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

February 2023
PRELIMINARY AND INCOMPLETE

Abstract

This paper develops a dynamic model of technology adoption featuring strategic complementarities: the benefits of usage increase with the number of adopters. Such an effect is inherent to several technologies, such as means of payments. We show that complementarities give rise to multiple equilibrium paths, multiple steady states, and suboptimal allocations. The model generates slow 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

*First draft: February 15, 2022. We thank Isaac Baley, Andres Blanco, Diego Comin (discussant), Joe Hazell, Rishabh Kirpalani, John Leahy (discussant), Claudio Michelacci, Ben Moll, Guillermo Noguera, Ezra Oberfield, Franck Portier, Rob Shimer, Nancy Stokey, Daniel Xu, Yu Zhu (discussant), 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, Bank of Canada, UCSB LAEF, 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. 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 technology adoption to study the role of such complementarities in the diffusion of a technology. The model allows us to analyze the efficiency of the equilibria and discuss optimal policy interventions.

In particular, we focus on the diffusion of new means of payments, such as mobile money and other peer-to-peer (P2P) payment instruments, which have been recently propelled by digitization (see e.g. Economides and Jeziorski (2017); Aron (2018)) and which have also appeared in several plans for central bank digital currency (see e.g. Auer et al. (2020); Carapella and Flemming (2020)). A central feature of our analysis is the presence of complementarities in adoption, which are an inherent property of payment instruments: the benefits of the payment instrument are larger if more people use it.

While the applied literature studying 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)), progress in this research area has been hindered by the challenges that arise when modeling adoption dynamically—a large state space, non-linear decisions, multiple steady states, and multiple equilibria, and by the lack of detailed data on technology diffusion. We develop a simple model featuring complementarities and fully fledged dynamic decisions: the current decision to adopt depends on the whole path of future adoptions, and transition paths can be characterized in closed form. The model embraces the possibility of multiple equilibria as well as multiple steady states. We discuss equilibrium existence and analyze its local 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; moreover, over 10% of the country's GDP is transacted through SINPE. This is a pertinent application of the theory because payment technologies intuitively feature strong network complementarities. In fact, we leverage a battery of granular administrative datasets to characterize adoption patterns

[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).

and to document the presence of strong complementarities in adoption.[2]

The model assumes the flow benefits of using the technology at time t depend on the number of agents who have already adopted the technology, $N(t)$, and on an idiosyncratic persistent random component, $x(t)$. Adoption entails a fixed cost and agents choose when to adopt taking the aggregate path of adoption as given. The model also includes an intensive margin for the usage of the technology. We show that when the idiosyncratic benefits are random the equilibrium features gradual adoption through a simple mechanism: agents wait for others to adopt.[3] This differs from previous contributions, discussed below, where gradualism is either absent, exogenously assumed (e.g., by means of staggered adoption opportunities), or due to learning. While gradualism could also be generated (or amplified) by a learning mechanism, we see strategic externalities as a key 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.[5] Second, we obtain a comparative static result with respect to the initial measure of adopters and the strength of the strategic complementarities. We show that there is a critical mass $\underline{N}$ such that, if the initial measure of adopters is below $\underline{N}$, then there is an equilibrium where no one will adopt in the future. Third, we show that besides the steady state with no adoption the model has two additional interior steady states, which we label low- and high-adoption steady states.

Fourth, we conduct a perturbation analysis with respect to the initial condition to study the stability of the interior steady states.[6] We find that the low-adoption steady state is

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

[3]We also analyze a model where x is heterogeneous across agents but fixed through time, mostly to compare with the existing literature. 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 slow 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]As expected, given the strategic complementarities, 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.

[6]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

locally unstable, while the high-adoption equilibrium is locally stable. Given these results, we focus on equilibria with either no activity or high activity in the steady state.

Fifth, since all equilibria are socially inefficient, we solve the planner's problem. This is a non-trivial problem that involves controlling the entire distribution of agents across time. We decentralize the planner's solution using a time-varying subsidy paid to those that use the technology. The optimal subsidy corrects the network externality, and if the initial condition has lower adoption than the steady steady of the planner problem, the subsidy is increasing across time.

We then use data on the diffusion of electronic payment methods and user networks to quantify these strategic complementarities. In particular, we leverage detailed data from SINPE, an app that connects users in Costa Rica and allows them to mobilize funds between their bank accounts. Information on all SINPE transactions has been collected since its inception in May 2015, which allows us to analyze the dynamics of adoption of this system of national payments in great detail. In turn, data on users—both receivers and senders—can be linked to several relevant networks, including the employer-employee network, family networks, and spatial networks to explore the role of neighbors.

The platform data allow us to test several model predictions. We document five new empirical facts which align with our model. First, we find the technology diffused slowly; while by 2021 over 60% of the adult population had adopted SINPE, adoption has grown at a constant rate since the app's launching. Second, most transactions are peer-to-peer; while firms can potentially use SINPE, over 90% of the transactions are between individuals, which aligns with a model like ours, where small agents trade with each other, rather than one with a few non-atomistic players (large firms). Third, individuals "belong" to networks; 75% of all transactions occur between coworkers, neighbors, or relatives. Fourth, there is evidence of selection at entry: we find that 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 align with our model where individuals with a high 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. We see these facts as consistent with the key assumptions of our model where agents are heterogenous and network effects are important.

We then proceed with a strategy to calibrate the model. Our quantitative analysis combines a model of strategic complementarities with a random diffusion of information model following the seminal work of Bass (1969). The calibration requires us to estimate the value

Souganidis (2022b).

of the parameter that governs the strength of the strategic complementarities. We do so by exploiting exogenous changes in the network of coworkers after a mass layoff. In particular, we examine how the share of coworkers who has adopted changes as someone moves from one firm to another one for plausibly exogenous reasons. This strategy allows us to 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. We examine how both the extensive and the intensive margin of adoption respond. The intensive margin, in particular, aids us in teasing out strategic complementarities from other channels. We calibrate other parameters in the model 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 to the literature. In contrast to the previous literature, which has studied *deterministic* problems (e.g. Stokey 2020), our model allows for both stochastic network connections and an initial arbitrary path of the distribution of adopters. Our theoretical approach has three advantages relative to the previous literature studying adoption *dynamically*. First, it allows for dynamics in technology adoption as observed in the data. In the stochastic version of the model, our model has dynamics even without the inclusion of frictions to delay the transition to steady state. Second, our model allows for multiple equilibria (e.g. Cabral 1990). 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. Lastly, and more importantly, casting the problem as a Mean Field Game allows us to solve the planning problem. This is relevant for policy since the presence of network externalities implies that the solution of the decentralized problem is not efficient. The solution of this problem is non-trivial since the planner needs to account for the law of motion of the density of non-adopters at each point of the state space and each time period. Our framework allows us to compute the optimal subsidy, which equates the solution of the decentralized problem with that of the planner and depends positively on the importance of strategic complementarities, which can be estimated directly in the data. Relevant studies include Benhabib et al. (2021), who model firms that can endogenously innovate and adopt at technology and the effect of these choices on productivity and balanced growth, but without conducting an analysis of the transitions between steady states; Crouzet et al. (2023), who develop a model with a unique equilibrium where the rate of adoption increases given a shock due to complementarities and where dynamics come from a sluggish adjustment a la Calvo (1983); and Buera et al. (2021), who study policies that can coordinate technology adoption across firms.

On the empirical front, SINPE data spans from 2015 to 2021. During this time, it went from zero adoption to over 60% of *national* adoption. These features allow us to study the general equilibrium effects of adoption across a long time period and complements previous studies, summarized by Suri (2017), that have relied on RCTs or shorter periods of time to analyze the patters of adoption of electronic methods of payment. In contrast to other large-scale studies, of which the closest one to our work is Crouzet et al. (2023), who rely on variation in the intensity with which Indian districts were exposed to the cash contraction induced by the 2016 Indian Demonetization, we are able to use individual-level data on adoption and SINPE usage, on individual earnings and demographic characteristics, and on each person's network of SINPE users, relatives, neighbors, and coworkers. This provides an opportunity to understand the characteristics and relevant networks of each user, identify the strength of complementarities and how they vary across networks and time—at both the extensive and intensive margins—and the dynamics of adoption over a long time period.

2 Model setup

We setup a model to study 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 2 in Appendix A for a detailed definition.

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

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

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

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

2.2 Aggregation

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

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

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

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

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

The second term on the right hand side is the fraction of agents who did not have the technology at time 0 and survived until time t without adopting. The third term considers the cohorts of agents that are born between 0 and t, and for each of these cohorts computes the fraction that survived without adopting up to t. We note that an equivalent version of equation (5) holds in a discretized version of the model.

We denote the resulting path of N as a function of $\bar{x}$ (the path of the adoption threshold) and of the initial condition m_0, namely $N = \mathcal{N}(\bar{x}; m_0)$.

2.3 Equilibrium

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

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

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

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

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

A definition of equilibrium equivalent to Definition 1 can be obtained as the four functions $\{D, m, \bar{x}, N\}$ satisfying the coupled of p.d.e.'s for D and m, and the respective boundary conditions, given by equation (9), equation (10), equation (11) and equation (12). This is the typical definition used in the Mean Field Game literature. 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 free boundary (for every period), akin to the Stefan problem 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 2 in Appendix A).[7] We have:

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

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

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).

4 No adoption Equilibrium

In this section we analyze the equilibrium in which there is no adoption i.e. $\bar{x}(t) = U$ for all t. To simplify 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 one stays 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.

5 Steady states

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.

5.1 Steady states 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 steady state 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 then 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

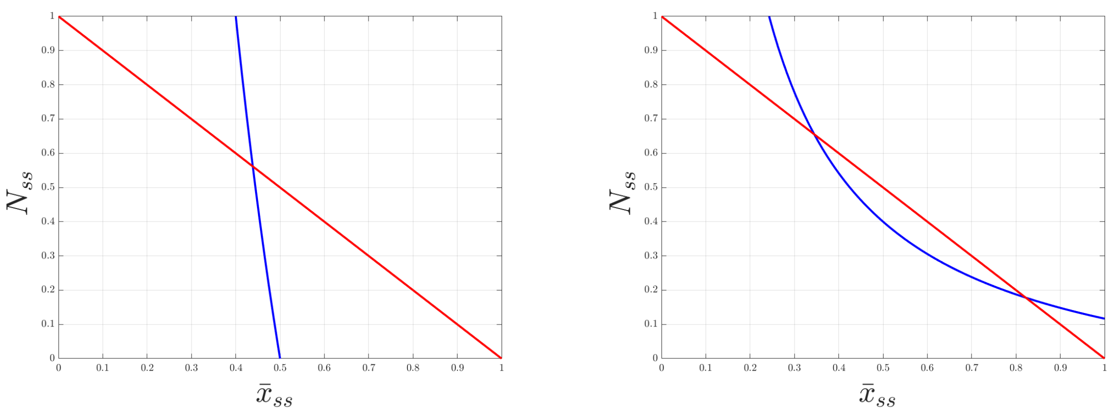
$$N_{ss} = 1 - \frac{\bar{x}_{ss}}{U} \tag{18}$$

Figure 1 plots the non-linear equation (17) and the linear equation equation (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:

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.

In words, multiple interior steady states occur when the complementarities are large relative to the intrinsic value of the technology, i.e. when θ_0 is small and θ_n is large.

Figure 1: Deterministic steady state solution



(i) $\rho c < \theta_0 U$ (ii) $\theta_0 U < \rho c$ and θ_n 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. 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

$N(t) = N_{ss}$. Interestingly, the stochastic version of the model will instead feature dynamics, namely a gradual adoption of the technology so that $N(t)$ is increasing through time.

5.2 Steady states in the stochastic model ($\sigma > 0$)

A steady state 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 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

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

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 it is given by:

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

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

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

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

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: $\mathcal{N}_{ss}(\bar{x}) = 1 - \frac{\bar{x}_{ss}}{U} + \frac{\sigma}{U\sqrt{2\nu}} + o(\sigma)$.

Thus, together equation (19) and equation (20) determine $\bar{x}_{ss}$ and N_{ss}. In particular, a steady state, described by the pair $\bar{x}_{ss}, N_{ss}$, which solve

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

Next we summarize the behaviour 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 $\bar{x}^H < \bar{x}^L$.

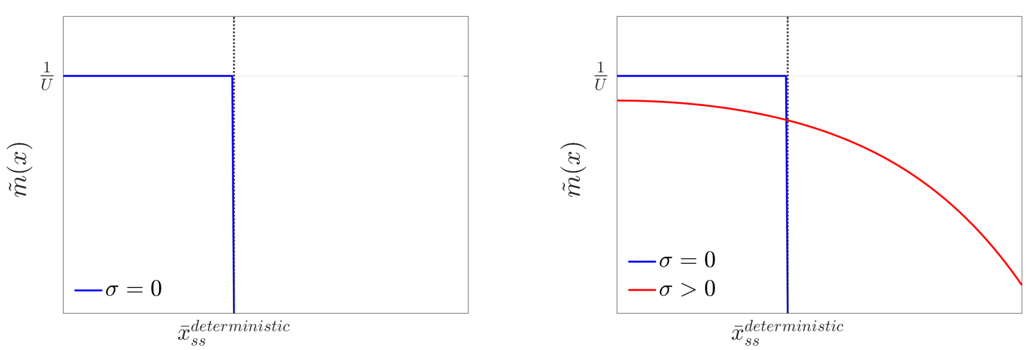
Proposition 6. Assume that $\nu > 0$ and that the parameters θ_0, θ_n, c and ρ are such that there are two interior steady states in the deterministic case of $\sigma = 0$, and label them as $\bar{x}_{ss}^H < \bar{x}_{ss}^L$. Then, (i) there exists a $\bar{\sigma} > 0$ such that for all $\sigma \in (0, \bar{\sigma})$ 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.

