_n \frac{2}{\bar{x}_{ss}} \int_t^T \sum_{j=0}^{\infty} e^{-\psi_j(\tau - t)} z(\tau) d\tau$$

where $c_j = 2\left(1 - \frac{\cos(\pi j)}{\pi(j+\frac{1}{2})}\right)$.

E.6.1 Perturbation Analysis of the Planning Problem

Recall that from equation (78), $\bar{y}(t)$ is equal to

$$\begin{aligned}
\bar{y}(t) &= \frac{-\ell_x(\bar{x}_{ss}, t)}{\tilde{\lambda}_{xx}(\bar{x}_{ss})} \\
&= \int_t^T \frac{\theta_n}{\tilde{\lambda}_{xx}(\bar{x}_{ss})} \sum_{j=0}^{\infty} c_j e^{-\psi_j(\tau - t)} n(\tau) d\tau + \int_t^T \frac{2\theta_n}{\tilde{\lambda}_{xx}(\bar{x}_{ss}) \bar{x}_{ss}} \sum_{j=0}^{\infty} c_j e^{-\psi_j(\tau - t)} z(\tau) d\tau \\
&= \int_t^T G_{yn}(\tau - t) n(\tau) d\tau + \int_t^T G_{yz}(\tau - t) z(\tau) d\tau
\end{aligned} \tag{79}$$

The expression for $n(t)$ is given by equation (69) and can be written as

$$n(t) = n_0(t) + \int_0^t H_{ny}(t - s) \bar{y}(s) ds.$$

where as before $n_0(t) \equiv -\sum_{j=0}^{\infty} \frac{\bar{x}_{ss}}{\pi(\frac{1}{2}+j)} \frac{\langle \varphi_j, \omega \rangle}{\langle \varphi_j, \varphi_j \rangle} e^{-\mu_j t}$. We can obtain a similar expression for $z(t)$ using the solution for $p(x, t)$ as

$$\begin{aligned}
z(t) &= -\int_0^{\bar{x}_{ss}} x p(x, t) dx \\
&= -\sum_{j=0}^{\infty} \hat{p}_j(t) \int_0^{\bar{x}_{ss}} x \varphi_j(x) dx \\
&= -\sum_{j=0}^{\infty} \frac{\bar{x}_{ss}^2(\pi(j+\frac{1}{2}) - \cos(j\pi)}{(\pi(\frac{1}{2}+j))^2} \frac{\langle \varphi_j, \omega \rangle}{\langle \varphi_j, \varphi_j \rangle} e^{-\mu_j t} + \tilde{m}_x(\bar{x}_{ss}) \sigma^2 \int_0^t \sum_{j=0}^{\infty} \frac{\pi(j+\frac{1}{2}) - \cos(j\pi))}{\pi(j+\frac{1}{2})} e^{-\mu_j(t-\tau)} \bar{y}(\tau) d\tau \\
&= z_0(t) + \int_0^t H_{zy}(t - s) \bar{y}(s) ds
\end{aligned}$$

where $z_0(t) \equiv -\sum_{j=0}^{\infty} \frac{c_j}{2} \frac{\bar{x}_{ss}^2}{\pi(j+\frac{1}{2})} \frac{\langle \varphi_j, \omega \rangle}{\langle \varphi_j, \varphi_j \rangle} e^{-\mu_j t}$ and $c_j \equiv \left(1 - \frac{\cos(\pi j)}{\pi(j+\frac{1}{2})}\right)$. Then, equation (79) can be written as

$$
\begin{aligned}
\bar{y}(t) &= \int_t^T G_{yn}(\tau - t)\left(n_0(\tau) + \int_0^t H_{ny}(\tau - s)\bar{y}(s)ds\right)d\tau \\
&+ \int_t^T G_{yz}(\tau - t)\left(z_0(\tau) + \int_0^t H_{zy}(\tau - s)\bar{y}(s)ds\right)d\tau \\
&= \int_t^T G_{yn}(\tau - t)n_0(\tau)d\tau + \int_t^T \int_0^t G_{yn}(\tau - t)H_{ny}(\tau - s)\bar{y}(s)ds\, d\tau \\
&+ \int_t^T G_{yz}(\tau - t)z_0(\tau)d\tau + \int_t^T \int_0^t G_{yz}(\tau - t)H_{zy}(\tau - s)\bar{y}(s)ds\, d\tau \\
&= \bar{y}_0(t) + \int_0^T M(t,s)\bar{y}(s)ds
\end{aligned}
$$

where

$$
\bar{y}_0(t) \equiv \int_t^T G_{yn}(\tau - t)n_0(\tau)d\tau + \int_t^T G_{yz}(\tau - t)z_0(\tau)d\tau
$$

and

$$
\begin{aligned}
\int_0^T M(t,s)\bar{y}(s)ds &\equiv \int_t^T \int_0^t G_{yn}(\tau - t)H_{ny}(\tau - s)\bar{y}(s)ds\, d\tau + \int_t^T \int_0^t G_{yz}(\tau - t)H_{zy}(\tau - s)\bar{y}(s)ds\, d\tau \\
&= \int\limits_0^T \int\limits_{\max\{t,s\}}^T G_{yn}(\tau - t)H_{ny}(\tau - s)\bar{y}(s)ds\, d\tau + \int\limits_0^T \int\limits_{\max\{t,s\}}^T G_{yz}(\tau - t)H_{zy}(\tau - s)\bar{y}(s)ds\, d\tau
\end{aligned}
$$

with

$$
\begin{aligned}
G_{yn}(w) &= \frac{\theta_n}{\tilde{\lambda}_{xx}(\bar{x}_{ss})}\sum_{j=0}^{\infty} c_j e^{-\psi_j(w)} \\
G_{yz}(w) &= \frac{2\theta_n}{\tilde{\lambda}_{xx}(\bar{x}_{ss})\bar{x}_{ss}}\sum_{j=0}^{\infty} e^{-\psi_j(w)} \\
H_{zy}(q) &= \frac{\tilde{m}_x(\bar{x}_{ss})\sigma^2}{2}\sum_{j=0}^{\infty} c_j e^{-\mu_j(q)} \\
H_{ny}(q) &= \frac{\tilde{m}_x(\bar{x}_{ss})\sigma^2}{\bar{x}_{ss}}\sum_{j=0}^{\infty} e^{-\mu_j(q)}
\end{aligned}
$$

where $e^{-rq}G_{yn}(w)H_{ny}(q) = G_{yz}(w)H_{zy}(q)e^{-rq}$. Using the definitions of $n_0(t)$ and $z_0(t)$ we first find the value of $\bar{y}_0(t)$ as

$$\begin{aligned}
\bar{y}_0(t) &\equiv \int_t^T G_{yn}(\tau-t)n_0(\tau)d\tau + \int_t^T G_{yz}(\tau-t)z_0(\tau)d\tau \\
&= \frac{-\theta_n}{\tilde{\lambda}_{xx}(\bar{x}_{ss})}\int_t^T \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} c_j \frac{\bar{x}_{ss}}{\pi(\frac{1}{2}+i)}\frac{\langle\varphi_i,\omega\rangle}{\langle\varphi_i,\varphi_i\rangle}e^{-\psi_j(\tau-t)}e^{-\mu_i\tau}d\tau \\
&+ \frac{-\theta_n}{\tilde{\lambda}_{xx}(\bar{x}_{ss})}\int_t^T \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} c_i \frac{\bar{x}_{ss}}{\pi(\frac{1}{2}+i)}\frac{\langle\varphi_i,\omega\rangle}{\langle\varphi_i,\varphi_i\rangle}e^{\psi_j t}e^{-\psi_j(\tau-t)}e^{-\mu_i\tau}d\tau \\
&= \frac{-\theta_n}{\tilde{\lambda}_{xx}(\bar{x}_{ss})}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}(c_j+c_i)\frac{\bar{x}_{ss}}{\pi(\frac{1}{2}+i)}\frac{\langle\varphi_i,\omega\rangle}{\langle\varphi_i,\varphi_i\rangle}e^{\psi_j t}\left(\frac{e^{-(\psi_j+\mu_i)t}-e^{-(\psi_j+\mu_i)T}}{\psi_j+\mu_i}\right) \qquad (80)
\end{aligned}$$

Then, we find

$$\begin{aligned}
\int_0^T M(t,s)\bar{y}(s)ds &= \int_0^T \left(\int_{\max\{t,s\}}^T G_{yn}(\tau-t)H_{ny}(\tau-s)\bar{y}(s)d\tau + \int_{\max\{t,s\}}^T G_{yz}(\tau-t)H_{zy}(\tau-s)\bar{y}(s)d\tau\right)ds \\
&= \tilde{\Theta}(\bar{x}_{ss})\int_0^T\int_{\max\{t,s\}}^T \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} c_j e^{-\psi_j(\tau-t)}e^{-\mu_i(\tau-t)}\bar{y}(s)d\tau\, ds \\
&+ \tilde{\Theta}(\bar{x}_{ss})\int_0^T\int_{\max\{t,s\}}^T \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} c_i e^{-\psi_j(\tau-t)}e^{-\mu_i(\tau-t)}\bar{y}(s)d\tau\, ds
\end{aligned}$$

where we let $\tilde{\Theta}(\bar{x}_{ss}) \equiv \frac{\theta_n\tilde{m}_x(\bar{x}_{ss})\sigma^2}{\tilde{\lambda}_{xx}(\bar{x}_{ss})\bar{x}_{ss}}$. Solving the integrals we get

$$\begin{aligned}
\int_0^T M(t,s)\bar{y}(s)ds &= \tilde{\Theta}(\bar{x}_{ss})\int_0^T\left(\sum_{j=0}^{\infty}\sum_{i=0}^{\infty} c_j e^{\psi_j t+\mu_i s}\left(\frac{e^{-(\psi_j+\mu_i)\max\{t,s\}}-e^{-(\psi_j+\mu_i)T}}{\psi_j+\mu_i}\right)\right)ds \\
&+ \tilde{\Theta}(\bar{x}_{ss})\int_0^T\left(\sum_{j=0}^{\infty}\sum_{i=0}^{\infty} c_i e^{\psi_j t+\mu_i s}\left(\frac{e^{-(\psi_j+\mu_i)\max\{t,s\}}-e^{-(\psi_j+\mu_i)T}}{\psi_j+\mu_i}\right)\right)ds \\
&= \tilde{\Theta}(\bar{x}_{ss})\int_0^T\left(\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}(c_j+c_i) e^{\psi_j t+\mu_i s}\left(\frac{e^{-(\psi_j+\mu_i)\max\{t,s\}}-e^{-(\psi_j+\mu_i)T}}{\psi_j+\mu_i}\right)\right)ds \\
&= \tilde{\Theta}(\bar{x}_{ss})\int_0^T \tilde{K}(t,s)ds.
\end{aligned}$$

Thus, equation (79) can be written as

$$\bar{y}(t) = \bar{y}_0(t) + \tilde{\Theta}(\bar{x}_{ss})\int_0^T \tilde{K}(t,s)\bar{y}(s)ds$$

Notice also that since $e^{-rt}M(t,s) = e^{-rs}M(t,s)$

$$\int_0^T e^{-rt}M(t,s)\bar{y}(s)ds = \tilde{\Theta}(\bar{x}_{ss}) \int_0^T \left(\sum_{j=0}^{\infty} \sum_{i=0}^{\infty} (c_j + c_i) e^{\mu_j t + \mu_i s} \left(\frac{e^{-(r+\mu_j+\mu_i)\max\{t,s\}} - e^{-(r+\mu_j+\mu_i)T}}{\mu_j + \mu_i + r} \right) \right) ds$$

F A "Pure" Learning Model

In this section, we develop a model with random diffusion of the technology across agents. 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. Newborn agents start as uninformed, and become informed by randomly matching with informed agents. 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 main conclusions are that the pure learning model differs from the model with strategic complementarity in that:

1. it has a unique equilibrium, and a unique stable steady state,
2. it has a logistic S shape adoption profile, provided the initial share of uninformed is small enough,
3. the use of the technology for those that adopt depends only on the cohort, and not the size of the network,
4. the equilibrium is constrained efficient: the optimal subsidy to use the technology is zero.