Importantly, the last term of equation (20) shows that there is a smaller density of non-adopters for $x \in [0, \bar{x}_{ss}]$ in the stochastic case relative to the deterministic case. This observations is key to understand why the stochastic case has dynamics. Panel (a) of Figure 2 shows the density of non-adopters around $\bar{x}$ for $\sigma \approx 0$. In this case, the density of adopters is very close to zero so that the distribution looks almost uniform, as in the deterministic case. Panel (b) shows the same figure for $\sigma > 0$. First, $\bar{x}_{ss}$ is larger due to the option value that is present in the stochastic model. More importantly, for $\sigma > 0$, the density of adopters below $\bar{x}_{ss}$ is non-zero, as a result the density of non-adopters is not uniform. The key novelty of the stochastic model is that there are agents with $x(t) < \bar{x}(t)$ 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 x moved down in time. As a result, $m(x) < 1/U$ when $\sigma > 0$. Given that the density takes time to adjust, the stochastic model features the presence of dynamics in the adoption of new technology as the optimal value of $\bar{x}$ is not independent of time.

Figure 2: Stochastic Steady State: Density of non-adopters: $m(x)$



(a) $\sigma \approx 0$ (b) $\sigma > 0$

6 Perturbation and stability of equilibrium steady states

In this section we analyze the stability of the steady states. We explore the question using a perturbation of the distribution of adopters in each of the two interior steady states. The analysis uses techniques from the Mean Field Game literature developed in Alvarez, Lippi and Souganidis (2022b). The analysis allows us to approximate $\mathcal{X}$ and $\mathcal{N}$ around the steady state and to inspect the local stability of the equilibrium.

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 produces a new p.d.e., and 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 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

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

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 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 $n_0(\omega)(t)$. Alternatively, $n_0(\omega)(t)$ is the effect at time t on the path $N(t)$ of 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 limit case of perturbation ω given by (the limit) of 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 second term in equation (22) contains the effect of the perturbation y on the path of the threshold, $\bar{x}(s)$. Alternatively, 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 fixed. Note also that, consistent with our general result for $\mathcal{N}$, the effect of the thresholds is negative, because $J > 0$ and $\tilde{m}_x(\bar{x}_{ss}) < 0$.

For future reference it is useful to understand the behaviour 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 $\mu_0 = \nu + \frac{\sigma^2}{8}\left(\frac{\pi}{\bar{x}_{ss}}\right)^2$ which is the dominant eigenvalue, which correspond to 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 strategy of the proof is similar to the one outlined for the previous proposition and is given in Appendix D.2.

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 (73) 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

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

where $n_0(\omega)(t)$ is given in Proposition 8 and $\Theta(\bar{x}_{ss}) \equiv \frac{\tilde{m}_x(\bar{x}_{ss})\sigma^2\theta_n}{\bar{x}_{ss}\tilde{D}_{xx}(\bar{x}_{ss})} > 0$. 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, $\text{Lip}_K \equiv \sup_t \int |K(t,s)|ds \le \left(\frac{\bar{x}_{ss}^2}{\sigma^2}\right)^2$. Furthermore, if $\Theta(\bar{x}_{ss})\,\text{Lip}_K < 1$ there exists a unique bounded solution to equation (27) which is the limit of

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

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

This theorem gives a linear system of equations that the perturbation of the equilibrium around a steady state must satisfy, as well as a partial characterisation of its solution. The proof can be found in Appendix D.3.

7 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 subsidies. Section 7.1 characterizes the steady state of this problem. Section 7.2 uses a linearized version of the problem to analyze dynamics around the steady state.

The planner solves a non-trivial dynamic problem, since it has as its state the entire distribution. At time zero the planner solves:

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

subject to

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

The objective function of the planner integrates the lifetime utility of agents using as a weight the discount factor e^{-rt} for the cohort born at t. The first term contains the utility flows of all those using the technology. The second term subtracts the cost of adoption across time, where $N_t(t) + \nu N(t)$ is the gross cost of adoption at time t. The planner decides at each time a threshold $\bar{x}(t)$ which determines adoption, and takes as given the initial condition $m_0(x)$. The planner takes as given the law of motion of the density m that is only affected through the choice of $\bar{x}$. The first constraint defines $N(t)$, second 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 form a 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 discrete-time discrete-state using a finite-difference approximation. In this set up we allow 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 F.2. The p.d.e's corresponding to the planning problem are summarized in the following proposition.

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

$$\begin{aligned} \rho\lambda(x,t) &= x\big(\theta_0 + \theta_n[1 - \int_0^{\bar{x}(t)} m(z,t)dz]\big) + \theta_n\big(\tfrac{U}{2} - \int_0^{\bar{x}(t)} m(z,t)z\,dz\big) \qquad (29) \\ &\quad + \tfrac{\sigma^2}{2}\lambda_{xx}(x,t) + \lambda_t(x,t) \text{ for } x \leq \bar{x}(t) \text{ and } t \geq 0 \\ \lambda(x,t) &= c \text{ for } x \geq \bar{x}(t) \text{ and } t \geq 0 \\ \lambda_x(\bar{x}(t),t) &= 0 \text{ for } t \geq 0 \qquad (30) \\ \lambda_x(0,t) &= 0 \text{ for } t \geq 0 \end{aligned}$$

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 equilibrium. Second, it indicates how to decentralise the optimal allocation as an equilibrium. Let $Z(t) \equiv \frac{U}{2} - \int_0^{\bar{x}(t)} m(x,t)x\,dx$. Comparing the p.d.e. the Lagrange multiplier λ with the p.d.e. for D, which characterizes the equilibrium, we see that the two 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 will 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 coincide with an equilibrium for the same initial conditions with a time varying subsidy paid to adopters. The flow subsidy paid at time t to those that have adopted at t or before is given by $\theta_n Z(t)$ where

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

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

For future reference we define as $Z = \mathcal{Z}(\bar{x}; m_0)$ as the solution of the path for Z as defined in equation (31). In particular, given $\bar{x}$ and m_0, using the KFE one solves for the path of m, and computing the integral in equation (31) gives Z.

Consider the path $\bar{x}$ that solves the p.d.e. $\rho\lambda(x,t) = x\,(\theta_0 + \theta_n N(t)) + \theta_n Z(t) + \frac{\sigma^2}{2}\lambda_{xx}(x,t) + \lambda_t(x,t)$ with the three boundaries given in equation (30) given the paths of N and Z and terminal condition $\lambda(x,T) = \lambda_T(x)$. For future reference, we define $\bar{x} = \mathcal{X}^P(N, Z; \lambda_T)$ to denote the functional, which is defined as the $\mathcal{X}$ in Section 2.1 and where the superscript P denotes the planning problem.

Note that, using the definitions for $\mathcal{X}^P, \mathcal{Z}$ and $\mathcal{N}$ the planner's problem must satisfy the fixed point $\bar{x}^* = \mathcal{H}(\bar{x}^*, \lambda_T, m_0)$ where $\mathcal{H}(\bar{x}; \lambda_T, m_0) \equiv \mathcal{X}^P(\mathcal{N}(\bar{x}; m_0), \mathcal{Z}(\bar{x}; m_0); \lambda_T)$. We can use the same type of analysis, based on monotonicity, to characterize the solution to this fixed point problem, and to compute it. To simplify we omit this analysis.

We turn next to the description of the steady state of the planning problem.

7.1 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 7. 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 F.3. 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 decentralized version. This proposition can be used to conclude that the steady state level of adoption in the planning problem is higher that the high-activity steady state in the equilibrium.

7.2 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 6.

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

$$\begin{aligned} \bar{y}(t) &\equiv \lim_{\epsilon \downarrow 0} \frac{\mathcal{X}^{\mathcal{P}}(N_{ss} + \epsilon n, Z_{ss} + \epsilon z; \tilde{\lambda})(t) - \mathcal{X}^{P}(N_{ss}, Z_{ss}; \tilde{\lambda})(t)}{\epsilon} \\ &= \int_t^T G_{yn}(\tau - t) n(\tau) d\tau + \int_t^T G_{yz}(\tau - t) z(\tau) d\tau \end{aligned} \tag{32}$$

where

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

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

$$\begin{aligned} z(t) &\equiv \lim_{\epsilon \downarrow 0} \frac{\mathcal{Z}(\bar{x}_{ss} + \epsilon y; \tilde{m} + \epsilon w)(t) - \mathcal{Z}(\bar{x}_{ss}; \tilde{m})(t)}{\epsilon} \\ &= z_0(\omega)(t) + \int_0^t H_{zy}(t - s)\bar{y}(s) ds \end{aligned} \tag{33}$$

where

$$z_0(\omega)(t) \equiv -\sum_{j=0}^{\infty} \frac{\bar{x}_{ss}^2(\pi j + \frac{1}{2} - \cos(j\pi))}{\pi(\frac{1}{2}+j)} \frac{\langle \varphi_j, \omega\rangle}{\langle \varphi_j, \varphi_j\rangle} e^{-\mu_j t} \text{ and} \tag{34}$$

$$H_{zy}(q) = \tilde{m}_x(\bar{x}_{ss})\sigma^2 \sum_{j=0}^{\infty} \eta_j e^{-\mu_j q} \tag{35}$$

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 F.5.

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

$$\bar{y}(t) = \bar{y}_0(t) + \tilde{\Theta}(\bar{x}_{ss}) \int_0^T \tilde{K}(t,s)\bar{y}(s)ds \text{ where} \tag{36}$$

$$\bar{y}_0(\omega)(t) \equiv \int_t^T G_{yn}(\tau - t) n_0(\omega)(\tau) d\tau + \int_t^T G_{yz}(\tau - t) z_0(\omega)(\tau) d\tau \tag{37}$$

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

$$\tilde{K}(t,s) = \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} (c_j + c_i) e^{\psi_j t + \mu_i s} \left(\frac{e^{-(\psi_j+\mu_i)\max\{t,s\}} - e^{-(\psi_j+\mu_i)T}}{\psi_j + \mu_i} \right) > 0 \tag{38}$$

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 (36) which is the limit of

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

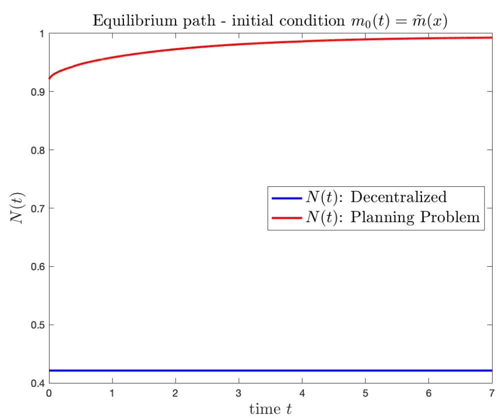
and where $\tilde{\mathcal{K}}^{j+1}(g)(t) \equiv \int_0^T \tilde{K}(t,s)\,\tilde{\mathcal{K}}^j(g)(s)\,ds$ for any bounded $g : [0,T] \to \mathbb{R}$. The operator

$\tilde{\mathcal{K}}$ is self-adjoint, and positive definite.

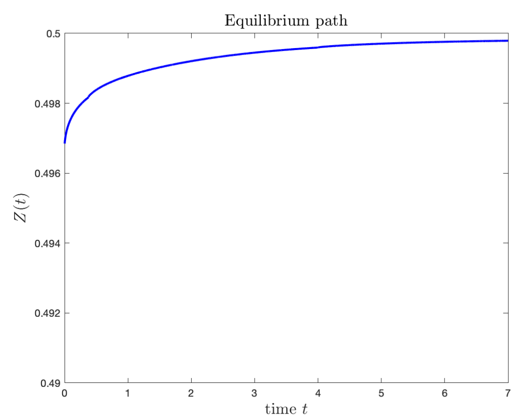
We again consider a perturbation to the steady state density of non-adopters. In this case, we let $m_0(x)$ equal the steady state distribution of no-adopters of the decentralized problem, so that the shock resembles a starting equilibrium with lower adoption than that prescribed by the planning solution.

In Panel (a) of Figure 3 we display the time path of the equilibrium $N(t)$ for both the decentralized and planning problems. Since the initial distribution of non-adopters corresponds to the steady state distribution of the decentralized problem, it is not surprising that the path of $N(t)$ of the decentralized problem is constant. On the other hand, the path of $N(t)$ of the planning problem jumps on impact and converges gradually to steady state. Panel (b) shows the equilibrium path of the optimal subsidy $Z(t)$, which jumps initially and increases over time thereafter. In this example, although both decentralized and planning problems have interior solutions, the planning solution mandates close to full adoption in steady state.

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



(a) $N(t)$



(b) $Z(t)$

8 Application: SINPE, a digital payment 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.[8] To use SINPE, users must have a bank account at a financial entity and link this account to their mobile number.

[8]SINPE is an acronym for the initials of "National Electronic Payment System" (*Sistema Nacional de Pagos Electrónicos*), in Spanish.

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 $161.[9] The average size of transactions in SINPE is about $46, and has slowly decreased over time, as shown is Figure I3.

8.1 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 using accounts across different banks. 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 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 was obtained from the Registry of Economic Variables of the Central Bank of Costa Rica, which tracks the universe of formal employment and labor earnings. This data set includes *monthly* details on each employee, including her occupation, earnings, and employment history between 2006 and 2021.[10] The average number of coworkers in our sample is 4.7 (median 1). Using this data set, we can identify which people are working at the same firm in a given

[9]Respectively, these limits in dollars correspond with approximately 200,000; 150,000; and 100,000 Costa Rican colones.