Learning Setup. We follow the canonical notation for an "SIR" model and assume that the population, normalized to have measure 1, is split between the uninformed, whose measure we denote by $S(t)$, and the informed, which have measure $I(t)$, so that $I(t)+S(t) = 1$. Those that are informed can be split in two groups, those that have adopted the technology, with measure $N(t)$, and those informed that have not adopted $M(t)$, so that $I(t) = M(t) + N(t)$.

The main assumption about learning about the technology is that agents do *not* need to use the technology to learn about it. In particular, agents that know about the technology will randomly meet agents that don't and transmit the information in such way. Recall that among the $I(t)$ informed agents, only a $N(t)$ have adopted, and $M(t)$ are informed but have decided not to adopt.

Optimal Adoption. Now we turn to the decision of agents. The uninformed agents have no decision to make. The decision problem of those that are informed is similar to the steady state problem in our model with strategic complementarities.

The value of an agent that already has adopted the technology is

$$\rho a(x) = \theta_0 x + \frac{\sigma^2}{2} a_{xx}(x) \text{ for } x \in [0, U]$$

with boundaries $a_x(0) = a_x(U) = 0$ The value function for an agent that is informed is:

$$\rho v(x) = \max \left\{ \frac{\sigma^2}{2} v_{xx}(x) \, , \, \rho(a(x) - c) \right\}$$

with time invariant threshold $\bar{x} < U$ solving, and boundary at zero:

$$v_x(\bar{x}) = a_x(\bar{x}) \text{ and } v(\bar{x}) = a(\bar{x}) - c \text{ and } v_x(0) = 0$$

The solution of v and a are identical to the steady state solutions of the baseline model $\tilde{v}$ and $\tilde{a}$ where we set $\theta_n = 0$. Likewise the solution for $\bar{x}$ is the same as the value $\bar{x}_{ss}$ for the model with $\theta_n = 0$.

Evolution of Distributions. Now we turn to the description of the distribution of agents across states. We let $s(x,t)$ the density of those uninformed at t with x, and $m(x,t)$ the density of those informed at t with x and that have not adopted yet. First we characterize g which satisfies:

$$s_t(x,t) = \frac{\sigma^2}{2} s_{xx}(x,t) - (\nu + \beta(S(t)))\, s(x,t) + \nu \frac{1}{U} \text{ all } t \geq 0 \text{ and } x \in [0, U]$$

with boundary conditions given by reflections at the boundary, i.e., $0 = s_x(0,t) = s_x(U,t)$ all $t \geq 0$ and initial condition independent of x:

$$s(x, 0) = s_0 \text{ all } x \in [0, U]$$

In this case $S(t)$ is the total measure of uninformed agents at time t, and $\beta(\cdot)$ is a function that gives the probability per uninformed of becoming informed:

$$S(t) = \int_0^U s(x,t) dx$$

We assume that $\beta(\cdot)$ is given by

$$\beta(S) = \beta_0\,(1 - S) = \beta_0\,I \text{ for some constant } \beta_0 > \nu > 0$$

The interpretation is that each agent has β_0 meeting per unit of time, and that a fraction $1 - S$ are with those informed of the technology.

We will return to solve for S and I below. Now we turn to the law of motion for m is:

$$\begin{aligned} m_t(x,t) &= \frac{\sigma^2}{2} m_{xx}(x,t) + \beta(S(t))s(x,t) - \nu m(x,t) \text{ all } t \geq 0 \text{ and } x \in [0, \bar{x}] \\ m(x,t) &= 0 \text{ all } t \geq 0 \text{ and } x \in [\bar{x}, U] \end{aligned}$$

Continuity of m implies that $m(\bar{x}, t) = 0$ all $t \geq 0$. The reflecting barrier of x at zero implies $0 = m_x(0,t)$ for all $t \geq 0$.

Comparing with the baseline model with constant $\bar{x}$, the evolution of the density m has one main difference. Instead of having the constant inflow ν/U, it has a time varying, and smaller, inflow $\beta(S(t))s(x,t)$. This smaller inflow, everything else the same, can substantially retard the adoption.

We define the total number that are uninformed as:

$$M(t) \equiv \int_0^{\bar{x}} m(x,t)dx \leq I(t) = 1 - S(t)$$

The initial condition that the density of those that have not adopted is smaller than the density of those that are informed, i.e.: $0 \leq M(0) \leq I(0)$ all $x \in [0, U]$. Note that by integrating across x and using the boundary conditions:

$$M_t(t) = \int_0^{\bar{x}} m_t(x,t)dx = \frac{\sigma^2}{2} m_x(\bar{x}, t) + \beta(S(t))S(t)\frac{\bar{x}}{U} - \nu M(t) \text{ all } t \geq 0 \text{ and } x \in [0, \bar{x}]$$

We are interested in: $N(t) = 1 - S(t) - M(t)$, which using the previous equations gives:

$$N_t(t) = -\frac{\sigma^2}{2} m_x(\bar{x}, t) - \nu N(t) + \beta(S(t))S(t)\left(1 - \frac{\bar{x}}{U}\right) \text{ for all } t \geq 0$$

with initial condition $N(0) = \left(1 - \frac{\bar{x}}{U}\right) I(0)$.

Note that since $m(x,t) > 0$ for $x < \bar{x}$ and $m(\bar{x},t) = 0$, then $m_x(\bar{x}, t) < 0$. The next proposition rewrite this expression which it is useful to interpret the determinants of the dynamics of $N(t)$.

Proposition 15. Assume that $s_0(x) = S_0/U$ for all $x \in [0, U]$, and that $\beta(S) = \beta_0(1 - S)$.

Then we can write $N(t)$ as function of path $I(t)$ and $m(\bar{x}, t)$ and the threshold $\bar{x}$:

$$N(t) = I(t)\left(1 - \frac{\bar{x}}{U}\right) + \int_0^t e^{-\nu(t-\tau)}\left[-\frac{\sigma^2}{2} m_x(\bar{x}, \tau)\right] d\tau \tag{81}$$

The expression in the right hand side of $N(t)$ in Proposition 15 has the following interpretation. The term $I(t)\left(1 - \frac{\bar{x}}{U}\right)$ has the fraction of those informed with values of x above the threshold $\bar{x}$. The second term takes into account the past flows of agents that were informed, whose value of x went from below $\bar{x}$ to higher than $\bar{x}$.

Solving for Path of $N(t), M(t), I(t), S(t)$ Given $\bar{x}$. The solution is recursive: we first solve for $S(t)$ and $I(t)$, and then using the path of $I(t)$ we solve for $N(t)$. This is done in the next two propositions.

Proposition 16. Assume that $\beta(S) = \beta_0(1 - S)$ for $\beta_0 > \nu$. Furthermore assume that $s_0(x) = S_0/U$ for all $x \in [0, U]$. For a given $I(0)$ we have that the unique solution of

$$\dot{I}(t) = \beta_0 I(t)\left[\left(1 - \frac{\nu}{\beta_0}\right) - I(t)\right]$$

is given by

$$I(t) = 1 - S(t) = \left(1 - \frac{\nu}{\beta_0}\right) \frac{e^{(\beta_0-\nu)t}}{\frac{\left(1-\frac{\nu}{\beta_0}\right)}{I(0)} - 1 + e^{(\beta_0-\nu)t}} \tag{82}$$

Thus, if $0 < I(0) < 1 - \frac{\nu}{\beta_0}$, then $I(t)$ converges monotonically to $I_{ss} = 1 - \frac{\nu}{\beta_0} \in (0, 1)$. If $I(0) < I_{ss}$, then

$$I(t) = \begin{cases} \text{is convex in } t & \text{if } t < \frac{\log((I_{ss}-I(0))/I(0))}{\beta_0-\nu} \text{ or } I(t) < \frac{I_{ss}}{2} \\ \text{is concave in } t & \text{if } t > \frac{\log((I_{ss}-I(0))/I(0))}{\beta_0-\nu} \text{ or } I(t) > \frac{I_{ss}}{2}. \end{cases}$$

As shown in Proposition 16, when $I(0)$ is small, then $I(t)$ displays a "logistic" type of path of technology adoption, but $I(t)$ is only the population that can adopt. We characterize the number of adopters in the next proposition.

Proposition 17. Assume that $s_0(x) = S_0/U$ for all $x \in [0, U]$. Take the path $I(t)$ as given, and the optimal threshold $\bar{x} < U$. Then the unique solution of $m(x,t)$ is:

$$m(x,t) = \sum_{j=0}^{\infty} \varphi_j(x)\hat{b}_j(t) \text{ where } \varphi_j(x) = \sin\left((j+\tfrac{1}{2})\pi\left(1-\frac{x}{\bar{x}}\right)\right)$$
$$\hat{b}_j(t) = \frac{2}{\pi(j+\frac{1}{2})}\left(e^{-\mu_j t}\frac{I(0)}{U} + \beta_0\int_0^t e^{-\mu_j(t-\tau)}\frac{I(\tau)\,(1-I(\tau))}{U}d\tau\right) \text{ and } \mu_j = \nu + \left((j+\tfrac{1}{2})\tfrac{\pi}{\bar{x}}\right)^2$$

and thus $N(t) = I(t) - M(t)$ is given by:

$$N(t) = I(t) - \tfrac{\bar{x}}{U}\left(H(t)I(0) + \beta_0\int_0^t H(t-\tau)I(\tau)\,(1-I(\tau))\,d\tau\right) \text{ where}$$
$$H(z) \equiv \sum_{j=0}^{\infty}\omega_j e^{-\mu_j z} \text{ with } \omega_j \equiv \frac{2}{\left(\pi(j+\frac{1}{2})\right)^2} > 0 \text{ and } \sum_{j=0}^{\infty}\omega_j = 1\,.$$

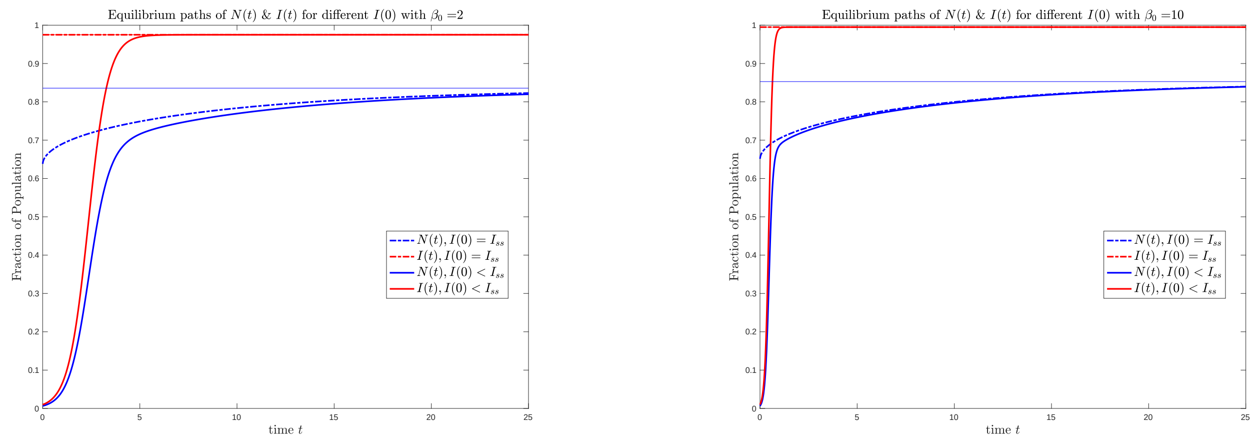
Combining the expression for $N(t)$ in Proposition 17 with the path of $I(t)$ solved for in Proposition 16 we obtain an explicit solution to $N(t)$. Next we analyze the invariant distribution in this model, which is the value at which it tends as $t \to \infty$. We denote $\tilde{m}$ the density for m which satisfies: $\nu\tilde{m}(x) = \frac{\sigma^2}{2}\tilde{m}_{xx}(x) + \beta_0(1-\frac{\nu}{\beta_0})\frac{\nu}{\beta_0}\frac{\bar{x}}{U}$ for all $x \in [0,\bar{x}]$ and $\tilde{m}_x(\bar{x}) = 0$ and $\tilde{m}(\bar{x}) = 0$. The next proposition gives the solution for $\tilde{m}$, as well as the steady state number of adopters N_{ss}.