[10]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).

month to construct networks of coworkers that can be matched to SINPE records. Networks of coworkers change at a monthly frequency, as people change their 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.

8.2 From Model to Data

We 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, maximizing the following expression:

$$\xi^*(x,N) = \arg\max_{\xi} \frac{1+p}{p}\left[\beta(x,N)\xi - \frac{\xi^{1+p}}{1+p}\right]$$

where $p > 0$ so that the problem is convex and $\beta(x,N) > 0$. The first order condition describes the optimal intensity in which the technology is used:

$$\xi^*(x,N) = \beta(x,N)^{1/p} \tag{39}$$

We can choose the function $\beta(x,N)$ such that the indirect utility function gives the specified flow benefit, i.e:

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

The solution turns out to be

$$\beta(x, N) = [(\theta_0 + \theta_n N)\, x]^{\frac{p}{p+1}} . \tag{40}$$

Combining equation (39) with equation (40), and taking logs

$$\ln \xi_t^* = \frac{1}{1+p} \ln [(\theta_0 + \theta_n N_t)] + \frac{1}{1+p} \ln x_t \tag{41}$$

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.

We can separate the effect of N_t and x_t by taking differences

$$\Delta \ln \xi_t^* = \frac{1}{1+p} \Delta \ln(\theta_0 + \theta_n N_t) + \frac{1}{1+p} \Delta \log x_t$$

which can be estimated using non-linear least squares. Alternatively, an approximation around $N^\star = 0$ yields

$$\Delta \ln \xi_t^* \approx \gamma + \widetilde{\theta} \Delta N_t + \varepsilon_t \tag{42}$$

where γ is a constant and $\widetilde{\theta} \equiv \frac{1}{1+p}\frac{\theta_n}{\theta_0}$. We can estimate equation (42) using a linear specification. The coefficient of interest is $\widetilde{\theta}$ since strategic complementarities in the adoption of the technology exist if $\widetilde{\theta} > 0$ (i.e. $\theta_0 > 0$ and $\theta_n > 0$).

8.3 Stylized Facts

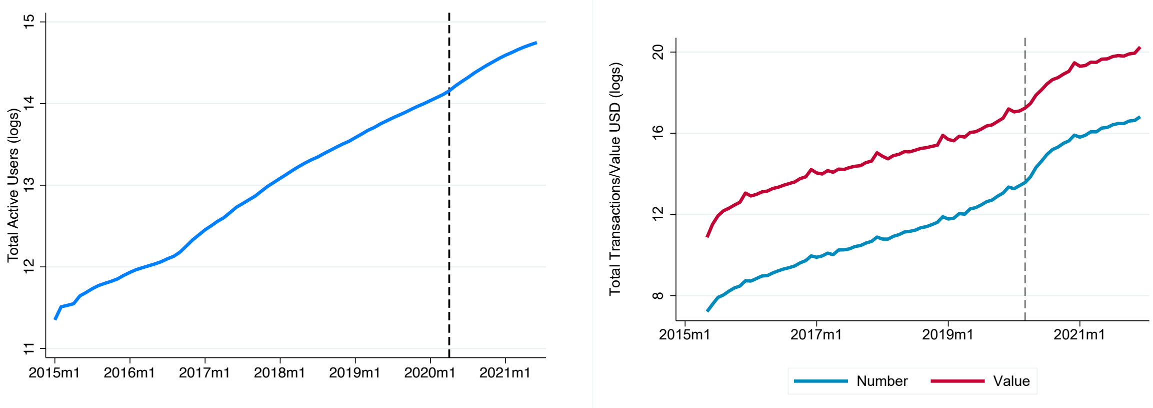
This section explores the diffusion process of SINPE across time and across networks, along with the relationship between individual characteristics and technology usage. We document five facts from the data that align with predictions of the model that we developed.

Fact 1: The technology diffused slowly. The adoption of SINPE has grown at a constant rate over time since its inception in 2015, as shown in Figure 4 using monthly data on the total number of adopters.[11] 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

[11] 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.

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

Fact 2: 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 I4, individual-to-individual transactions account for over 95% of all transactions, regardless of the time period considered.[12] This motivates us to study adoption through the lens of our model while focusing on peer-to-peer transactions only.

Fact 3: 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

[12] This finding holds if we instead consider unweighted number of transactions, as shown in Figure I5.

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 al least one of these networks.[13] Half of all transactions have a neighbor as counterpart, about 45% of all transactions are among coworkers, and 41% 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, 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 48, i.e. people do not necessarily transact with the *same* six peers each month.[14] The average transaction size is $46, and has decreased slowly over time, as shown in Figure I3.

Table 1: Share of Transactions Within Network

	Neighborhood (1)	Firm (2)	Family (3)	Union of all three (4)
Neighborhood	0.43			0.72
Firm	0.62	0.44		
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.

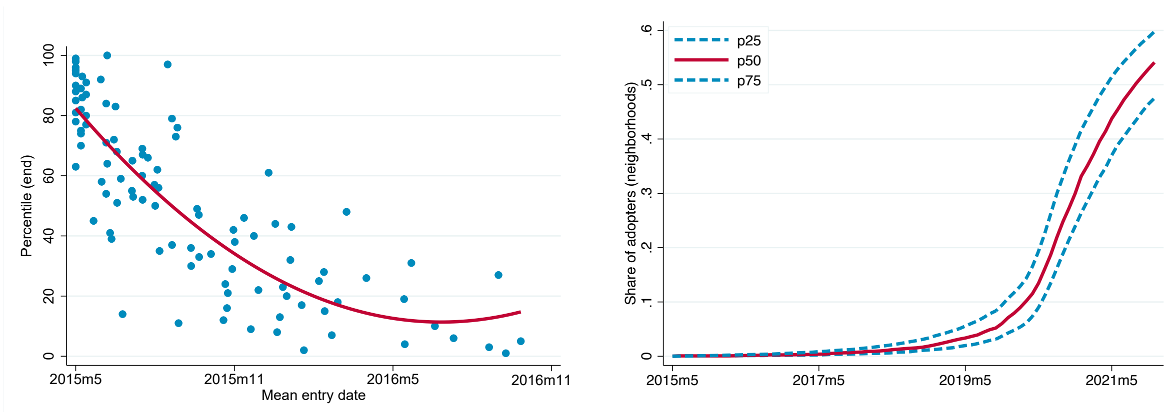
Fact 4: The adoption of the technology across networks was staggered. We empirically explore the dynamics of adoption for SINPE across networks and show that the early stages feature an S shaped profile. This profile is qualitatively similar to what is produced by the learning model, suggesting that initial awareness of SINPE was uneven across networks. We classify networks (i.e. neighborhoods) 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. Panel (a) of Figure 5 shows the timing of adoption across different percentiles. We measure the timing of adoption as the period in which we first see an individual within a network adopting. Panel (a) shows that networks with the

[13] 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.

[14] Average monthly patterns are documented in Figure I6.

largest shares of adopters also adopted the technology first; in fact, networks with the highest penetration of the technology adopted instantly after the technology was launched. On the other hand, networks with the lowest penetration took more than a year and a half to start adopting the technology. Panel (b) shows the diffusion path of the technology for the median neighborhood. It shows that the technology was adopted gradually within networks. Taken of together, these panels show that networks which adopted the technology early also tend to have higher penetration throughout our sample period. While Figure 5 is computed based on networks of neighbors, the same patterns emerge when analyzing networks of coworkers and relatives, as shown in Figure I10.

Figure 5: Entry and Diffusion Across and Within Networks of Neighbors



(a) Entry Across Networks

(b) Diffusion Within Network

Notes: Panel (a) shows the timing of adoption across networks, defined as neighborhoods. It shows the entry date (the first time an individual within a network adopts the technology) 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 (b) uses the same classification of percentiles to show the patterns of diffusion of the technology within networks.

Fact 5: 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 the ones that follow, we classify an individual as an adopter from the time she first used the app onward. First, as shown in Figure 6, 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)).[15] Early adopters are also younger, on average, than

[15]We classify an occupation as high-skill if it requires education or training beyond a high-school diploma.

later adopters, as shown in Figure I7.

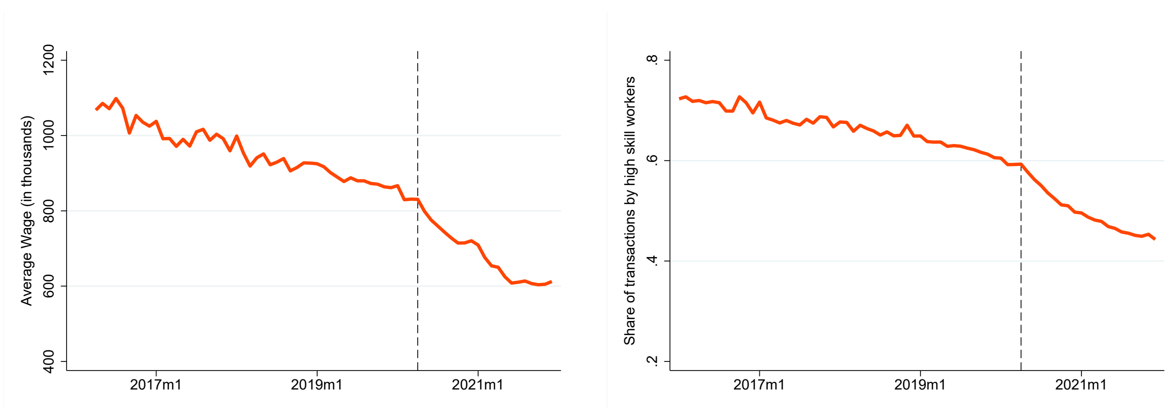
Second, recall the model-derived relationship between intensity of usage (ξ^n_{it}) and the share of user's i network who had adopted *when she used the app for the first time* ($N^n_{i,entry}$):

$$\ln \xi^n_{it} = \gamma + \beta N^n_{i,entry} + \lambda^n_t + \nu^n_{it},$$

where $n \in$\{neighbors, coworkers, relatives\} and ξ^n_{it} is defined as number of transactions of user i each month t. Our model predicts that $\beta < 0$, 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. The regression also includes network-time fixed effects, so that the relationship described by β is not mechanical.

We estimate $\hat{\beta}$ to be -2.2, with a standard error of 0.004 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. It is worth mentioning that this estimation includes network-time effects, thus, the inverse relationship that we document is not just mechanical. 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 6: Average Wage and Skill at the Time of Adoption



(a) Average Wage (in USD)　　(b) Share of High- Skill Users

Notes: Panel (a) shows the cross-sectional distribution of SINPE users' wages. 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. The figures include a vertical dashed line to mark the start of the COVID-19 pandemic (March 2020).

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

Dependent variable: Number of Transactions (logs)

	(1)	(2)	(3)
Size of Neighbors' Network at Entry	-2.215***		
	(0.004)		
Size of Coworkers' Network at Entry		-1.098***	
		(0.004)	
Size of Family Network at Entry			-1.061***
			(0.004)
Observations	27,648,863	13,263,171	12,001,989
Adjusted R^2	0.200	0.238	0.150
Network×Time FE	Yes	Yes	Yes

Notes: The dependent variable in this estimation is the number of transactions each month for each user, which we transform using the inverse hyperbolic sine function: $\ln(\xi + \sqrt{\xi^2 + 1})$. The coefficients describe 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.

Fact 6: 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 test for the presence—albeit suggestive—of these externalities along the intensive margin of adoption, we consider the following version of equation (42):

$$\Delta \ln \xi_{it}^n = \gamma + \widetilde{\theta} \Delta N_t^n + \psi \Delta size_t^n + \lambda_t + \nu_{it}^n, \tag{43}$$

where $\ln \xi_{it}^n$, the intensity with which individual i uses the technology, can be interpreted as either the value or the number of SINPE transactions in a given month t, N_t^n is share of user i's network that has adopted the app, $\Delta size_t^n$ is the change in the level of network n, and we include time fixed-effects, λ_t, which includes the effect of aggregate adoption increasing over time. Again, networks can be defined in different ways, and as such $n \in$\{neighbors, coworkers, relatives\}. This regression has several advantages. First, it 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.[16] Second, as the regression is in changes, individual effects which might affect usage cancel out, including the effect of the idiosyncratic taste (x), as it follows a random walk.[17]

[16]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.

[17]In line with the model, as x behaves differently when it reflects on the barriers 0 and U, we trim the top and bottom 1% of transactions in the empirical exercise.

Table 3 shows results when considering n as a user's network of neighbors, network of coworkers, and network of relatives. The dependent variable in this table refers to the number of SINPE transactions transformed using the inverse hyperbolic sine function. Results are robust to considering alternative transformations.[18]

Across specifications, we find that $\widetilde{\theta}$ 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. All findings remain unchanged if we consider the monthly value of transactions of each user as our dependent variable instead of the number of transactions, as reported in Table I4. Similarly, results are robust to including controls for COVID-19 and cohort fixed-effects, as shown in Table I5, Table I6, and Table I7.

Table 3: Changes in Number of Transactions and Network Changes

Dependent variable: IHS(Δ Number of Transactions)				
	(1)	(2)	(3)	(4)
Δ Share Neighborhood Adopters	1.062***			0.897***
	(0.025)			(0.044)
Δ Share Coworkers Adopters		0.228***		0.248***
		(0.006)		(0.008)
Δ (Log) Wage		0.046***		0.051***
		(0.001)		(0.001)
Δ Share Relatives Adopters			0.470***	0.476***
			(0.008)	(0.009)
Observations	25,632,610	16,208,557	11,275,971	7,230,892
Adjusted R^2	0.021	0.026	0.023	0.029
RMSE	0.793	0.765	0.793	0.759
Time FE	Yes	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. Standard errors are in parentheses. Extreme values (one and 99 percentile) were trimmed from the dependent variables. Results are robust to alternative transformations, as shown in Table I3, and to no trimming.