Proposition 18. Assume that $s_0(x) = S_0/U$ for all $x \in [0,U]$, that $\bar{x} < U$, $\beta(S) = \beta_0(1-S)$, and that $\beta_0 > \nu > 0$. Then the steady state density $\tilde{m}$ is given by:

$$\tilde{m}(x) = (1-\tfrac{\nu}{\beta_0})\tfrac{1}{U}\left(1-\frac{\cosh(\gamma x)}{\cosh(\gamma\bar{x})}\right) \text{ where } \gamma = \sqrt{2\nu}/\sigma \text{ and thus}$$
$$N_{ss} = I_{ss} - \int_0^{\bar{x}}\tilde{m}(x)dx = (1-\tfrac{\nu}{\beta_0})\left[1-\tfrac{\bar{x}}{U}\left(1-\frac{\tanh(\gamma\bar{x})}{\gamma\bar{x}}\right)\right] \tag{83}$$

It is interesting to see that even if $I(0) = I_{ss} \equiv 1-\frac{\nu}{\beta_0}$, then $N(0) < N_{ss}$, and convergence will take time. In words, even if all agents are informed about the technology it takes time for the selection process to yield N_{ss}. In particular equation (83) implies that $N_{ss} > I_{ss}(1-\frac{\bar{x}}{U})$, since among the adopters there are agents who had $x \geq \bar{x}$ in the past and currently have $x < \bar{x}$.

Figure F2: Equilibrium paths of N and I of Pure Learning Model



Slow learning, $\beta_0 = 2$.

Fast Learning, $\beta_0 = 10$.

Figure F2 illustrates the main results of this section. The left and right panel differ in the value of β_0, with the left panel with a slow learning $\beta_0 = 2$, and the right panel a high value, $\beta_0 = 10$. In each panel we consider two initial condition for $I(0)$: one with $I(0) = I_{ss}$ (dotted lines), and with $I(0) = I_{ss}/100$ (solid lines). The remaining parameters are all the same. The paths for N are in blue, and the ones for A are in red. Focusing first in the slow learning case (left panel), note that when $I(0)$ is small, so that early on adoption is restricted by the information about the technology, the fraction that adopt $N(t)$ follows an approximate logistic path, as explained above. Instead, if $I(0) = I_{ss}$, then the path of $N(t)$ is concave in time, and starts at a high value at $t = 0$. In the case of fast learning, i.e., in the right panel, the same dynamics of learning are also present, but in a much abbreviated period of time.

Optimality of Equilibrium. The equilibrium path is constrained efficient. In particular, if the planner can only give a subsidy to those that use the technology, then the optimal subsidy is zero. This is because, given our assumptions about learning, such subsidy does not affect the fraction of people that learn about the application. Furthermore, since we assume that there is no complementary in the use of the technology, the individual decision will coincide with the planner decision for $\bar{x}$.

F.1 Proofs for the Learning Model

Proof. (Proposition 15) We start by integrating the differential equation for N to obtain

$$N(t) = e^{-\nu t} N(0) + \int_0^t e^{-\nu(t-s)} \left[-\frac{\sigma^2}{2} m_x(\bar{x}, s) + \beta(S(s)) S(s) \left(1 - \frac{\bar{x}}{U}\right) \right] ds$$
$$N(0) = \left(1 - \tfrac{\bar{x}}{U}\right) I(0)$$

Using that $\dot{I}(t) = \beta(S(t))S(t) - \nu I(t)$, so

$$\int_0^t e^{-\nu(t-s)} \beta(S(s)) S(s) ds = \int_0^t e^{-\nu(t-s)} \dot{I}(t) ds + \int_0^t e^{-\nu(t-s)} \nu I(t) ds$$

Integrating by parts:

$$\begin{aligned} \int_0^t e^{-\nu(t-s)} \beta(S(s)) S(s) ds &= I(t) - I(0)e^{-\nu t} - \int_0^t \nu e^{-\nu(t-s)} I(s) ds + \int_0^t e^{-\nu(t-s)} \nu I(t) ds \\ &= I(t) - I(0) e^{-\nu t} \end{aligned}$$

Thus:

$$\begin{aligned} N(t) &= e^{-\nu t} \left(1 - \tfrac{\bar{x}}{U}\right) I(0) + \int_0^t e^{-\nu(t-s)} \left[-\frac{\sigma^2}{2} m_x(\bar{x}, s) \right] ds + \left[I(t) - I(0) e^{-\nu t} \right] \left(1 - \frac{\bar{x}}{U}\right) \\ &= I(t) \left(1 - \frac{\bar{x}}{U}\right) + \int_0^t e^{-\nu(t-s)} \left[-\frac{\sigma^2}{2} m_x(\bar{x}, s) \right] ds \end{aligned}$$

□

Proof. (of Proposition 16) Integrating the p.d.e. for g we get:

$$S_t(t) \equiv \int_0^U s_t(x,t) dx = \frac{\sigma^2}{2} \int_0^U s_{xx}(x,t) dx - (\nu + \beta(S(t))) \int_0^U s(x,t) dx + \nu \frac{\int_0^U dx}{U}$$

and using its boundary conditions at $x = 0$ and $x = U$:

$$S_t(t) = -\left(\nu + \beta(S(t))\right) S(t) + \nu \text{ all } t \geq 0$$

with initial condition:

$$s(0) = S_0 \text{ for some constant } 0 \leq S_0 = 1 - I(0) \leq 1$$

Since we assume that $s_0(x)$ is constant across x, i.e. if

$$s_0(x) = \frac{S_0}{U} \text{ all } x \in [0, U]$$

then the solution satisfies

$$s(x,t) = \frac{S(t)}{U} \text{ all } t \geq 0 \text{ for all } x \in [0, U]$$

Thus we obtain

$$\begin{aligned} S' &= -\left(\nu + \beta_0(1-S)\right) S + \nu = (1-S)\left(\nu - \beta_0 S\right) \\ &= \nu\left(1-S\right)\left(1 - \frac{S}{S^*}\right) \end{aligned}$$

It is convenient to solve for the path of I, the fraction of agents informed of the technology, $I(t) + S(t) = 1$ for all $t \geq 0$, so:

$$I' = -I\left(\nu - \beta_0(1-I)\right) = \beta_0 I\left(I_{ss} - I\right) \text{ where } I_{ss} = 1 - \frac{\nu}{\beta_0}$$

Let $\tilde{I} = \beta_0 I$, so that:

$$\tilde{I}' = \tilde{I}\left(\tilde{I}_{ss} - \tilde{I}\right) = \tilde{I}_{ss}\tilde{I} - (\tilde{I})^2 \text{ where } \tilde{I}_{ss} = \beta_0 - \nu$$

Then we get that its solution is given by:

$$\tilde{I}(t) = \frac{\tilde{I}_{ss} e^{\tilde{I}_{ss} t}}{\frac{\tilde{I}_{ss}}{\tilde{I}(0)} - 1 + e^{\tilde{I}_{ss} t}}$$

Note that

$$\begin{aligned} I_{ss}\frac{d}{dt}\frac{\tilde{I}_{ss} e^{\tilde{I}_{ss} t}}{\frac{\tilde{I}_{ss}}{\tilde{I}(0)} - 1 + e^{\tilde{I}_{ss} t}} &= \tilde{I}_{ss}\frac{\tilde{I}_{ss} e^{\tilde{I}_{ss} t}}{\frac{\tilde{I}_{ss}}{\tilde{I}(0)} - 1 + e^{\tilde{I}_{ss} t}} - \frac{\tilde{I}_{ss} e^{\tilde{I}_{ss} t}\tilde{I}_{ss} e^{\tilde{I}_{ss} t}}{\left(\frac{\tilde{I}_{ss}}{\tilde{I}(0)} - 1 + e^{\tilde{I}_{ss} t}\right)^2} \\ &= \tilde{I}_{ss}\tilde{I}(t) - (\tilde{I}(t))^2 \end{aligned}$$

which verifies the answer. Using $I = \tilde{I}/\beta_0$ we obtain the desired result.

□

Proof. (of Proposition 17) Given the path $\{S(t)\}$ define

$$B(t) \equiv \beta(S(t))S(t)\frac{1}{U}$$

We start with

$$m(x,t) = \sum_{j=0}^{\infty} \varphi_j(x)\hat{b}_j(t) \text{ where } \varphi_j(x) = \sin\left((j+\tfrac{1}{2})\pi\left(1-\frac{x}{\bar{x}}\right)\right)$$

Note that each φ_j satisfies the lateral boundary conditions for $m(x,t)$ at $x = 0$ and $x = \bar{x}$ for all t. Then the p.d.e. can be written as:

$$\begin{aligned}
0 &= m_t(x,t) - \frac{\sigma^2}{2}m_{xx}(x,t) + \nu m(x,t) - B(t) \text{ or} \\
0 &= \sum_{j=0}^{\infty} \varphi_j(x)\left[\hat{b}_j'(t) + \nu\hat{b}_j(t) + \left((j+\tfrac{1}{2})\tfrac{\pi}{\bar{x}}\right)^2 b_j(t) - B(t)\tfrac{\langle\varphi_j,1\rangle}{\langle\varphi_j,\varphi_j,\rangle}\right]
\end{aligned}$$

or for each $j = 0, 1, \ldots$:

$$\hat{b}_j'(t) = -\left[\nu + \left((j+\tfrac{1}{2})\tfrac{\pi}{\bar{x}}\right)^2\right] b_j(t) + B(t)\tfrac{\langle\varphi_j,1\rangle}{\langle\varphi_j,\varphi_j,\rangle}$$

or letting $\mu_j = \left((j+\tfrac{1}{2})\tfrac{\pi}{\bar{x}}\right)^2$

$$\hat{b}_j(t) = \hat{b}_j(0)e^{-\mu_j t} + \tfrac{\langle\varphi_j,1\rangle}{\langle\varphi_j,\varphi_j,\rangle}\int_0^t e^{-\mu_j(t-s)}B(s)ds$$

On the other hand $\{\hat{b}_j(0)\}$ are given so that

$$M(0) = \frac{\bar{x}}{U}I(0)$$

so that $M(0) = \int_0^{\bar{x}} m_0(x)dx$ and if $m_0(x)$ does not depend on x we have $M(0) = \bar{x}\,m_0(x)$:

$$m_0(x) = \frac{M(0)}{\bar{x}} = \frac{I(0)}{U}$$

$$\hat{b}_j(0) = \tfrac{\langle\varphi_j,1\rangle}{\langle\varphi_j,\varphi_j\rangle}\frac{I(0)}{U}$$

which ensures:

$$\sum_{j=0}^{\infty}\hat{b}_j(0)\varphi_j(x) = \frac{I(0)}{U}$$

so

$$\hat{b}_j(t) = \tfrac{\langle\varphi_j,1\rangle}{\langle\varphi_j,\varphi_j\rangle}\left(e^{-\mu_j t}\frac{I(0)}{U} + \int_0^t e^{-\mu_j(t-s)}B(s)ds\right)$$

Finally,

$$\langle\varphi_j, 1\rangle = \frac{\bar{x}}{\pi(j+\frac{1}{2})} \text{ and } \langle\varphi_j,\varphi_j\rangle = \frac{\bar{x}}{2}$$

Thus,

$$\hat{b}_j(t) = \frac{2}{\pi(j+\frac{1}{2})}\left(e^{-\mu_j t}\frac{I(0)}{U} + \int_0^t e^{-\mu_j(t-s)}B(s)ds\right)$$

Thus, if we compute:

$$M(t) = \int_0^{\bar{x}} m(x,t)dx = \sum_{j=0}^{\infty}\hat{b}_j(t)\int_0^{\bar{x}}\varphi_j(x)dx = \sum_{j=0}^{\infty}\hat{b}_j(t)\langle\varphi_j,1\rangle$$

substituting the expression for $\hat{b}_j(t)$:

$$\begin{aligned} M(t) &= \sum_{j=0}^{\infty}\tfrac{(\langle\varphi_j,1\rangle)^2}{\langle\varphi_j,\varphi_j\rangle}\left(e^{-\mu_j t}\frac{I(0)}{U} + \int_0^t e^{-\mu_j(t-s)}B(s)ds\right) \\ &= \sum_{j=0}^{\infty}\frac{2}{\left(\pi(j+\frac{1}{2})\right)^2}\left(e^{-\mu_j t}\frac{I(0)}{U} + \int_0^t e^{-\mu_j(t-s)}B(s)ds\right) \end{aligned}$$

since

$$\frac{(\langle\varphi_j,1\rangle)^2}{\langle\varphi_j,\varphi_j\rangle} = \left(\frac{\bar{x}}{\pi(j+\frac{1}{2})}\right)^2\frac{1}{\bar{x}/2} = \bar{x}\frac{2}{\left(\pi(j+\frac{1}{2})\right)^2}$$

To check, note that at $t=0$:

$$M(0) = I(0)\frac{\bar{x}}{U}\sum_{j=0}^{\infty}\tfrac{(\langle\varphi_j,1\rangle)^2}{\langle\varphi_j,\varphi_j\rangle} = I(0)\frac{\bar{x}}{U}\sum_{j=0}^{\infty}\frac{2}{\left(\pi(j+\frac{1}{2})\right)^2}$$

since $1 = \sum_{j=0}^{\infty} \frac{2}{\left(\pi(j+\frac{1}{2})\right)^2}$ Thus

$$
\begin{aligned}
N(t) &= I(t) - \sum_{j=0}^{\infty} \tfrac{(\langle \varphi_j, 1\rangle)^2}{\langle \varphi_j, \varphi_j\rangle} \left(e^{-\mu_j t} \frac{I(0)}{U} + \int_0^t e^{-\mu_j (t-s)} B(s) ds \right) \\
&= I(t) - \sum_{j=0}^{\infty} \bar{x} \frac{2}{\left(\pi(j+\frac{1}{2})\right)^2} \left(e^{-\mu_j t} \frac{I(0)}{U} + \int_0^t e^{-\mu_j (t-s)} B(s) ds \right) \\
&= I(t) - \frac{\bar{x}}{U} \sum_{j=0}^{\infty} \frac{2}{\left(\pi(j+\frac{1}{2})\right)^2} \left(e^{-\mu_j t} I(0) + \beta_0 \int_0^t e^{-\mu_j (t-s)} I(s)(1-I(s)) ds \right)
\end{aligned}
$$

So we can write:

$$
\begin{aligned}
& N(t) = I(t) - \tfrac{\bar{x}}{U} \left(\sum_{j=0}^{\infty} \omega_j e^{-\mu_j t} I(0) + \beta_0 \int_0^t \sum_{j=0}^{\infty} \omega_j e^{-\mu_j (t-s)} I(s) \left(1 - I(s)\right) ds \right) \text{ where} \\
& \omega_j \equiv \frac{2}{\left(\pi(j+\frac{1}{2})\right)^2} > 0 \text{ and } \sum_{j=0}^{\infty} \omega_j = 1\,.
\end{aligned}
$$

Defining

$$
H(z) \equiv \sum_{j=0}^{\infty} \omega_j e^{-\mu_j z}
$$

we can write:

$$
\begin{aligned}
& N(t) = I(t) - \tfrac{\bar{x}}{U} \left(H(t) I(t) + \beta_0 \int_0^t H(t-s) I(s) \left(1 - I(s)\right) ds \right) \text{ where} \\
& \omega_j \equiv \frac{2}{\left(\pi(j+\frac{1}{2})\right)^2} > 0 \text{ and } \sum_{j=0}^{\infty} \omega_j = 1\,.
\end{aligned}
$$

□

Proof. (of Proposition 18) We can rewrite the o.d.e. for $\tilde{m}$ as:

$$
\tilde{m}(x) = \tfrac{\sigma^2}{2\nu} \tilde{m}_{xx}(x) + \left(1 - \tfrac{\nu}{\beta_0}\right) \tfrac{1}{U} \text{ for all } x \in [0, \bar{x}]
$$

The solution is given by a sum of particular solution, $(1-\frac{\nu}{\beta_0})\frac{1}{U}$, and two homogenous solutions. The homogenous solutions are exponentials $\exp(\pm\gamma x)$. The requirement that $\tilde{m}_x(0) = 0$ implies that the coefficient that multiplies each of the exponentials has the same absolute value but opposite sign, i.e., the two homogenous solutions combine into a cosh. Then, imposing that $\tilde{m}(\bar{x}) = 0$ we get:

$$\tilde{m}(x) = (1 - \tfrac{\nu}{\beta_0})\tfrac{1}{U}\left(1 - \frac{\cosh(\gamma x)}{\cosh(\gamma \bar{x})}\right) \text{ where } \gamma = \sqrt{2\nu}/\sigma$$

Thus, using that $\int_0^{\bar{x}} \frac{\cosh(\gamma x)}{\cosh(\gamma \bar{x})} = \frac{\tanh(\gamma \bar{x})}{\gamma}$ we obtain the desired result.

$\square$

G HJB Equations for $a(x, t)$ and $v(x, t)$

Moreover, $a(x, t)$ solves the p.d.e. and boundary conditions for all $t \geq 0$:

$$\rho a(x,t) = x(\theta_0 + \theta_n N(t)) + \frac{\sigma^2}{2} a_{xx}(x,t) + a_t(x,t) \text{ if } x \in [0, U]$$
$$a_x(0,t) = a_x(U,t) = 0$$

where the boundary conditions arise from our assumption of reflecting barriers. Throughout, we assume $0 \leq a(x,t) \leq \frac{U(\theta_0+\theta_n)}{\rho}$ for all x, t, and $0 < c < \frac{U(\theta_0+\theta_n)}{\rho}$.

Adoption Decision: The value function of an agent that has not adopted solves the following variational inequality:

$$\rho v(x,t) = \max\left\{\frac{\sigma^2}{2} v_{xx}(x,t) + v_t(x,t) \,,\, \rho\left(-c + a(x,t)\right)\right\}$$

for all $t \geq 0$ and $x \in [0, U]$. We conjecture that the optimal decision rule is given by a path for the threshold $\bar{x}(t) \in (0, U)$ such so that, for each $t \geq 0$, the following holds

$$\rho v(x,t) = \frac{\sigma^2}{2} v_{xx}(x,t) + v_t(x,t) \text{ if } 0 \leq x \leq \bar{x}(t)$$
$$v(x,t) = -c + a(x,t) \text{ if } \bar{x}(t) \leq x \leq U$$

If $v(\cdot, t)$ is C^1 we have the following boundary conditions for all $t \geq 0$:

$$v(\bar{x}(t), t) = a(\bar{x}(t), t) - c \qquad \text{Value Matching}$$
$$v_x(\bar{x}(t), t) = a_x(\bar{x}(t), t) \qquad \text{Smooth Pasting}$$
$$v_x(0, t) = 0 \qquad \text{Reflecting}$$

where the first one is the value matching condition, the second the smooth pasting condition, and the last one arises from the reflecting barrier at $x = 0$.

H The Dynamics of the Deterministic Model

For each x, let $a(x,t) = x\,\alpha(t)$ where $\rho\alpha(t) = \theta_0 + \theta_n N(t) + \alpha_t(t)$. We fix an x and a path $\alpha(t)$ for $t \geq 0$. Let $t^*(x)$ be the optimal t that solves the adoption problem:

$$t^*(x) = \arg\max_{t\geq 0} G(t,x) \text{ with } G(x,t) \equiv e^{-\rho t}(\alpha(t)x - c)$$

The necessary first order conditions if $\alpha(t)$ is differentiable at $t = t^*(x) < \infty$ are:[38]

$$-\rho(\alpha(t^*)x - c) + x\alpha_t(t^*) = 0 \text{ if } t^*(x) > 0$$
$$-\rho(\alpha(0)x - c) + x\alpha_t(0) \leq 0 \text{ if } t^*(x) = 0$$

Furthermore, since not adopting is feasible (i.e., $t = \infty$) and yields a zero payoff, then

$$\alpha(t^*(x))x \geq c \text{ for all } x.$$

Given $t^*(x)$ we can define $\bar{x}(t)$ as the smallest value of x that makes $t = t^*(x)$ optimal for any $t \geq 0$.[39] We will look for an equilibrium where at any $t \geq 0$ someone adopts, so

$$\rho(\alpha(t)\bar{x}(t) - c) = \bar{x}(t)\alpha_t(t)$$

provided α is differentiable at t.[40] The following two lemmas are useful to characterize the solution for the deterministic case. The proof of both lemmas can be found in Appendix **??**.

Lemma 7. *Assume that for $t > 0$ there is some x for which $0 < t^*(x) < \infty$, we denote the smallest of such x as $\bar{x}(t)$. Then if the first and second order necessary conditions holds, then $N(t)$ and $\alpha(t)$ are weakly increasing in time.*

[38] If $\alpha(t)$ is not differentiable at $0 < t = t^*(x) < \infty$:

$$-\rho(\alpha(t^*)x - c) + x\alpha_t^+(t^*) \leq 0 \leq -\rho(\alpha(t^*)x - c) + x\alpha_t^-(t^*)$$

where $\alpha_t^-(t^*)$ and $\alpha_t^+(t^*)$ are the left and right derivatives of $\alpha(t)$ at $t = t^*(x)$. Note that α is differentiable at t provided that $N(t)$ does not jump at t. If $N(t)$ jumps at t, then α has right and left derivatives.

[39] Since in equilibrium $0 \leq N(t) \leq 1$, then $\frac{\theta_0}{\rho} \leq \alpha(t) \leq \frac{\theta_0+\theta_n}{\rho}$ and $0 < \frac{\rho c}{\theta_0+\theta_n} \leq \bar{x}(t) \leq \frac{\rho c}{\theta_0}$. Note that we allow $\bar{x}(t) > U$, but in this case everybody is adopting at t. Thus, we assume that $\frac{(\theta_0+\theta_n)}{\rho} > \frac{c}{U}$, otherwise nobody can ever adopt, and $c > 0$, so that some type will never adopt.

[40] If α is not differentiable at t, then: $\bar{x}(t)\alpha_t^+(t) \leq \rho(\alpha(t)\bar{x}(t) - c) \leq \bar{x}(t)\alpha_t^-(t)$.

The fact that the threshold $\bar{x}(t)$ is weakly decreasing rules out a solution where the number of adopters is decreasing through time.