It is also possible to use an alternative measure of N, which will *by construction* comprehend all transactions. Namely, we take the last period in our sample (December 2021)—in which most adults have already adopted—as our starting point, and then look back in time at all transactions which have occurred. Then, for each individual, we define her network as the collection of people with which she transacted at some point in time. Thus, for instance,

[18] Table I3 reports results using logs and also following Davis and Haltiwanger (1992), who define $\Delta x_t = 2\frac{x_t - x_{t-1}}{x_t + x_{t-1}}$, where x is the value of transactions for each individual. Moreover, while all these table trim extreme values of the data (one and 99 percentiles), our results are robust to no trimming. Our preferred specification, however, considers trimming as specifications hold only when x is not at the reflecting barriers.

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 I10 shows the results of estimating equation (43) using this alternative network and the number of transactions per user as our dependent variable.[19] The positive and correlation between changes in usage and in share of adopters within network is always present across specifications.

8.4 Identification: Changes in Networks of Coworkers After a Mass Layoff

Fact 5 in the previous section documented a correlation between the intensity with which someone uses the app and the share of individuals in her network who have adopted it. In this section, we consider an identification strategy to claim that this relationship is causal. This strategy focuses on the network of coworkers and implements a movers design where we follow workers fired during a mass layoff.

We focus on the workers displaced during mass layoffs to examine the effect of network changes on the extensive and intensive margins of adoption.[20]

Extensive Margin of Adoption For the extensive margin, we consider the change in the probability of adoption for displaced workers *who had not downloaded the app by the time they were rehired* depending 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 larger incentives to adopt via the effect of strategic complementarities. We consider:

$$Adopt_i = \alpha + \hat{\theta}\Delta N_i^{coworkers} + \hat{\gamma}\Delta \ln wage_i + \hat{\psi}\Delta \ln size_i + \hat{\lambda} date\, hired_i + \hat{\omega}\Delta Covid_i + \epsilon_i, \quad (44)$$

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; and $\Delta Covid_i$ controls for the change in the cumulative

[19] Table I8 displays results considering instead the value of transactions per user.

[20] 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 had a stable workforce before this episode and did not recover in the following 12 months. More details are provided in Appendix J.

Table 4: Extensive Margin of Adoption and Changes in Coworkers' Network After a Mass Layoff

Dependent Variable: $Adopt_i$ (Logit)

	(1)	(2)	(3)
$\Delta N_i^{coworkers}$	8.298***	5.088***	4.764***
	(0.119)	(0.147)	(0.152)
	[0.450]	[0.340]	[0.315]
$\Delta \ln wage_i$		-0.004	-0.014
		(0.033)	(0.034)
$\Delta Covid_i$			0.113***
			(0.016)
Observations	22,249	17,658	17,658
Pseudo R^2	0.495	0.515	0.518
Time FE	No	Yes	Yes

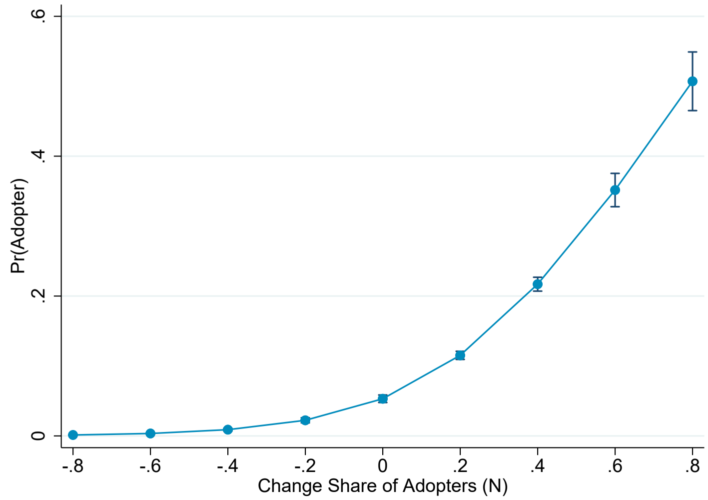
Notes: The unit of observation is the individual. We run regressions using data on mass layoffs that occurred between May 2015, when the technology was introduced, and December 2021. Standard errors are in parentheses. Marginal effects for the main variable of interest are reported in brackets.

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 J provides more details on each of these variables.

Table 4 shows the results estimating equation (44) 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 Table 4, is shown in Figure 7. This marginal effect is monotonous and—as expected—is present only when the change in the share of adopters who had adopted 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

Figure 7: Marginal Effect of Network Changes on Adoption Probability



Notes: This figure plots the marginal effect of $\Delta N_i^{coworkers}$ in the specification described by Column (3) of Table 4. Bars denote 95% confidence intervals.

in the previous subsection, it is possible to derive the relationship in equation (43) from our theoretical model, which speaks to the technology's intensity of usage. Similarly as in equation (43), we consider:

$$\Delta \ln \tilde{\xi}_i = \tilde{\alpha} + \tilde{\theta}\Delta N_i^{coworkers} + \tilde{\gamma}\Delta \ln wage_i + \hat{\psi}\Delta \ln size_i + \tilde{\lambda} date\, hired_i + \tilde{\omega}\Delta Covid_i + \quad (45)$$

$$\tilde{\delta} cohort_i + \tilde{\nu} \ln \sum^t \tilde{\xi}_i + \tilde{\epsilon}_i, \quad (46)$$

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, $cohort_i$ controls for the date when individual i adopted SINPE, and $\ln \sum^t \tilde{\xi}_i$ is the sum of all historical transactions made by agent i since she adopted the app. The last two variables aim to control for learning how to use the app due to having more people in your network who have adopted it. Other variables are defined in the same way as in equation (44).[21]

Table 5 displays our results using the number of transactions per user as our dependent variable.[22] 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. Figure 8 displays the marginal effect of these network changes following the specification described by Column (3) of Table 5. As Figure 8 shows, not only is the relationship between usage and network changes positive, but also whenever a worker moves to a firm with a lower adoption

[21] Appendix J provides more details on these variables and the choices made to conduct this exercise.
[22] Table I9 reports the same results with the value of transactions as dependent variable.

rate, her usage decreases (i.e. the change on the vertical axis is negative).[23]

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.

Table 5: 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***	1.647***	1.087***	0.995***
	(0.153)	(0.171)	(0.180)	(0.183)
$\Delta \ln wage_i$		0.401***	0.349***	0.363***
		(0.046)	(0.044)	(0.048)
$\Delta Covid_i$			0.167***	0.155***
			(0.020)	(0.023)
Observations	1,554	1,554	1,554	1,554
Adjusted R^2	0.141	0.221	0.257	0.280
Time FE	No	Yes	Yes	Yes
Cohort FE	No	No	No	Yes

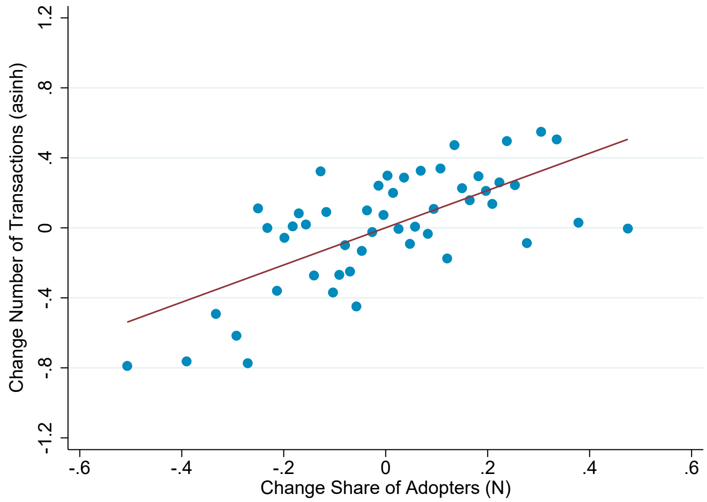
Notes: The unit of observation is the individual. We run regressions using data on mass layoffs that occurred between May 2015, when the technology was introduced, and December 2021. Standard errors, clustered by individuals, are in parentheses.

The analysis can be taken to an even more detailed level if, instead of considering all transactions in out left-hand-side variable, we focus only on those transactions that had a coworker as a counterpart. These 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, the results are remarkably similar to those using all transactions, as shown in Figure 9 and Table I11.[24]

[23] The marginal effect considering the value of transactions as dependent variable, as opposed to the number of transactions, is reported in Figure I8.

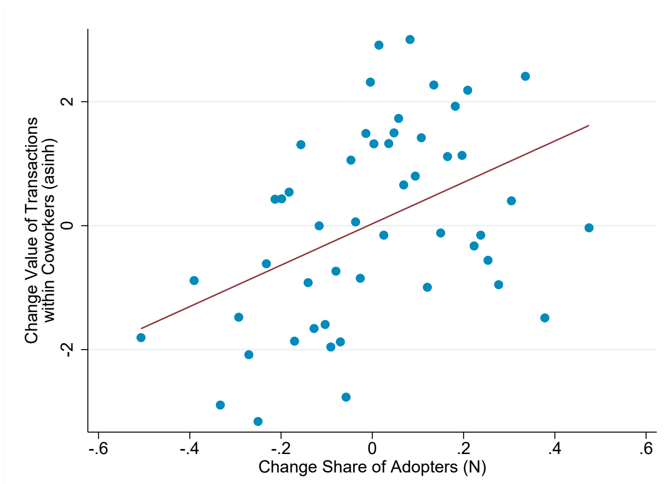
[24] The corresponding results using the value instead of the number of transactions is reported in Figure I9.

Figure 8: Marginal Effect of Network Changes on Usage Intensity



Notes: This figure plots the marginal effect of $\Delta N_i^{coworkers}$ in the specification described by Column (4) of Table 5. 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.

Figure 9: Marginal Effect of Network Changes on Usage Intensity Among Coworkers Only



Notes: This figure plots the marginal effect of $\Delta N_i^{coworkers}$ in the specification described by Column (4) of Table 5. Bars denote 95% confidence intervals. The dependent variable in this estimation is the number of transactions which has a coworker as a counterpart (transformed using the inverse hyperbolic sine function) on each period for each user.

9 Quantitative Performance and Optimal Subsidy

In this section, we calibrate our model and evaluate its performance relative to the data of SINPE Mobile. We begin by describing an extension of the model that combines the model of strategic complementarities with the learning model. This extension is helpful to make the model consistent with some features of the data. We then describe our calibration procedure in detail.

A Learning Model with Strategic Complementarities: Using the derivations of the previous section, it is straightforward to extend our benchmark model of strategic complementarities to include random diffusion of the technology across agents. The variational inequality of the adoption decision, the value of an agent that already has adopted the technology $a(x,t)$ and the value of waiting $v(x,t)$, are the same as in the model with strategic complementarities since these decision are made after agents are aware of the technology. On the other hand, the law of motion of m needs to be modified to include the inflow of informed agents as in the learning model.

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

where $I(t)$ is given by equation (79) and, as before, β_0 indicates the number of meetings per unit of time. The reflecting barrier of x at zero implies $0 = m_x(0,t)$ for all $t \geq 0$ and continuity of m implies that $m(\bar{x},t) = 0$ all $t \geq 0$.

Calibration: By Lemma 1, the problem with strategic complementarities features 5 independent parameters as U and θ_0 can be normalized without affecting the nature of the solution: $\nu, \rho, \widetilde{\theta}, \tilde{\sigma}$, and $\widetilde{c}$, where the tilde indicates the normalized parameters.[25] In addition, the model that includes learning requires an additional parameter to be calibrated, β_0, as well as an initial condition for the population that is informed, $I(0)$.

We calibrate ν to 0.0278 in order to match the rate at which agents stop using SINPE in the data. This is the average fraction of agents in 2019-2021 that had adopted SINPE but did not conduct a single transaction using the application the following 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; this parameter can admit higher values if we assume

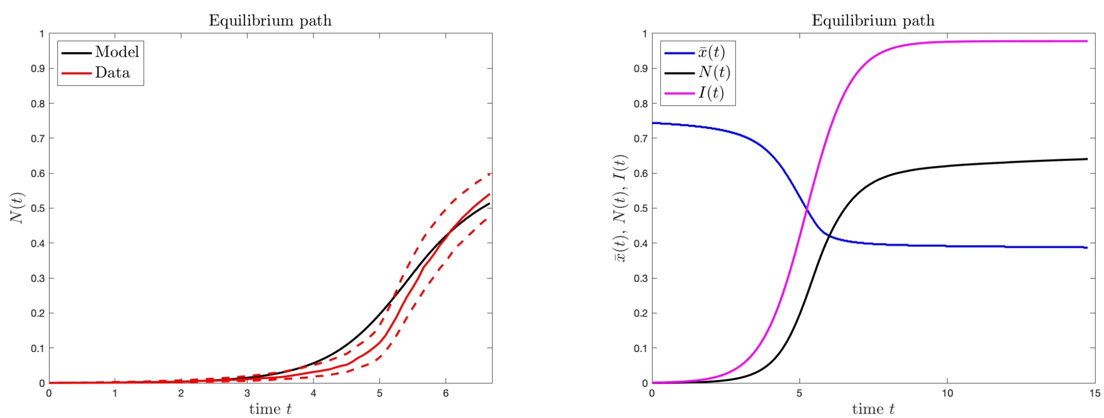
[25] $\widetilde{c} \equiv \frac{c}{U\theta_0}$, $\tilde{\sigma} \equiv \frac{\sigma}{U}$, and $\widetilde{\theta} \equiv \frac{\theta_n}{\theta_0}$.

that agents expect new technologies to arrive in the future and replace SINPE. The values of ν and r imply $\rho = r + \nu = 0.0778$.