Lemma 8. Assume that $\bar{x}(t)$ is continuously differentiable with respect to time, that $N(t)$ is weakly increasing in time, and that the initial condition satisfies $M_0(x) \equiv \int_0^x m_0(z)dz \leq F(x)$ for all x. Then, $\bar{x}(t)$ must be decreasing in time, and if in an interval $N(t)$ is strictly decreasing, then $\bar{x}(t)$ must be strictly decreasing. Thus,

$$M(x,t) \equiv \int_0^x m(z,t)dz = \begin{cases} (1-e^{-\nu t})F(x) + e^{-\nu t}M_0(x) & \text{for } x \leq \bar{x}(t) \\ (1-e^{-\nu t})F(\bar{x}(t)) + e^{-\nu t}M_0(\bar{x}(t)) & \text{for } x > \bar{x}(t) \end{cases}$$

The previous lemmas has the following immediate implication.

Lemma 9. Consider the initial condition $m_0(x) = f(x)$ holding for all $x < \bar{x}(0)$. Then: $m(x,t) = f(x)$ and $N(t) = 1 - F(\bar{x}(t))$.

The last lemma states that if we start the economy with a threshold $\bar{x}(0)$ and no agent below that threshold has adopted, then all agents with $x > \bar{x}(0)$ will immediately adopt and the distribution of the non adopters is time invariant. This result is intuitive and it is at the heart of the lack of dynamics in the equilibrium of the deterministic problem.

Assuming Lemma 9 holds and that $m_0(x) = f(x)$, combining the first order conditions with the law of motion for $\alpha(t)$, $\rho\alpha(t) - \alpha_t(t) = \theta_0 + \theta_n(1 - F(\bar{x}(t)))$, we get

$$\bar{x}(t)\left[\theta_0 + \theta_n(1 - F(\bar{x}(t)))\right] - \rho c = 0 \tag{84}$$

Note that in equation (84) the solution for $\bar{x}(t)$ does not depend on t. Thus, we can construct equilibrium where $\bar{x}(0) = \bar{x}(t)$ for all $t \geq 0$. We summarize this result in the following proposition

Proposition 19. Consider the initial condition $m_0(x) = f(x)$ for all $x < \bar{x}(0)$. Then the solution implies a time invariant threshold $\bar{x}(t) = \bar{x}$ solving equation (84), immediate adoption for all agents with $x \geq \bar{x}$, and a time-invariant fraction of adopters $N(t) = N = 1 - F(\bar{x})$.

Note that equation (84) may have multiple solutions. Given that from Lemma 7 we know that $\bar{x}(t)$ is weakly decreasing and $\alpha(t)$ must be strictly increasing in time, then the lower root is the stable solution in the sense that, if the economy is at that point it will remain there forever. We show below that other paths are also possible in the presence of multiple solutions, with the fraction of adopters $N(t)$ ratcheting up at discrete moments in time.

Let us consider the case with two possible stationary equilibria, denoted by $\bar{x}_H > \bar{x}_L$, with associated adoption rates $N_H < N_L$, and the stationary value function $\rho\alpha(t) = \theta_0 + \theta_n N(t)$ (recall $\alpha_t = 0$) with solution $\bar{\alpha}_H = (\theta_0 + \theta_n N_H)/\rho$ and $\bar{\alpha}_L = (\theta_0 + \theta_n N_L)/\rho$, where $\alpha_H < \alpha_L$ since a low threshold yields higher utility due to the larger adoption rate.

We can now check that indeed $t^*(x)$ are optimal for a steady state equilibrium. Since $\alpha(t) = \bar{\alpha}_i$ for $i = L, H$, then the adoption problem becomes:

$$t^*(x) = \arg\max_{t \geq 0} e^{-\rho t} (\bar{\alpha}_i x - c)$$

and the solution is:

$$t^*(x) = \begin{cases} \infty & \text{if } x < \frac{c}{\bar{\alpha}_i} \\ 0 & \text{if } x \geq \frac{c}{\bar{\alpha}_i} \end{cases}$$

Hence, there are no dynamics in the deterministic case. Nonetheless, an equilibrium can be constructed where $\bar{x}(t)$ is piecewise continuous and jumps from $\bar{x}_H$ to $\bar{x}_L$ at some arbitrary time T and where the value of $N(t)$ also jumps. For instance, let $\bar{x}(t) = \bar{x}_H$ for $t \in [0, T)$ and let $\bar{x}(t) = \bar{x}_L$ for $t \in [T, \infty)$, where $T > 0$ is arbitrary. For $t \geq T$, set $\alpha(t) = \bar{\alpha}_L$ and for $t \in [0, T)$, solve $\alpha_t(t) = \rho(\alpha(t) - \alpha_H)$ with boundary condition $\alpha(T) = \bar{\alpha}_L$. This gives

$$\alpha(t) = \bar{\alpha}_H + (\bar{\alpha}_L - \bar{\alpha}_H)\, e^{-\rho(T-t)}$$

Note that $\alpha'(t) > 0$ for $t \in [0, T)$ and $\alpha(0) > \bar{\alpha}_H$. The equilibrium so constructed satisfies the first and second order condition for $t^*(x)$. The following proposition describes such "ratcheting" equilibria:

Proposition 20. Assume that $m_0(x) = f(x)$ for all $x \in [0, U]$. Let $\bar{X}$ be the set of steady state equilibria, i.e.

$$\bar{X} \equiv \{0 < \bar{z}_i \leq U : \rho c = \bar{z}_i [\theta_0 + \theta_n (1 - F(\bar{z}_i))]\}$$

An equilibrium is described by a path $\bar{x}(t)$ that at times $0 = t_0 \leq t_1 < t_2 < t_m < \infty$:

$$\bar{x}(t) = \bar{z}_i \in \bar{X} \text{ for } t_{i-1} \leq t < t_i \text{ for all } i = 1, 2, \ldots, m$$

and where $\bar{z}_i > \bar{z}_{i+1}$ for all $i = 1, 2, \ldots, m$.

In words, an equilibrium is given by a piece-wise constant path for $\bar{x}(t)$, such that at each discontinuity point $\bar{x}(t)$ jumps down to a value that is one of the steady-state solutions of equation (84). The set of equilibria thus includes the fully static one where $\bar{x}(0) = \bar{z}_m$, as well as several other time-varying paths where the elements of $\bar{X}$ (the steady state solutions) and the time sequence t_i are arbitrarily selected subject to the constraint that the path for $\bar{x}(t)$ must be weakly decreasing.

H.1 Proofs of the Deterministic Model

Proof. (of Lemma 7) The necessary second order condition for $0 < t^*(x) < \infty$ is:

$$G_{tt}(t,x)|_{t=t^*(x)} = e^{-\rho t^*}\left(-\rho\alpha_t(t^*)x + \alpha_{tt}(t^*)x\right)$$

Differentiating with respect to time the law of motion for α (i.e., $\rho\alpha(t) = \theta_0 + \theta_n N(t) + \alpha_t(t)$) we have:

$$\rho\alpha_t(t) = \theta_n N_t(t) + \alpha_{tt}(t)$$

Evaluating the second order condition at $(t, \bar{x}(t))$ and replacing $\alpha_{tt}(t)$:

$$\begin{aligned} G_{tt}(t,\bar{x}(t)) &= e^{-\rho t}\left(-\rho\alpha_t(t)\bar{x}(t) + \bar{x}(t)\alpha_{tt}(t)\right) \\ &= e^{-\rho t}\left(-\rho\alpha_t(t)\bar{x}(t) + \bar{x}(t)\left(\rho\alpha_t(t) - \theta_n N_t(t)\right)\right) = -e^{-\rho t}\bar{x}(t)\theta_n N_t(t) \end{aligned}$$

Thus, if the necessary first order condition holds, i.e if $G_{tt}(t,\bar{x}(t)) \leq 0$, then $N_t(t) \geq 0$ and hence it is weakly increasing.

Furthermore, using the first order condition at $t > 0$

$$\rho(\alpha(t)\bar{x}(t) - c) = \alpha_t(t)\bar{x}(t)$$

Note that if $\alpha_t(t)$ is strictly decreasing then $\alpha(t)\bar{x}(t) - c < 0$, which is a contradiction with $\alpha(t)\bar{x}(t) - c \geq 0$. Thus, for a t where α is differentiable (no jump on N), then $\alpha(t)$ must be weakly increasing.

Lastly, notice that

$$\begin{aligned} N(t) &= 1 - \int_0^{\bar{x}(t)} m(z,t)dz \\ &= 1 - \left[M_0(\bar{x}(t))e^{-\nu t} + F(\bar{x}(t))\left(1 - e^{-\nu t}\right)\right] \end{aligned}$$

where the second line uses that $\bar{x}(t)$ is decreasing in time. Taking the derivative of $N(t)$ with respect to time:

$$N_t(t) = \left[m_0(\bar{x}(t))e^{-\nu t} + f(\bar{x}(t))\left(1 - e^{-\nu t}\right)\right]\bar{x}_t(t) - \nu e^{-\nu t}\left[F(\bar{x}(t)) - M_0(\bar{x}(t))\right]$$

□

Proof. (of Lemma 8) The proof has two parts. The first part establishes that $\bar{x}$ is decreasing. The second one uses that property to obtain the law of motion of M. Differentiating the definition of N

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

with respect to t, and using that $m(\cdot,t)$ is zero for $x > \bar{x}(t)$ is in general strictly positive at the left limit $m(\bar{x}(t)^-,t)$, we have:

$$N_t(t) = \bar{x}_t(t)1_{\{\bar{x}_t(t)\le 0\}}m(\bar{x}(t)^-,t) - \int_0^{\bar{x}(t)} m_t(x,t)dx$$

Using the law of motion of m we have:

$$\begin{aligned}
N_t(t) &= \bar{x}_t(t)1_{\{\bar{x}_t(t)\le 0\}}m(\bar{x}(t)^-,t) + \nu\int_0^{\bar{x}(t)}\big(m_(x,t) - f(x)\big)\,dx \\
&= \bar{x}_t(t)1_{\{\bar{x}_t(t)\le 0\}}m(\bar{x}(t)^-,t) + \nu\int_0^{\bar{x}(t)}\big(m_(x,t) - f(x)\big)\,dx \\
&= \bar{x}_t(t)1_{\{\bar{x}_t(t)\le 0\}}m(\bar{x}(t)^-,t) + \nu\left(M(x,t) - F(x,t)\right)
\end{aligned}$$

But since, $M(x,t) \le F(x,t)$ for all x, then if $N_t(t) \ge 0$, then $\bar{x}_t(t) \le 0$.

Now, let use that $\bar{x}(t)$ is decreasing. In this case if $x < \bar{x}(t)$ then it must be that $\bar{s} \le \bar{x}(s)$ for all $s \le t$. Then for such x we have:

$$m_t(x,s) = -\nu\left(m(x,s) - f(x)\right) \text{ for all } s \le t$$

We can solve this o.d.e. for each x, using the boundary $m(x,s) = m_0(x)$. This gives

$$m(x,t) = \begin{cases} (1 - e^{-\nu t})f(x) + e^{-\nu t}m_0(x) & \text{for all } x \le \bar{x}(t) \\ 0 & \text{for all } x > \bar{x}(t) \end{cases}$$

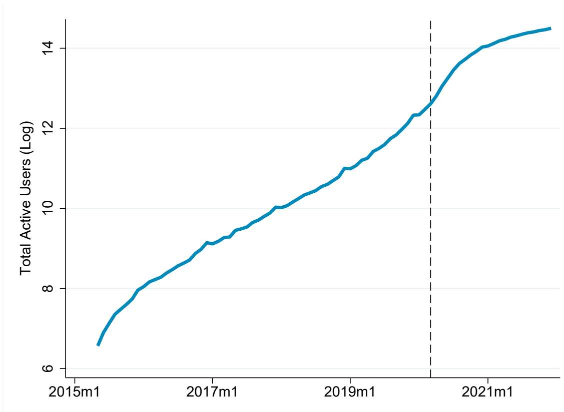
Integrating this density we get the desired result of its CDF $M(x,t)$. □

I Empirical Appendix

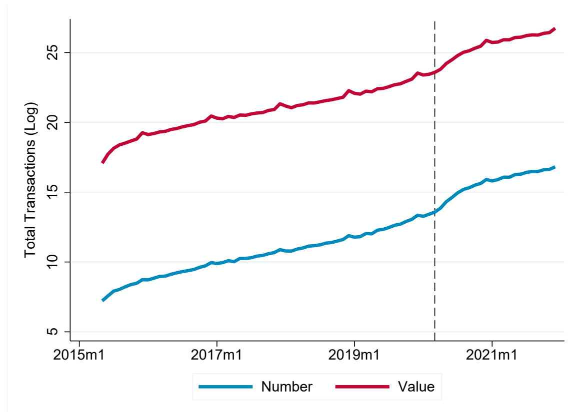
I.1 Descriptive Figures and Summary Statistics: SINPE

The technology diffused slowly. The aggregate adoption of SINPE has grown at a constant rate over time since its inception in 2015, as shown in Figure I1 using monthly data on the total number of adopters.[41] By 2021, close to 79% of the adult population in the country owned a bank account, and over 60% of adults were SINPE subscribers who had not deactivated their account. Moreover, the value of annual transactions in SINPE is approximately 10% of GDP. Thus, this setting has the unique feature of allowing us to study the adoption of mobile payments in the entire population of the country, across many years since the inception of the technology, and until it reached almost the universe of the country's adult population. The fact that adoption occurs gradually coincides with the dynamics of our dynamic stochastic model, and rules out the deterministic case in which adoption happens on impact.

Figure I1: Users, Transactions, and Value of Transactions



(a) Total Users

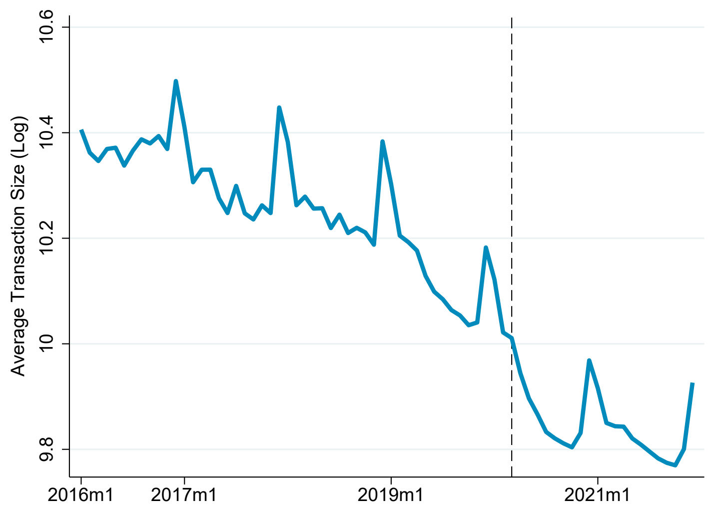


(b) Number and Value of Transactions

Notes: Panel (a) shows total active SINPE users. We include only active subscriptions by individuals, as users have the option of deactivating their account. Panel (b) shows both total transactions in the application and total value of transactions by individuals. Both figures include a vertical dashed line to mark the start of the COVID-19 pandemic (March 2020).

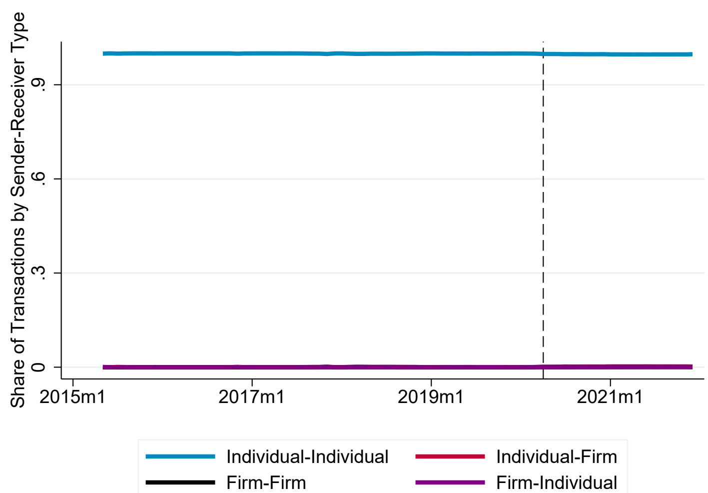
[41] The figures include a vertical dashed line at the beginning of the COVID-19 pandemic (March 2020). As shown, it did not dramatically change the adoption rate.

Figure I2: Average Transaction Size



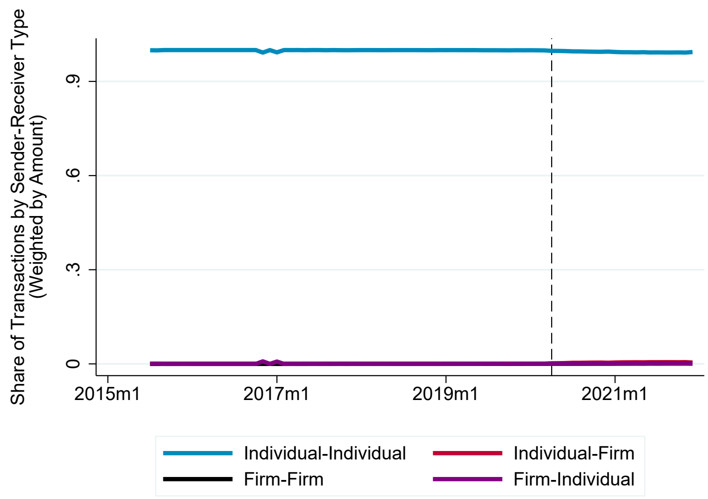
Notes: The figure shows the evolution of the average transaction size in SINPE.

Figure I3: Transactions by Sender-Receiver Type



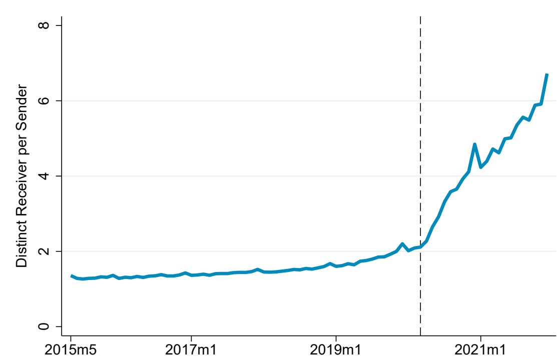
Notes: Transactions are classified according to the type of user. Individuals correspond with Costa Rican adult citizens. Firms correspond with formal enterprises.

Figure I4: Share of Transactions Between Types of Users (Weighted by Amount)

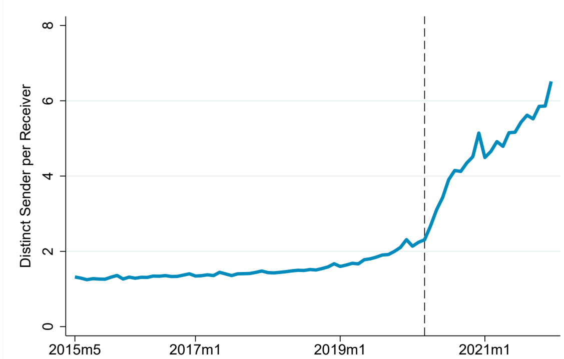


Notes: The figure shows total number of SINPE transactions between four different types of users, as a share of all of their transactions.

Figure I5: Mean Number of Connections per User



(a) Receivers



(b) Senders

Figure I6: Average Age at the Time of Adoption

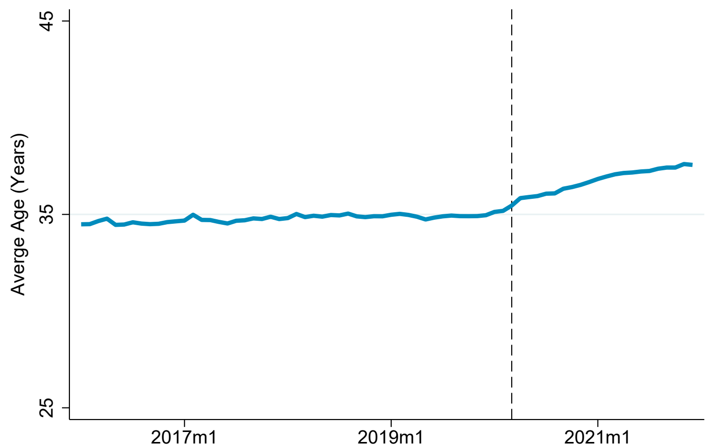


Table I1: Mean Share of Transactions Within Network (2015-2021)

	Neighborhood	Firm	Family	Union of all three
Neighborhood	0.39			
Firm	0.56	0.39		0.65
Family	0.50	0.58	0.25	

Notes: We construct average shares using data from May 2015, when the technology was introduced, to December 2021. Shares using data from the middle of the period (year 2018) only are shown in Table 1.

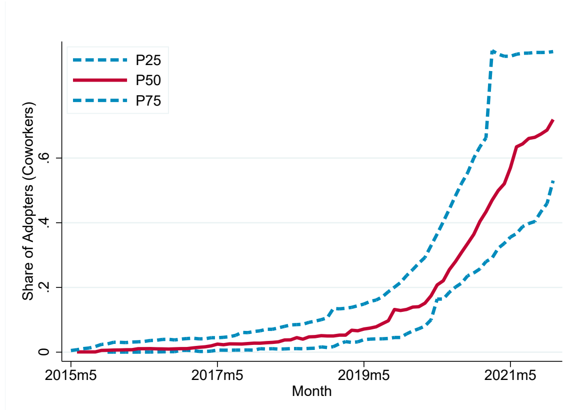
I.2 Evidence on Selection at Entry: Robustness

Table I2: Amount Transacted and Size of Network at Entry

	Dependent variable: Amount transacted (IHS)		
Size of Neighbors' Network at Entry	-5.805***		
	(0.014)		
Size of Coworkers' Network at Entry		-2.663***	
		(0.013)	
Size of Family Network at Entry			-2.077***
			(0.240)
Observations	7,135,126	163,050	6,742,411
R-squared	0.022	0.006	0.003
Network×Time/Cohort FE	Yes	Yes	Yes

Notes: The dependent variable in this estimation is the amount transacted each month for each user, which we transform using the inverse hyperbolic sine function. The coefficient describes the effect of increasing the share of an individual's network who had adopted the app at the time when she downloaded it. We run regressions using data from May 2015, when the technology was introduced, to December 2021.

Figure I7: Gradual Diffusion Across a Networks of Coworkers



(Coworkers)

Notes: The figure shows the dynamics of diffusion across networks defined as r coworkers. Percentiles are calculated in the period with highest adoption in the sample given the share of individuals that had adopted the technology.

I.3 Evidence on Strategic Complementarities: Robustness

Table I3: Changes in Number of Transactions and Network Changes

Dependent variable: Δ Number of Transactions

(a) Logs	(1)	(2)	(3)	(4)
Δ Share Neighborhood Adopters	0.597***			0.470***
	(0.023)			(0.032)
Δ Share Coworkers Adopters		0.205***		0.202***
		(0.007)		(0.007)
Δ (Log) Wag		0.046***		0.046***
		(0.001)		(0.001)
Δ Share Relatives Adopters			0.228***	0.250***
			(0.003)	(0.004)
Observations	24,025,266	12,374,020	22,775,723	11,727,213
Time/Cohort FE	Yes	Yes	Yes	Yes
Adjusted R-squared	0.020	0.025	0.021	0.026
(b) Davis & Haltiwanger				
Δ Share Neighborhood Adopters	1.286***			1.090***
	(0.027)			(0.037)
Δ Share Coworkers Adopters		0.302***		0.293***
		(0.009)		(0.009)
Δ (Log) Wage		0.047***		0.047***
		(0.001)		(0.001)
Δ Share Relatives Adopters			0.305***	0.339***
			(0.004)	(0.005)
Observations	28,160,145	14,311,886	26,663,615	13,549,708
Time/Cohort FE	Yes	Yes	Yes	Yes
Adjusted R-squared	0.014	0.018	0.015	0.019

Notes: The unit of observation is the individual. We run regressions using data from May 2015, when the technology was introduced, to December 2021. Standard errors (clustered by individual) are in parentheses. All regressions control for network size (in levels).