We interpret the flow benefit of agents who adopt the technology as being proportional to how many transactions they conduct (i.e. how intensively they use SINPE). Thus, we set $\widetilde{\theta} = 1.8$, which is consistent with our estimates in Table 5.[26] We set $\tilde{\sigma} = 0.032$ to reflect the variation of transactions conditional on the size of the network of neighbors. To be more precise, we use the residuals obtained from the estimation of equation (43) and calculate the standard deviation. We adjust this estimate to reflect that the regression is estimated in logs and for the range of transactions (i.e. U). We obtain estimates of U using the distribution of transactions. Specifically, in the data, the average number of transactions per year is 95 and the 99^th percentile is 775. We find U by noticing that the upper bound of the distribution of transactions must equal to $\left[U(1+\frac{\theta_n}{\theta_0})\right]^{\frac{1}{1+p}}$ (i.e. $x = U$ and $N = 1$).

We set $\widetilde{c} = 9$ to match the fraction of the population that has adopted the technology by the end of 2021. This value implies that approximately 90% of the population adopts in steady state. Lastly, we set $\beta_0 = 1.3$ to match the path of technology adoption for the median neighborhood in the data. We display the path of adopters starting at $I(0) = 0.001$, that is, 0.1 percent of population is informed about SINPE Mobile at the time it was launched.

Figure 10: Path of Adopters (Short-Run and Long-Run)



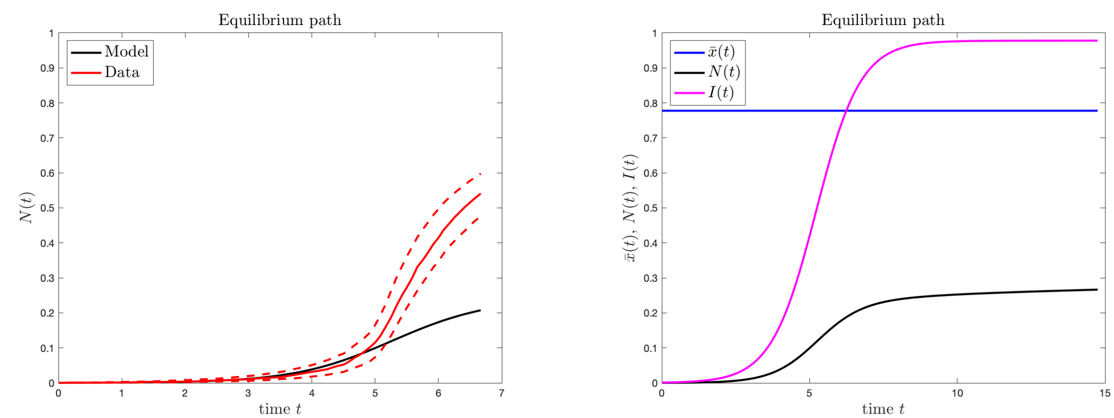
(a) Model vs Data (b) Long-Run Path

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.

[26] We assume $p \approx 0$.

Panel (a) of Figure 10 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 under our baseline calibration, both the speed and the levels 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 figure shows that most people are informed about the technology within the first 7 years. In steady state, approximately 97.5% of the population know about the application. The figure also shows the long-run level of $N(t)$. The model predicts that in steady state 90% of the population living in the median neighborhood will adopt the application. Lastly, the path of $\bar{x}(t)$ indicates that, consistent with our empirical evidence, there is selection in the model; agents that benefit the most from the technology adopt first. This can be seen by the declining path of $\bar{x}(t)$.

Figure 11: Path of Adopters: Only Learning (Short-Run and Long-Run)



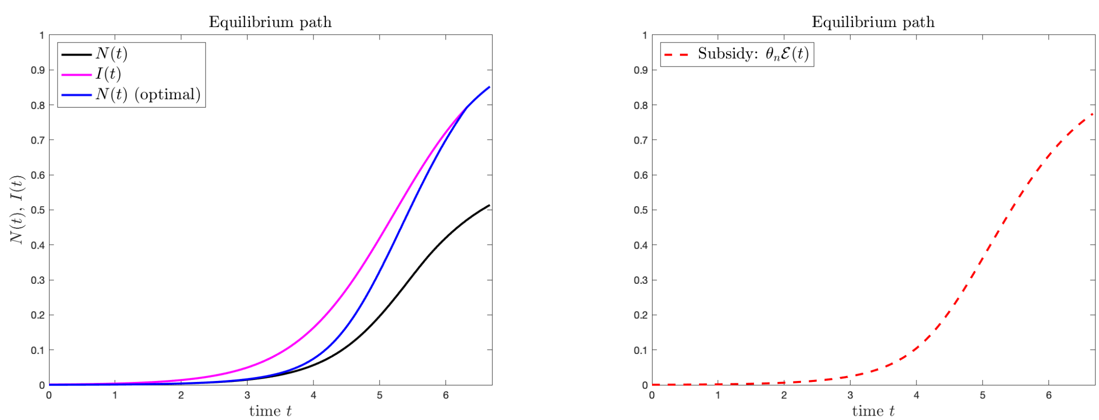
(a) Model vs Data (b) Long-Run Path

Notes: Panel(a) compares the path of adopters in the model and in the data when $\theta_n = 0$. 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 and $\theta_n = 0$.

Figure 11 shows the performance of the model under our baseline calibration but with $\theta_n = 0$. It shows that, without strategic complementarities, the levels of adoption by the end of 2021 (6 and a half years after the launch of SINPE) would be around 20%. Panel (b) shows that the path of $\bar{x}(t)$ in the model with only learning is flat, which indicates that this version of the model does not feature selection in the adoption of the technology as observed in the data.

Optimal Subsidy: Panel (a) of Figure 12 shows the optimal adoption path relative to the path of adopters from the decentralized equilibrium. During the first four years after the launch of the technology, the optimal level of adoption are similar to those of the equilibrium without subsidy. The optimal path of adopters from the planning problem is higher than that of the equilibrium. In fact, by the end of 2021, it is equal to the total number of informed agents in the economy, 30 percentage points higher than the levels of adoptions observed in the data. Panel (b) shows the path of the optimal subsidy. As the share of adopters increase, so does the adoption externality. As a result, the optimal subsidy, which is the same across agents, also increases over time. The optimal subsidy eventually pushes the economy to universal adoption.

Figure 12: Planning Problem: Solution and Optimal Subsidy



(a) Optimal Path of Adopters

(b) Optimal Subsidy

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, $Z(t)(\theta_0 + \theta_n N(t))$.

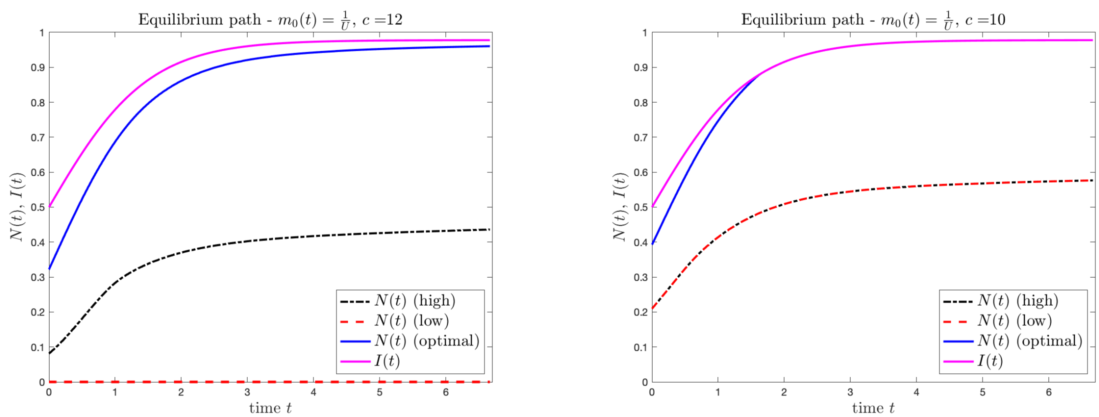
High Adoption Cost: Panel (a) of Figure 13 shows an example with a higher adoption cost (i.e. $\tilde{c} = 12$) and a higher fraction of the population informed about the technology at launch. It is motivated by the recent experience in El Salvador, where 70% of the population knew about the payment application introduced by the government (i.e. Chivo Wallet) 7 months after its initial launch.[27] Panel (a) shows the paths of adopters $N(t)$ in the decentralized equilibrium. It shows that, when the adoption cost is larger, the equilibrium where no one

[27] The Salvadorean government also launched an app called "Chivo Wallet," which allows users to digitally trade both bitcoin and dollars without paying any transaction fees.

adopts the technology is not ruled out. In this case, the optimal adoption path increases gradually and takes more than 7 years to converge to the share of informed agents.

Panel (b) shows the same paths for a lower adoption cost (i.e $c = 10$). The figure shows that a large enough adoption subsidy can eliminate the no adoption equilibrium. Thus, even if the optimal subsidy is not implemented, a permanent subsidy can in fact solve the coordination failure and send the decentralized economy to the high adoption equilibrium.[28]

Figure 13: Planning Problem: High Adoption Cost



(a) Optimal Path of Adopters (b) Optimal Subsidy

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 for values of $c = 12$ and 70% of the population informed 7 months after the launch of the technology. Panel (b) shows the same variables for $c = 10$ and 70% of the population informed 7 months after the initial launch.

10 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 leveraging tools form mean field game theory. The model provides a framework to characterize the process of learning, generates slow adoption through a novel mechanism—waiting for others to adopt—and allows us to derive predictions that can be tested empirically.

[28]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. However, our model indicates that the subsidy was not large enough to rule out the no adoption equilibrium. More details about the app and the implementation of the subsidy can be found in Alvarez, Argente and Van Patten (2022a).

We also 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 equilibrium. We show that, in our framework, the optimal subsidy increases over time but it is flat, thus, easily implementable.

Our application consists of analyzing 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 most 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 our both the descentralized equilibrium and the planning problem in our model, which allows us to estimate, for instance, 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.

The methodology we develop can be useful for wide set of multidimensional dynamic problems, and the model can be applied to studying any technology that features strategic complementarities, learning, or both. Moreover, the setup is simple but can lead to interesting extensions, for instance, a drift in the idiosyncratic benefit of adoption could be key to understand other technologies with adoption costs that decrease or increase over time.

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.

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.

APPENDIX

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 2. 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:

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

Thus we define:

$$p = \frac{\sigma^2}{2} \frac{\Delta_t}{(\Delta_x)^2} \frac{1}{(1-\rho\Delta_t)} \tag{47}$$

and write:

$$a_{i,j-1} = \Delta_t\, x_i\, (\theta_0 + \theta_n N_j) + (1-\rho\Delta_t) [p a_{i-1,j} + (1-2p) a_{i,j} + p a_{i+1,j}] \tag{48}$$

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

$$a_{1,j-1} = \Delta_t\, x_1\, (\theta_0 + \theta_n N_j) + (1-\rho\Delta_t) [(1-p) a_{1,j} + p a_{2,j}] \tag{49}$$

$$a_{I,j-1} = \Delta_t\, x_I\, (\theta_0 + \theta_n N_j) + (1-\rho\Delta_t) [p a_{I-1,j} + (1-p) a_{I,j}] \tag{50}$$

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

$$a_{i,J} = a_{i,T} \text{ for } i = 1, 2, \ldots, I \tag{51}$$

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:

$$\tilde{a}_i = \Delta_t\, x_i\, (\theta_0 + \theta_n N_{ss}) + (1-\rho\Delta_t) [p\tilde{a}_{i-1} + (1-2p)\tilde{a}_i + p\tilde{a}_{i+1}] \tag{52}$$

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

$$\tilde{a}_1 = \Delta_t\, x_1\, (\theta_0 + \theta_n N_{ss}) + (1-\rho\Delta_t) [(1-p)\tilde{a}_1 + p\tilde{a}_2] \tag{53}$$

$$\tilde{a}_I = \Delta_t\, x_I\, (\theta_0 + \theta_n N_{ss}) + (1-\rho\Delta_t) [p\tilde{a}_{I-1} + (1-p)\tilde{a}_I] \tag{54}$$

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

$$v_{i,j-1} = \max\left\{-c + a_{i,j}\ ,\ (1-\rho\Delta_t)\left[pv_{i-1,j} + (1-2p)v_{i,j} + pv_{i+1,j}\right]\right\} \tag{55}$$

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

$$v_{1,j-1} = \max\left\{-c + a_{1,j}\ ,\ (1-\rho\Delta_t)\left[(1-p)v_{1,j} + pv_{2,j}\right]\right\} \tag{56}$$

$$v_{I,j-1} = \max\left\{-c + a_{I,j}\ ,(1-\rho\Delta_t)\left[pv_{I-1,j} + (1-p)v_{I,j}\right]\right\} \tag{57}$$

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:

$$\bar{x}_j = \min_{\{i=1,\ldots,I\}} \{x_i : v_{i,j} = -c + a_{i,j}\} \text{ for all } j = 1, 2, \ldots, J$$

$$\bar{i}_j = \min_{\{i=1,\ldots,I\}} \{i : v_{i,j} = -c + a_{i,j}\} \text{ for all } j = 1, 2, \ldots, J \text{ so that}$$

$$\bar{x}_j = x_{\bar{i}_j} \text{ for all } j = 1, 2, \ldots, J$$

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:

$$\tilde{v}_i = \max\left\{-c + \tilde{a}_i\ ,\ (1-\rho\Delta_t)\left[p\tilde{v}_{i-1} + (1-2p)\tilde{v}_i + p\tilde{v}_{i+1}\right]\right\} \tag{58}$$

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

$$\tilde{v}_1 = \max\left\{-c + \tilde{a}_1\ ,\ (1-\rho\Delta_t)\left[(1-p)\tilde{v}_1 + p\tilde{v}_2\right]\right\} \tag{59}$$

$$\tilde{v}_I = \max\left\{-c + \tilde{a}_I\ ,\ (1-\rho\Delta_t)\left[p\tilde{v}_{I-1} + (1-p)\tilde{v}_I\right]\right\} \tag{60}$$

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

$$q = \frac{\sigma^2}{2}\frac{\Delta_t}{(\Delta_x)^2}\frac{1}{(1 - \nu\Delta_t)} \tag{61}$$

we can write it as:

$$m_{1,j+1} = (1 - \nu\Delta_t)\left(qm_{2,j} + (1 - q)m_{1,j}\right) + \nu\Delta_t\frac{1}{U} \tag{62}$$
$$m_{i,j+1} = (1 - \nu\Delta_t)\left(qm_{i+1,j} + (1 - 2q)m_{i,j} + qm_{i-1,j}\right) + \nu\Delta_t\frac{1}{U} \text{ for } i = 2, 3, \ldots, \bar{i}_j - 1 \tag{63}$$
$$m_{i,j+1} = 0 \text{ for } i = \bar{i}_j, \ldots, I \tag{64}$$

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

$$m_{i,1} = m_0(x_i) \text{ and } i = 1, 2, \ldots, I \tag{65}$$

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

$$N_j = 1 - \left(\sum_{i=1}^{I} m_{i,j}\Delta_x - m_{1,j}\Delta_x/2 - m_{\bar{i}_j-1,j}\Delta_x/2\right) \text{ for } j = 1, 2, \ldots, J \tag{66}$$

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

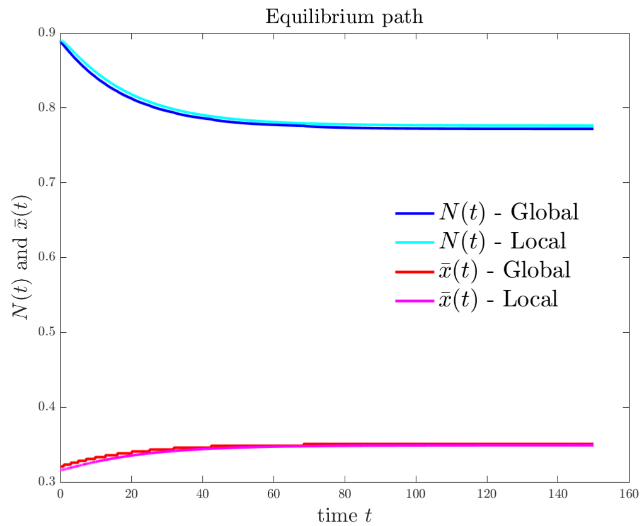
$$N^{k+1} = \mathcal{F}\left(N^k; m_0, D_T\right) \equiv \mathcal{N}\left(\mathcal{X}\left(N^k; D_T\right); m_0\right) \text{ for } k = 0, 1, 2, \ldots$$

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

$$-c + \mathbb{E}\left[\int_0^\infty e^{-\rho t} x(t)\,(\theta_0 + \theta_n N(t))\,dt \,|\, x(0) = x_1\right] \geq \qquad (67)$$
$$\mathbb{E}\left[-ce^{-\rho\tau_1} + \int_{\tau_1}^\infty e^{-\rho t} x(t)\,(\theta_0 + \theta_n N(t))\,dt \,|\, x(0) = x_1\right].$$

or equivalently that

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

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

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

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:

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

Using this strict inequality we get a contradiction with equation (67), 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

$$P(x,s,t) = Pr\left[X(r) \le \bar{x}(r),\ \text{for all } r \in [s,t] \mid X(s) = x\right] e^{-\nu(t-s)} \tag{68}$$

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 (68) 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 $\bar{x}' \ge \bar{x}$ 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 $\{X(r) \le \bar{x}(r),\ \text{for all } r \in [s,t]\}$ is included in the set of events $\{X(r) \le \bar{x}'(r),\ \text{for all } r \in [s,t]\}$, since $\bar{x}(r) \le \bar{x}'(r)$, and hence $P'(x,s,t) \ge P(x,s,t)$. Thus $M'(t) \ge M(t)$. Since $N'(t) = 1 - M'(t)$ and $N(t) = 1 - M(t)$, obtaining the desired result that $N'(t) \le N(t)$.

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. □

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) $\partial q(x)/\partial\theta_n = x\left(1 - x/U\right) > 0$ if $x \in (0, U)$, (2) $\theta_n = 0$ then, $q(x) = x\theta_0 - \rho c = U\left(\theta_0 x/U - \rho c/U\right) < U\left(\theta_0 - \rho c/U\right) < 0$, 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. □

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

$$N^a\left(\bar{x}^j(c)\right) = N^o\left(\bar{x}^j(c), c\right) \tag{69}$$

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 (69) 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:

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

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

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

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

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

which gives equation (19) in the main text.

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

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

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 equilibrium conditions

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

$$m_0(x) = \tilde{m}(x) + \epsilon\, \omega(x) \ . \tag{70}$$

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

$$\rho d(x,t) = x\theta_n n(t) + \frac{\sigma^2}{2} d_{xx}(x,t) + d_t(x,t) \tag{71}$$

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:

$$\begin{aligned} d(\bar{x}_{ss}, t) &= 0 \\ \tilde{D}_{xx}(\bar{x}_{ss})\bar{y}(t) + d_x(\bar{x}_{ss}, t) &= 0 \\ d_x(0, t) &= 0 \end{aligned} \tag{72}$$

for $t \in [0, T]$ and $d(x, T) = 0$ for $x \in [0, \bar{x}_{ss}]$. Note that equation (72) 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 (71) with respect to x and combining it with equation (72) we find

$$\bar{y}(t) = \frac{\theta_n}{\tilde{D}_{xx}(\bar{x}_{ss})} \int_t^T G(\tau - t) n(\tau) d\tau \tag{73}$$

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 (71), and the boundary conditions in equation (72), 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 (71) 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 (72), 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:

$$p_t(x,t) = \frac{\sigma^2}{2} p_{xx}(x,t) - \nu p(x,t) \tag{74}$$

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

$$\begin{aligned} \tilde{m}_x(\bar{x}_{ss})\bar{y}(t) + p(\bar{x}_{ss}, t) &= 0 \\ p_x(0,t) &= 0 \end{aligned} \tag{75}$$

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

$$p(0,x) = \omega(x) \tag{76}$$

The solution for p satisfies the p.d.e given in equation (74), its boundary conditions in equation (75), and the initial condition in equation (76). We have

Lemma 6. The solution for the KFE equation for p, satisfying the p.d.e given in equation (74), the boundary conditions in equation (75), and the initial condition in equation (76), 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:

$$\begin{aligned} n(t) &= -\int_0^{\bar{x}_{ss}} p(x,t)dx \\ &= n_0(t) + \frac{\tilde{m}_x(\bar{x}_{ss})\sigma^2}{\bar{x}_{ss}}\int_0^t J(t-\tau)\bar{y}(\tau)d\tau \end{aligned} \tag{77}$$

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 (73), $\bar{y}(t)$ 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 (77) 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 (73) and equation (77) 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 $\text{Lip}_K \equiv \sup_{t\in[0,T]} \int_0^T |K(t,s)|ds$, 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 $T \to \infty$

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

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

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

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

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

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

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

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 $|\Theta(\bar{x}_{ss})|\text{Lip}_K < 1$. 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 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 set up. 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 14. 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 $\bar{x}$:

$$
N(t) = I(t)\left(1 - \frac{\bar{x}}{U}\right) + \int_0^t e^{-\nu(t-\tau)}\left[-\frac{\sigma^2}{2} m_x(\bar{x},\tau)\right] d\tau \tag{78}
$$

The expression in the right hand side of $N(t)$ in Proposition 14 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 15. 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

$$I(t)=1-S(t)=\left(1-\frac{\nu}{\beta_0}\right)\frac{e^{(\beta_0-\nu)t}}{\frac{\left(1-\frac{\nu}{\beta_0}\right)}{I(0)}-1+e^{(\beta_0-\nu)t}} \tag{79}$$

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 15, 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 16. 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:

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

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

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

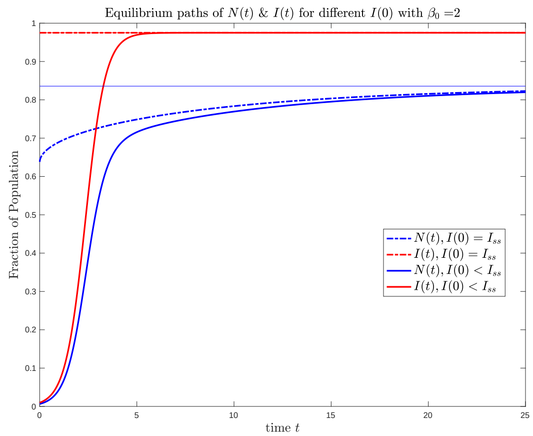
Combining the expression for $N(t)$ in Proposition 16 with the path of $I(t)$ solved for in Proposition 15 we obtain an explicit solution to $N(t)$. Next we analyze the invariant distribution in this model, which is the value at which it tends as $t \to \infty$. We denote $\tilde{m}$ the density for 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 17. 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:

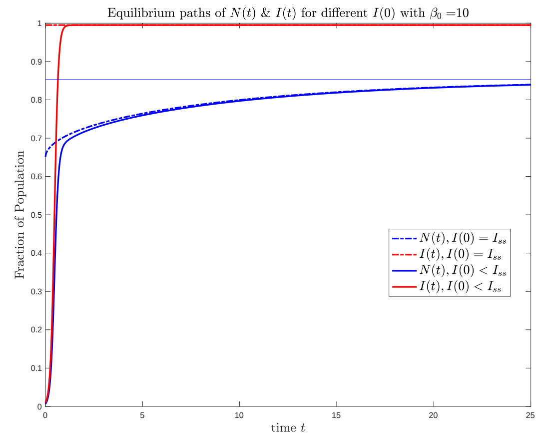
$$\tilde{m}(x) = (1 - \tfrac{\nu}{\beta_0})\tfrac{1}{U}\left(1 - \frac{\cosh(\gamma x)}{\cosh(\gamma\bar{x})}\right) \text{ where } \gamma = \sqrt{2\nu}/\sigma \text{ and thus}$$
$$N_{ss} = I_{ss} - \int_0^{\bar{x}} \tilde{m}(x)dx = (1 - \tfrac{\nu}{\beta_0})\left[1 - \tfrac{\bar{x}}{U}\left(1 - \frac{\tanh(\gamma\bar{x})}{\gamma\bar{x}}\right)\right] \tag{80}$$

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 (80) 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 E2: Equilibrium paths of N and I of Pure Learning Model



Slow learning, $\beta_0 = 2$.



Fast Learning, $\beta_0 = 10$.

Figure E2 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 $\bar{x}$.

E.1 Proofs for the learning model

Proof. (Proposition 14) 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 15) 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 16) 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:

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

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

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

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 17) We can rewrite the o.d.e. for $\tilde{m}$ as:

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

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.

□

F Planning Problem

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

F.1 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.

F.2 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:

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

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:

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

which can be written in vector notation as:

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

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.

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

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:

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

and subject to non-negativity:

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

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:

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

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 $2 \leq j \leq I-1$:

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

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:

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

dividing by Δ and further rearranging the expressions:

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

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

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

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

For the top boundary $j=I$:

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

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

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

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

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

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:

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

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 $+\infty$).