Table I4: Intensity of Usage and Network Changes (Value of Transactions)

Dependent variable: Δ Value of Transactions (IHS)

Δ Share Neighborhood Adopters	3.954***			3.573***
	(0.101)			(0.141)
Δ Share Coworkers Adopters		0.824***		0.796***
		(0.033)		(0.034)
Δ (Log) Wage		0.126***		0.125***
		(0.003)		(0.003)
Δ Share Relatives Adopters			0.867***	0.997***
			(0.014)	(0.019)
Observations	32,391,602	16,232,003	30,633,379	15,355,945
Time/Cohort FE	Yes	Yes	Yes	Yes
Adjusted R-squared	0.008	0.009	0.008	0.009

Notes: The unit of observation is the individual. Regressions control for firm size (in levels). We run regressions using data from May 2015, when the technology was introduced, to December 2021. Standard errors (clustered by individual) are in parentheses.

Table I5: Changes in Intensity of Usage and 2021 Network Changes

Dependent variable: %Δ Number of Transactions

	Logs	Davis & Haltiwanger	Inverse hyperbolic sine
Δ Share Adopters in 2021 Network	1.815***	1.950***	1.580***
	(0.007)	(0.008)	(0.006)
Observations	23,512,962	27,532,941	31,682,276
R-squared	0.022	0.017	0.017
Time/Cohort FE	Yes	Yes	Yes

Notes: The unit of observation is the individual. We run regressions using data from May 2015, when the technology was introduced, to December 2021. These regressions do not control for network size (in levels) as by construction in this exercise size is constant across time. We run regressions using data from May 2015, when the technology was introduced, to December 2021. Standard errors, clustered by individual, are in parentheses.

Table I6: Weighted Changes in Intensity of Usage and 2021 Network Changes)

Dependent variable: %Δ Value of Transactions			
	Logs	Davis & Haltiwanger	Inverse hyperbolic sine
Δ Share Adopters in 2021 Network	5.375***	1.802***	1.978***
	(0.031)	(0.012)	(0.009)
Observations	32,355,374	24,022,540	28,150,990
Time FE/Cohort	Yes	Yes	Yes
Adjusted R-squared	0.009	0.019	0.018

Notes: The unit of observation is the individual. We run regressions using data from May 2015, when the technology was introduced, to December 2021. These regressions do not control for network size (in levels) as by construction in this exercise size is constant across time. Standard errors, clustered by individual, are in parentheses.

Table I7: Leave-One-Out Instrument

First Stage. Dependent Variable: $\Delta N_{it}^{neighborhood}$	
$\Delta N_{-i,t}^{district}$	0.694***
	(0.054)
Observations	32,391,602
Clusters	1,987,052
Time/Cohort FE	Yes
F-statistic	31,717.94
Second Stage. Dependent variable: Δ Number of Transactions (IHS)	
$\Delta N_{-i,t}^{district}$	0.694***
	(0.054)
Observations	32,391,602
Clusters	1,987,052
Time/Cohort FE	Yes

Notes: The unit of observation is the individual. Robust standard errors are in parentheses in the first panel; standard errors clustered by individual are in parentheses in the second panel. Regressions control for network size (in levels).

Table I8: Leave-One-Out Instrument (Value of Transactions)

First Stage. Dependent Variable: $\Delta N_{it}^{neighborhood}$	
$\Delta N_{-i,t}^{district}$	1.596***
	(0.009)
Observations	32,391,602
Clusters	1,987,052
Time/Cohort FE	Yes
F-statistic	31,717.94
Second Stage. Dependent variable: Δ *Number of Transactions (IHS)*	
$\Delta N_{-i,t}^{district}$	0.694***
	(0.054)
Observations	32,391,602
Clusters	1,987,052
Time/Cohort FE	Yes

Notes: The unit of observation is the individual. Robust standard errors are in parentheses in the first panel; standard errors clustered by individual are in parentheses in the second panel. Regressions control for network size (in levels).

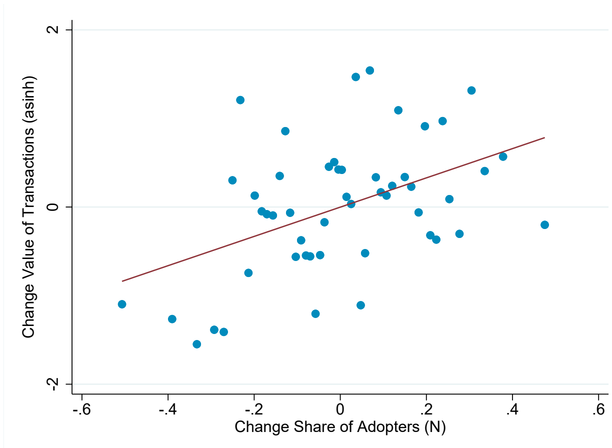
Table I9: Intensity of Usage and Network Changes: Balanced Panel of 2016 Users

Dependent variable: Δ *Number of Transactions (IHS)*				
Δ Share Neighborhood Adopters	0.768***			0.599***
	(0.134)			(0.157)
Δ Share Coworkers Adopters		0.157***		0.139***
		(0.031)		(0.031)
Δ (Log) Wage		0.035***		0.035***
		(0.003)		(0.003)
Δ Share Relatives Adopters			0.441***	0.491***
			(0.020)	(0.025)
Observations	1,073,880	743,321	1,026,384	710,142
Time/Cohort FE	Yes	Yes	Yes	Yes
Adjusted R-squared	0.014	0.016	0.015	0.017

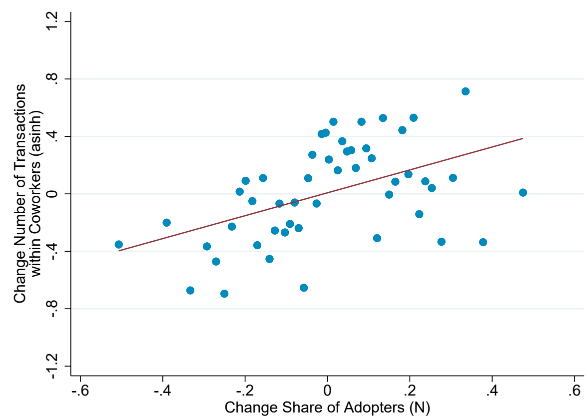
Notes: The unit of observation is the individual, and we consider only a subsample of users who had already adopted by 2016, and follow them from then onwards, until December 2021. Regressions control for firm size (in levels). Standard errors (clustered by individual) are in parentheses.

I.4 Mass Layoffs and Adoption Changes: Robustness and Details

Figure I8: Marginal Effect of Network Changes on Usage Intensity (Value of Transactions)



(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 value 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 value of transactions *which have a coworker as a counterpart* (inverse hyperbolic sine function) on each period for each user.

Table I10: Intensity of Usage and Changes in Coworkers' Network After a Mass Layoff (Value of Transactions)

Dependent Variable: Δ Value of transactions (IHS)

$\Delta N_i^{coworkers}$	4.879***	3.134***	2.075***	1.519**
	(0.509)	(0.552)	(0.588)	(0.594)
$\Delta \ln wage_i$		0.833***	0.736***	0.866***
		(0.157)	(0.157)	(0.166)
$\Delta Covid_i$			0.317***	0.238***
			(0.065)	(0.076)
Observations	1,554	1,554	1,554	1,554
Time FE	No	Yes	Yes	Yes
Cohort FE	No	No	No	Yes
Adjusted R-squared	0.063	0.116	0.130	0.170

Notes: 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, until December 2021. All regressions control for network size (in levels). Standard errors are in parentheses.

Table I11: Intensity of Usage *Among Coworkers* and Changes in Coworkers' Network After a Mass Layoff

Dependent Variable: Δ Number of transactions (IHS)

$\Delta N_i^{coworkers}$	1.338***	0.776***	0.690***	0.640***
	(0.138)	(0.172)	(0.185)	(0.192)
$\Delta \ln wage_i$		0.361***	0.352***	0.367***
		(0.041)	(0.042)	(0.046)
$\Delta Covid_i$			0.029	0.034
			(0.020)	(0.024)
Observations	1,554	1,554	1,554	1,554
Time FE	No	Yes	Yes	Yes
Cohort FE	No	No	No	Yes
Adjusted R-squared	0.057	0.099	0.099	0.086

Notes: 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, until December 2021. Standard errors are in parentheses.

Table I12: Intensity of Usage *Among Coworkers* and Changes in Coworkers' Network After a Mass Layoff (Value of Transactions)

Dependent Variable: Δ Value of transactions (IHS)

$\Delta N_i^{coworkers}$	5.639***	4.235***	3.763***	3.564***
	(0.584)	(0.663)	(0.728)	(0.772)
$\Delta \ln wage_i$		1.628***	1.585***	1.589***
		(0.176)	(0.178)	(0.190)
$\Delta Covid_i$			0.141	0.198*
			(0.088)	(0.101)
Observations	1,554	1,554	1,554	1,554
Time FE	No	Yes	Yes	Yes
Cohort FE	No	No	No	Yes
Adjusted R-squared	0.058	0.108	0.109	0.100

Notes: 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, until December 2021. All regressions control for network size (in levels). Standard errors are in parentheses.

I.4.1 Details on Mass Layoffs

This section provides additional details on the choices made to construct the variables and sample used in Section 7.3

Definition of a Mass Layoff 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* which did not recover in the following 12 months. We define a recovery as a firm which went back to its initial size (or above) within the following 12 months. Given this definition, the descriptive statistics of firms and workers impacted by a mass layoff are reported in Table I13.

Table I13: Mass Layoffs: Descriptive Statistics

Number of firms	856	
Number of displaced workers who had not adopted SINPE when fired	32,620	
Number of displaced workers who had adopted SINPE when fired	2,585	
Average firm size	529	(2147)
Average monthly wage pre-layoff, laid-off workers	\$504	(\$623)
Average monthly wage pre-layoff, all workers	\$663	(\$487)

Notes: Standard deviations for mean variables are reported in parenthesis. We consider layoffs that reduce in 30 workers or more the size of firms with at least 50 workers, and limit the analysis to workers with a period of unemployment of 6 months or less. Wages were calculated based on an exchange rate of 634 colones per dollar and the last month in which workers were employed. We include mass layoffs which occurred between May 2015, when the technology was introduced, and December 2021. The last row includes the average monthly wage pre-layoff for all workers who were employed at those firms at the time of the mass layoff.

Definition of Variables We construct several variables that are used in equation (37). We now provide more details on each of them.

- $Adopt_i$ equals one if individual i adopted SINPE within 6 months after arriving to her new firm, and zero otherwise. This variable is only computed for individuals who found a job within 6 months of being fired. Results are robust to considering shorter unemployment spells, including conducting the analysis using only job-to-job transitions.

- $\Delta N_i^{coworkers}$ is the change between the share of coworkers who had adopted at the old and the new employer. We compute this variable by calculating the difference between (i) the share of adopters at the old firm on the last month in which the individual was employed and (ii) the share of adopters at the new firm in month i, and considering only months i after the individual was hired at the new firm.

- $\Delta \ln wage_i$ 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 levels) at the new firm versus the old firm.

- $date\,hired_i$ controls for the month 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 prior to the move, which aims to control for factors, other than strategic complementarities, which might facilitate adoption at the new vs. the old firm.