F.3 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

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

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:

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

and

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

F.4 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

$$\rho\ell(x,t) = x\theta_n n(t) + \theta_n z(t) + \frac{\sigma^2}{2}\ell_{xx}(x,t) + \ell_t(x,t) \tag{81}$$

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:

$$\begin{aligned} \ell(x,T) &= 0 \\ \ell_x(0,t) &= 0 \\ \ell(\bar{x}_{ss},t) &= 0 \\ \tilde{\lambda}_{xx}(\bar{x}_{ss})\bar{y}(t) + \ell_x(\bar{x}_{ss},t) &= 0 \end{aligned} \tag{82}$$

Proposition 18. 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 $j = 1, 2, ...$ 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

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

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

F.5 Perturbation analysis of the Planning Problem

Recall that from equation (82), $\bar{y}(t)$ is equal to

$$\begin{aligned}
\bar{y}(t) &= \frac{-\ell_x(\bar{x}_{ss}, t)}{\tilde{\lambda}_{xx}(\bar{x}_{ss})} \\
&= \int_t^T \frac{\theta_n}{\tilde{\lambda}_{xx}(\bar{x}_{ss})} \sum_{j=0}^{\infty} c_j e^{-\psi_j(\tau - t)} n(\tau) d\tau + \int_t^T \frac{2\theta_n}{\tilde{\lambda}_{xx}(\bar{x}_{ss}) \bar{x}_{ss}} \sum_{j=0}^{\infty} c_j e^{-\psi_j(\tau - t)} z(\tau) d\tau \\
&= \int_t^T G_{yn}(\tau - t) n(\tau) d\tau + \int_t^T G_{yz}(\tau - t) z(\tau) d\tau \qquad (83)
\end{aligned}$$

The expression for $n(t)$ is given by equation (77) 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 (83) can be written as

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

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

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

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

$$
\begin{aligned}
\bar{y}_0(t) &\equiv \int_t^T G_{yn}(\tau - t)n_0(\tau)d\tau + \int_t^T G_{yz}(\tau - t)z_0(\tau)d\tau \\
&= \frac{-\theta_n}{\tilde{\lambda}_{xx}(\bar{x}_{ss})} \int_t^T \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} c_j \frac{\bar{x}_{ss}}{\pi(\frac{1}{2}+i)} \frac{\langle \varphi_i, \omega \rangle}{\langle \varphi_i, \varphi_i \rangle} e^{-\psi_j(\tau - t)} e^{-\mu_i \tau} d\tau \\
&+ \frac{-\theta_n}{\tilde{\lambda}_{xx}(\bar{x}_{ss})} \int_t^T \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} c_i \frac{\bar{x}_{ss}}{\pi(\frac{1}{2}+i)} \frac{\langle \varphi_i, \omega \rangle}{\langle \varphi_i, \varphi_i \rangle} e^{\psi_j t} e^{-\psi_j(\tau - t)} e^{-\mu_i \tau} d\tau \\
&= \frac{-\theta_n}{\tilde{\lambda}_{xx}(\bar{x}_{ss})} \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} (c_j + c_i) \frac{\bar{x}_{ss}}{\pi(\frac{1}{2}+i)} \frac{\langle \varphi_i, \omega \rangle}{\langle \varphi_i, \varphi_i \rangle} e^{\psi_j t} \left(\frac{e^{-(\psi_j+\mu_i)t} - e^{-(\psi_j+\mu_i)T}}{\psi_j + \mu_i} \right) \qquad (84)
\end{aligned}
$$

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

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

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

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

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

Thus, equation (83) 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 $e^{-rt}M(t,s) = e^{-rs}M(t,s)$

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

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:[29]

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

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

[29]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.

any $t \geq 0$.[30] 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.[31] 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.

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

$$\bar{x}(t)\left[\theta_0 + \theta_n(1 - F(\bar{x}(t)))\right] - \rho c = 0 \tag{85}$$

[30] 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.

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

Note that in equation (85) 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 (85), 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 (85) 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.

Both cases are shown in Figure 1, which shows the solutions of equation (85). In Panel (a) the solution with low $\bar{x}_{ss}$ is the stable solution, since $\bar{x}(t)$ is weakly decreasing from Lemma 7, and the one with higher adoption, since $N(t) = 1 - F(\bar{x}(t))$. Panel (b) shows that with low strategic complementarities (i.e. low θ_n), there is only one steady state.

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:

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

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

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

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 (85). 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:

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

Using the law of motion of m we have:

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

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

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

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

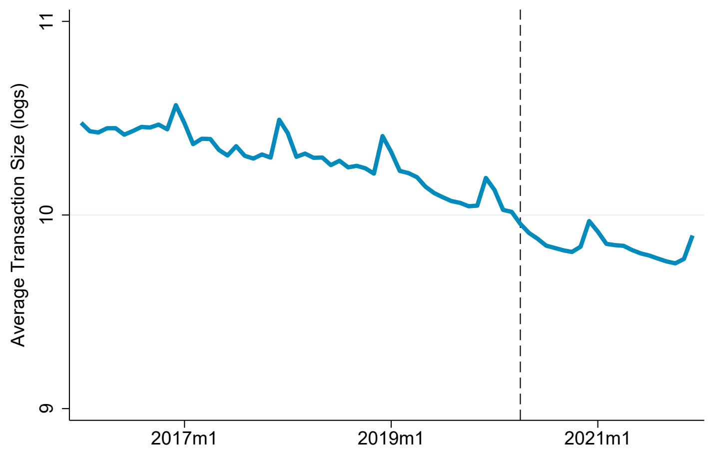
We can solve this o.d.e. for each x, using the boundary $m(x,s) = m_0(x)$. This gives

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

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

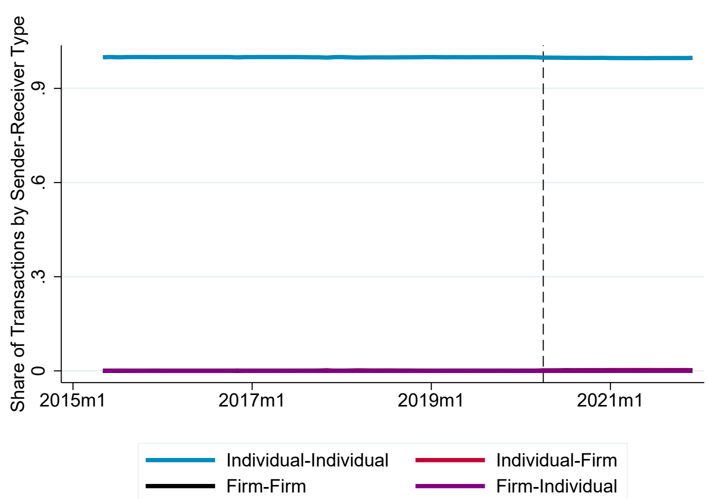
I Additional Figures and Tables from SINPE's Application

Figure I3: Average Transaction Size



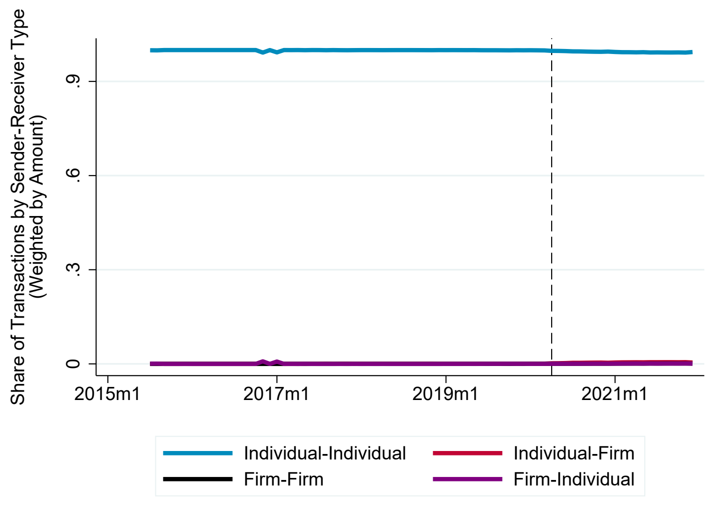
Notes: The figure shows the evolution of the average transaction size in SINPE. The figure includes a vertical dashed line to mark the start of the COVID-19 pandemic (March 2020).

Figure I4: Transactions by Sender-Receiver Type



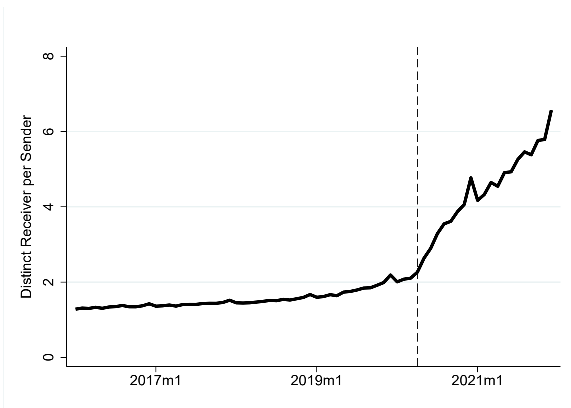
Notes: Transactions are classified according to the type of user. Individuals correspond with Costa Rican adult citizens. Firms correspond with formal enterprises.

Figure I5: Share of Transactions Between Types of Users (Weighted by Amount)

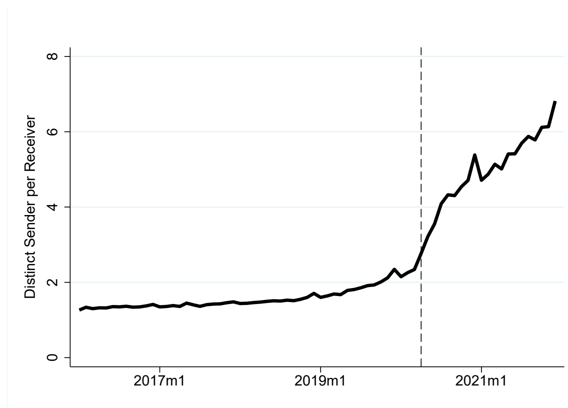


Notes: The figure shows total number of SINPE transactions between four different types of users, as a share of all of their transactions.

Figure I6: Mean Number of Connections per User



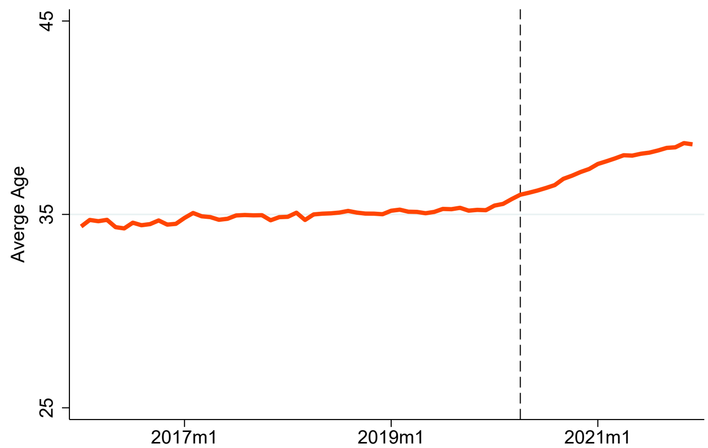
(a) Receivers



(b) Senders

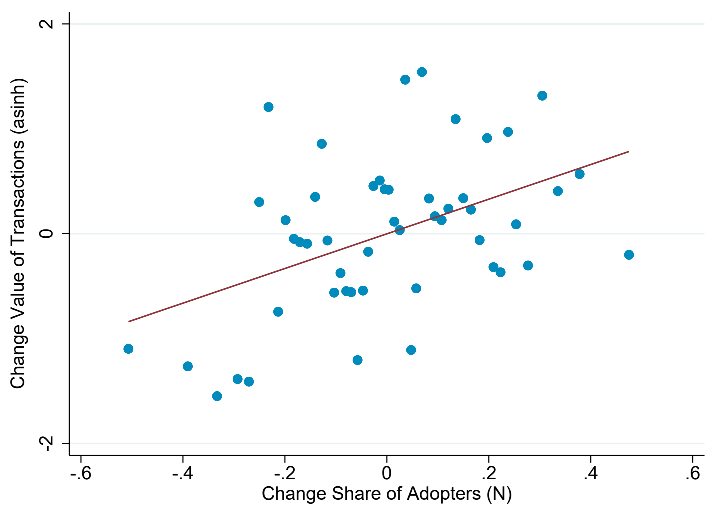
Notes: The figures include a vertical dashed line to mark the start of the COVID-19 pandemic (March 2020).

Figure I7: Average Age at the Time of Adoption



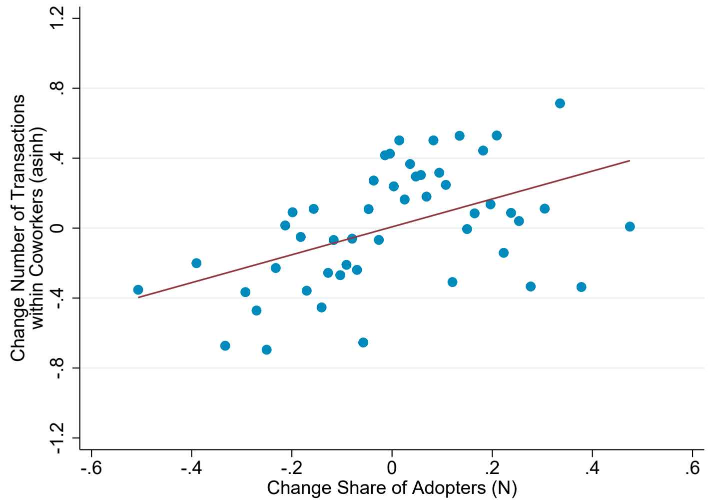
Notes: The figure includes a vertical dashed line to mark the start of the COVID-19 pandemic (March 2020).

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



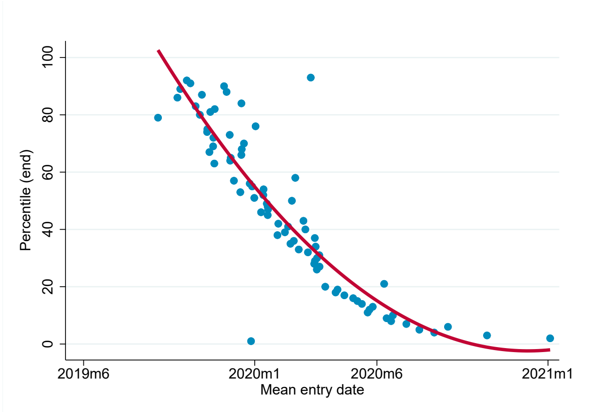
Notes: This figure plots the marginal effect of $\Delta N_i^{coworkers}$ in the specification described by Column (3) of Table 5. Bars denote 95% confidence intervals. The dependent variable in this estimation is the number of transactions (transformed using the inverse hyperbolic sine function) on each period for each user.