- $\Delta Covid_i$ 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. This change is zero for pre-pandemic years, thus, this variable is introduced using an inverse hyperbolic sine transformation, as opposed to a logarithm.

The regression described in equation (38) relies on the same variables that we described above, but also includes additional ones which we now describe.

- $\Delta \ln \tilde{\xi}_i$ 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. We only compute this variable for workers who had adopted SINPE more than 6 months before being fired, in order to attenuate any effect coming from a "learning curve." We transform $\tilde{\xi}_i$ using the inverse hyperbolic sine function, as zeros are common in the monthly data. Note that this inflates coefficients, particularly, for large values of intensity, which are likely to appear when the left-hand-side variable describes the total value (as opposed to the number) of transactions.

- $cohort_i$ controls for the month when individual i adopted SINPE. We include this variable to attenuate any effect coming from learning how to better use the app.

- $\ln \sum^t \tilde{\xi}_i$ is the sum of all historical transactions made by agent i since she adopted the app. This variable has no zeros by construction, as our definition of adoption is that the individual has used the app at least once. Similarly to $cohort_i$, the variable intends to control for learning how to use the app thanks to having more people in your network who have adopted it.

J Quantitative Exercises

J.1 Calibration

In this section, we describe the procedure to calibrate σ and θ_0 using simulated method of moments. Intuitively, we want to choose these two parameters so that they are consistent with the distribution of transactions in the data. To do so, in the data, we focus on a balanced sample of users that were active by 2019, and compute moments for the distribution of transactions over the years 2020-2021 in neighborhoods close to steady state (i.e., the top 5 percentile of neighborhoods in terms of adoption). We simulate the model replicating the same characteristics of our empirical sample. This is, we start from steady state and simulate a panel of 5000 users for two years.[42] We then compute the distribution of transactions both in the data and the model and choose σ and θ_0 to minimize the distance between the two distributions. We provide further details of this procedure below.

We begin by simulating the model for a panel of agents for different values of σ and θ_0. Our simulation takes as given the values of ν, r, ρ, β_0, and $\widetilde{\theta}$, since they are calibrated either externally or using other reduced form evidence. Initial conditions $x(0)$ are drawn from the steady state distribution of adopters. To find this distribution, we first find $\bar{x}$ (given $N_{ss} = 0.93$) using the following equation:

$$N_{ss} = (1 - \tfrac{\nu}{\beta_0}) \left[1 - \tfrac{\bar{x}}{U} \left(1 - \frac{\tanh(\gamma \bar{x})}{\gamma \bar{x}}\right)\right].$$

Then, given $\bar{x}$, we find the distribution of adopters using the steady state distribution of non-adopters:

$$\tilde{m}(x) = (1 - \tfrac{\nu}{\beta_0}) \tfrac{1}{U} \left(1 - \frac{\cosh(\gamma x)}{\cosh(\gamma \bar{x})}\right) \text{ where } \gamma = \sqrt{2\nu}/\sigma$$

using that $N_{ss} = I_{ss} - M_{ss}$ and $I_{ss} = (1 - \frac{\nu}{\beta_0})$. In the simulation, agents die at rate ν and they become inactive in the application just as in the data. The process of x follows a Brownian motion, independent across agents, with variance per unit of time σ, no drift, and reflecting barriers at $x = 0$ and $x = U$. We then interpret the flow benefit of agents who adopt the technology as being proportional to how intensively they use SINPE. Thus, we compute

$$\xi_t = [\theta_0 (1 + \widetilde{\theta} N_{ss}) x_t]^{\frac{1}{1+p}}. \tag{85}$$

Given the discreteness of the number of transactions in the data, ξ_t is interpreted as the mean of a Poisson distribution; transactions each period are drawn from a Poisson probability distribution with mean ξ_t.

Parameters σ and θ_0 are chosen to minimize a sum of the percent deviations of simulated

[42]Our estimates are not sensitive to simulating a larger sample of users.

moments from target moments:

$$\min \sum_{i}^{4} \frac{|\text{Model}(i) - \text{Data}(i)|}{\text{Data}(i)}$$

where Model(i) is a simulated i-th moment and Data(i) is a target value of i-th moment. Table J1 reports the empirical and simulated moments.

Table J1: Moments: Distribution of Transactions

Moment	Data	Model
Mean Number of Transactions	8.76	8.69
Median Number of Transactions	7.97	8.48
Absolute Value Changes in Transactions	3.87	3.22
IQR Changes in Transactions	4.56	4.00

Intuitively, equation (85) implies

$$\sqrt{\text{Var}\left(\frac{\xi_{t+\Delta}^{1+p}}{1+\widetilde{\theta} N_{ss}} - \frac{\xi_t^{1+p}}{1+\widetilde{\theta} N_{ss}}\right)} = \theta_0 \sqrt{\Delta} \sigma.$$

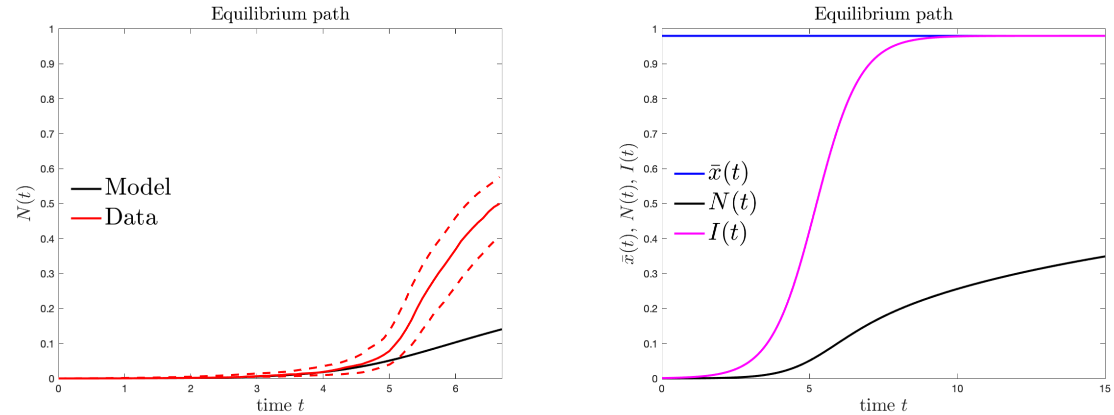
Thus, the dispersion in the changes of transactions contains relevant information to pin down σ. Similarly, the average number of transactions provides information relevant to pin down θ_0. This can be seen taking expectations of equation (85)

$$\mathbb{E}(\xi^{1+p}) = \theta_0 (1 + \widetilde{\theta} N_{ss}) \mathbb{E}(x).$$

J.2 Only Learning: $\widetilde{\theta} = 0$

In this section, we examine the behavior of a model without strategic complementarities. Not surprisingly, if we keep all parameter at their baseline value and set $\theta_n = 0$, the model predicts lower adoption in steady state, $N_{ss} = 0.59$. The adoption in this model is purely determined by the idiosyncratic benefits of the technology. Panel (b) shows that convergence to steady state takes longer in a model without complementarities. Recall that the model matches the fraction of agents informed about the technology three years after it was launched. Panel (b) suggests that in a pure learning model, adoption would be much slower than that observed in the data. Panel (b) also shows that the path of $\bar{x}(t)$ in the model with only learning is flat, which indicates there is no selection in the adoption of the technology as observed in the data. Importantly, this version of the model is constrained efficient: the optimal subsidy to use the technology is zero.

Figure J1: Path of Adopters - Only Learning (Short-Run and Long-Run)



(a) Model vs Data

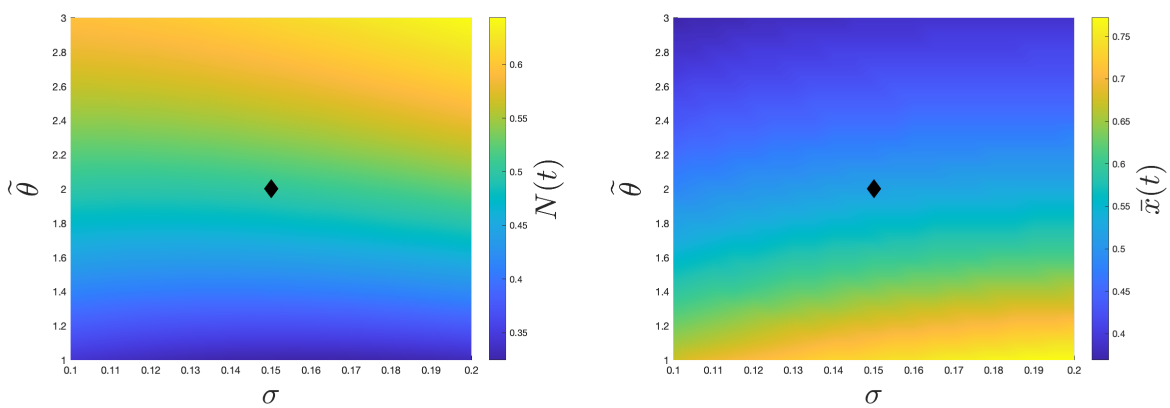
(b) Long-Run Path

Notes: Panel(a) compares the path of adopters in the model with $\theta_n = 0$ 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 but setting $\theta_n = 0$.

J.3 Comparative Statics

J.3.1 Stochastic Model: Short-Run

Figure J2: Adoption: $N(t)$ and $\bar{x}(t)$



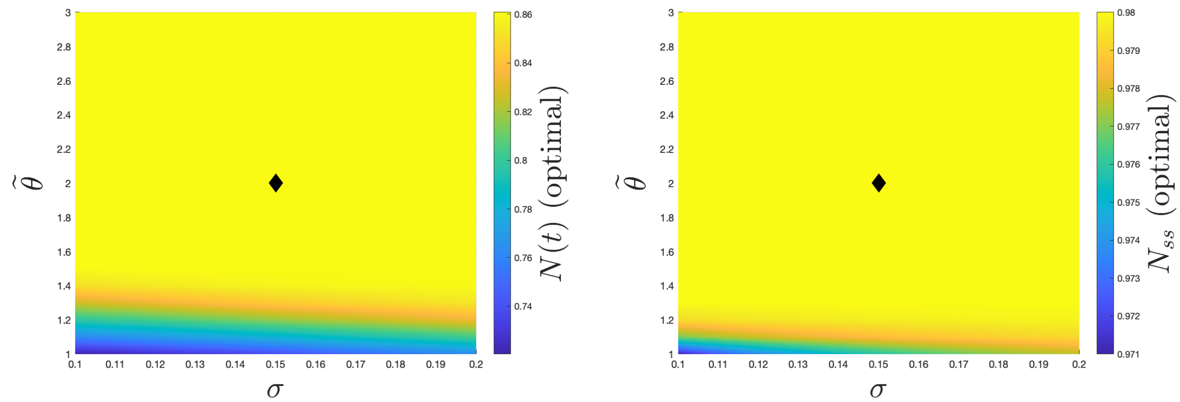
(a) Comparative Statics: $N(t)$

(b) Comparative Statics: $\bar{x}(t)$

Notes: Panel (a) and (b) show how $N(t)$ and $\bar{x}(t)$ change with $\widetilde{\theta}$ and σ, keeping the rest of the parameters constant 7 years after the technology was launched. The black diamonds indicate the levels of $\widetilde{\theta}$ and σ in our baseline calibration.

J.3.2 Stochastic Model: Planning Problem

Figure J3: Optimal Adoption: $N(t)$ and N_{ss}



(a) Comparative Statics: $N(t)$ (b) Comparative Statics: N_{ss}

Notes: Panel (a) shows how $N(t)$ changes 7 years after the technology was launched with $\widetilde{\theta}$ and σ, keeping the rest of the parameters constant. The black diamonds indicate the levels of $\widetilde{\theta}$ and σ in our baseline calibration. Panel (b) shows the same comparative static for N_{ss}.