Figure I9: Marginal Effect of Network Changes on Usage Intensity (Value of Transactions) Among Coworkers Only

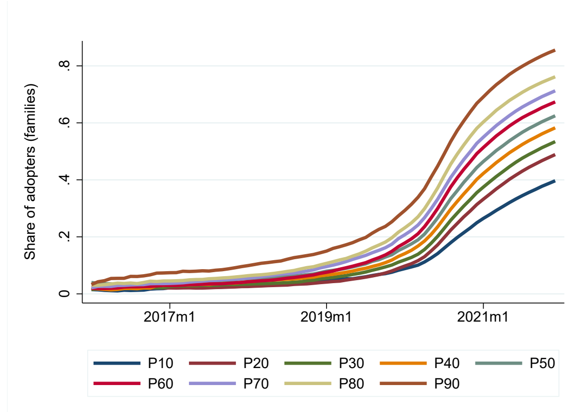


Notes: This figure plots the marginal effect of $\Delta N_i^{coworkers}$ in the specification described by Column (3) of Table 5. Bars denote 95% confidence intervals. The dependent variable in this estimation is the number of transactions (transformed using the inverse hyperbolic sine function) on each period for each user.

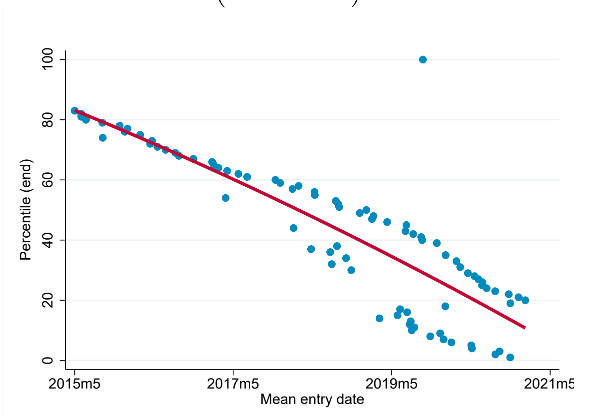
Figure I10: Entry and Diffusion Across and Within Networks of Relatives and Coworkers



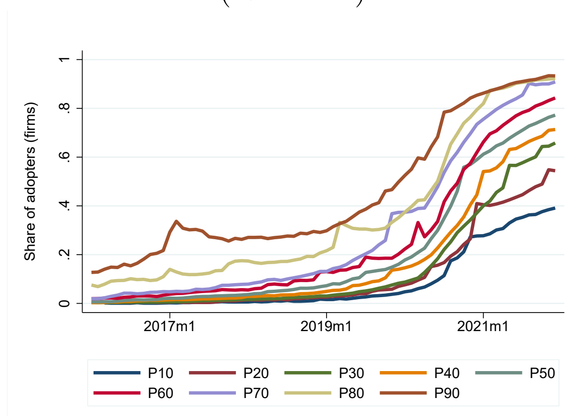
(a) Entry Across Networks (Relatives)



(b) Diffusion Within Network (Relatives)



(c) Entry Across Networks (Coworkers)



(d) Diffusion Within Network (Coworkers)

Notes: Panel (a) and Panel (c) show the timing of adoption across networks, defined as relatives and coworkers, respectively. They show the entry date (first time an individual within a network adopts the technology) across different percentiles of the distribution of networks. Percentiles are calculated in the period with highest adoption in the sample given the share of individuals that had adopted the technology. Panel (b) and Panel (d) use the same classification of percentiles to show the patterns of diffusion of the technology within networks.

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

	(1) Neighborhood	(2) Firm	(3) Family	(4) 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.

Table I2: Amount Transacted and Size of Network at Entry

Dependent variable: Amount transacted (logs)

	(1)	(2)	(3)
Size of Neighbors' Network at Entry	-3.125***		
	(0.005)		
Size of Coworkers' Network at Entry		-1.494***	
		(0.005)	
Size of Family Network at Entry			-1.475***
			(0.006)
Observations	27,655,426	13,243,530	12,000,848
Adjusted R^2	0.147	0.171	0.090
Network×Time 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.

Table I3: Changes in Number of Transactions and Network Changes—Alternative Transformations

Dependent variable: Δ Number of Transactions

(a) Logs	(1)	(2)	(3)	(4)
Δ Share Neighborhood Adopters	0.917***			0.833***
	(0.023)			(0.041)
Δ Share Coworkers Adopters		0.160***		0.175***
		(0.006)		(0.008)
Δ (Log) Wage		0.043***		0.048***
		(0.001)		(0.001)
Δ Share Relatives Adopters			0.401***	0.411***
			(0.008)	(0.009)
Observations	23,675,631	15,069,613	10,456,216	6,750,537
Adjusted R^2	0.018	0.022	0.019	0.023
RMSE	0.746	0.721	0.745	0.715
Time FE	Yes	Yes	Yes	Yes
(b) Davis & Haltiwanger				
Δ Share Neighborhood Adopters	0.916***			0.701***
	(0.026)			(0.046)
Δ Share Coworkers Adopters		0.236***		0.255***
		(0.006)		(0.008)
Δ (Log) Wage		0.042***		0.047***
		(0.001)		(0.001)
Δ Share Relatives Adopters			0.420***	0.426***
			(0.008)	(0.009)
Observations	25,632,610	16,208,557	11,275,971	7,230,892
Adjusted R^2	0.025	0.032	0.028	0.035
RMSE	0.795	0.770	0.788	0.760
Time FE	Yes	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. Standard errors are in parentheses. Extreme values (one and 99 percentile) were trimmed from the dependent variable. Results are robust to alternative transformations (Table 3) and to no trimming.

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

Dependent variable: Δ Value of Transactions

(a) Logs	(1)	(2)	(3)	(4)
Δ Share Neighborhood Adopters	1.095***			1.005***
	(0.037)			(0.066)
Δ Share Coworkers Adopters		0.202***		0.214***
		(0.009)		(0.013)
Δ (Log) Wage		0.076***		0.084***
		(0.001)		(0.002)
Δ Share Relatives Adopters			0.469***	0.492***
			(0.012)	(0.015)
Observations	23,683,186	15,040,729	10,455,020	6,734,204
Adjusted R^2	0.016	0.019	0.016	0.021
RMSE	1.186	1.157	1.185	1.149
Time FE	Yes	Yes	Yes	Yes
(b) Davis & Haltiwanger				
Δ Share Neighborhood Adopters	0.834***			0.733***
	(0.037)			(0.055)
Δ Share Coworkers Adopters		0.234***		0.250***
		(0.007)		(0.010)
Δ (Log) Wage		0.061***		0.067***
		(0.001)		(0.001)
Δ Share Relatives Adopters			0.417***	0.431***
			(0.009)	(0.011)
Observations	16,179,194	16,179,194	11,274,524	7,214,334
Adjusted R^2	0.029	0.029	0.025	0.032
RMSE	0.936	0.936	0.953	0.928
Time FE	Yes	Yes	Yes	Yes
(c) Inverse hyperbolic sine				
Δ Share Neighborhood Adopters	2.084***			1.208***
	(0.104)			(0.187)
Δ Share Coworkers Adopters		0.768***		0.831***
		(0.023)		(0.034)
Δ (Log) Wage		0.111***		0.125***
		(0.002)		(0.003)
Δ Share Relatives Adopters			1.061***	1.066***
			(0.029)	(0.035)
Observations	25,639,560	16,179,194	11,274,524	7,214,334
Adjusted R^2	0.027	0.035	0.031	0.040
RMSE	3.032	2.926	2.958	2.847
Time FE	Yes	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. Standard errors are in parentheses. Extreme values (one and 99 percentile) were trimmed from the dependent variables.

Table I5: Changes in Intensity of Usage using Logs and Network Changes

Dependent variable: $\Delta \ln$ *Value of Transactions*

(a) Neighbors Network	(1)	(2)	(3)	(4)
Δ Share Neighborhood Adopters	3.730***	0.917***	3.075***	0.587***
	(0.014)	(0.023)	(0.017)	(0.023)
Δ (Log) COVID-19 Cases			0.006***	
			(0.000)	
Observations	23,675,631	23,675,631	20,537,868	23,675,631
Adjusted R^2	0.002	0.018	0.002	0.020
Time FE	No	Yes	No	Yes
Cohort FE	No	No	No	Yes

(b) Coworkers Network	(1)	(2)	(3)	(4)
Δ Share Coworkers Adopters	0.494***	0.160***	0.436***	0.123***
	(0.005)	(0.006)	(0.006)	(0.006)
Δ (Log) COVID-19 Cases			0.013***	
			(0.000)	
Δ (Log) Wage	0.053***	0.043***	0.057***	0.043***
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	15,069,613	15,069,613	12,799,581	15,069,613
Adjusted R^2	0.001	0.022	0.001	0.024
Time FE	No	Yes	No	Yes
Cohort FE	No	No	No	Yes

(c) Family Network	(1)	(2)	(3)	(4)
Δ Share Relatives Adopters	0.734***	0.420***	0.568***	0.378***
	(0.007)	(0.009)	(0.009)	(0.009)
Δ (Log) COVID-19 Cases			0.013***	
			(0.000)	
Observations	10,456,216	6,750,537	5,645,335	6,750,537
Adjusted R^2	0.001	0.023	0.002	0.026
Time FE	No	Yes	No	Yes
Cohort FE	No	No	No	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. Standard errors are in parentheses.

Table I6: Changes in Intensity of Usage using Davis and Haltiwanger (1992) and Network Changes

Dependent variable: %Δ Value of Transactions

(a) Neighbors Network	(1)	(2)	(3)	(4)
Δ Share Neighborhood Adopters	2.853***	0.918***	4.442***	0.246***
	(0.019)	(0.030)	(0.022)	(0.029)
Δ (Log) COVID-19 Cases			0.011***	
			(0.000)	
Observations	25,632,610	25,632,610	21,858,049	25,632,610
Adjusted R^2	0.001	0.023	0.002	0.028
Time FE	No	Yes	No	Yes
Cohort FE	No	No	No	Yes
(b) Coworkers Network				
Δ Share Coworkers Adopters	0.716***	0.229***	0.682***	0.159***
	(0.007)	(0.007)	(0.007)	(0.007)
Δ (Log) COVID-19 Cases			0.023***	
			(0.000)	
Δ (Log) Wage	0.068***	0.061***	0.080***	0.061***
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	16,208,557	16,208,557	13,482,422	16,208,557
Adjusted R^2	0.001	0.029	0.002	0.034
Time FE	No	Yes	No	Yes
Cohort FE	No	No	No	Yes
(c) Family Network				
Δ Share Relatives Adopters	0.781***	0.416***	0.708***	0.338***
	(0.009)	(0.009)	(0.009)	(0.009)
Δ (Log) COVID-19 Cases			0.023***	
			(0.000)	
Observations	11,275,971	11,275,971	9,452,870	11,275,971
Adjusted R^2	0.001	0.025	0.001	0.029
Time FE	No	Yes	No	Yes
Cohort FE	No	No	No	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. Standard errors are in parentheses.

Table I7: Changes in Intensity of Usage using Inverse Hyperbolic Sine and Network Changes

Dependent variable: %Δ Value of Transactions

(a) Neighbors Network	(1)	(2)	(3)	(4)
Δ Share Neighborhood Adopters	5.168***	2.077***	15.829***	-0.537***
	(0.078)	(0.104)	(0.080)	(0.103)
Δ (Log) COVID-19 Cases			0.038***	
			(0.001)	
Observations	25,632,610	25,632,610	21,858,049	25,632,610
Adjusted R^2	0.000	0.026	0.003	0.034
Time FE	No	Yes	No	Yes
Cohort FE	No	No	No	Yes

(b) Coworkers Network	(1)	(2)	(3)	(4)
Δ Share Coworkers Adopters	2.432***	0.764***	0.682***	0.498***
	(0.024)	(0.023)	(0.007)	(0.023)
Δ (Log) COVID-19 Cases			0.023***	
			(0.000)	
Δ (Log) Wage	0.106***	0.112***	0.080***	0.111***
	(0.003)	(0.002)	(0.001)	(0.002)
Observations	16,208,557	16,208,557	13,482,422	16,208,557
Adjusted R^2	0.001	0.035	0.002	0.043
Time FE	No	Yes	No	Yes
Cohort FE	No	No	No	Yes

(c) Family Network	(1)	(2)	(3)	(4)
Δ Share Relatives Adopters	2.141***	1.063***	2.257***	0.767***
	(0.029)	(0.029)	(0.029)	(0.029)
Δ (Log) COVID-19 Cases			0.082***	
			(0.001)	
Observations	11,275,971	11,275,971	9,452,870	11,275,971
Adjusted R^2	0.001	0.030	0.002	0.037
Time FE	No	Yes	No	Yes
Cohort FE	No	No	No	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. Standard errors are in parentheses.

Table I8: Weighted Changes in Intensity of Usage and 2021 Network Changes)

Dependent variable: %Δ Value of Transactions

	(1) Logs	(2) Davis & Haltiwanger	(3) Inverse hyperbolic sine
Δ Share Adopters in 2021 Network	2.010***	1.766***	4.922***
	(0.012)	(0.008)	(0.030)
Observations	23,680,468	25,635,895	25,635,895
Adjusted R^2	0.018	0.026	0.028
Time 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. Standard errors, clustered by individual, are in parentheses.

Table I9: Intensity of Usage and Changes in Coworkers' Network After a Mass Layoff

Dependent Variable: Δ Value of transactions (inverse hyperbolic sine)

	(1)	(2)	(3)	(4)
$\Delta N_i^{coworkers}$	4.879***	3.166***	2.048***	1.478**
	(0.509)	(0.552)	(0.588)	(0.595)
$\Delta \ln wage_i$		0.842***	0.739***	0.871***
		(0.157)	(0.157)	(0.166)
$\Delta Covid_i$			0.333***	0.247***
			(0.065)	(0.076)
Observations	1,554	1,554	1,554	1,554
Adjusted R^2	0.063	0.111	0.127	0.168
Time FE	No	Yes	Yes	Yes
Cohort FE	No	No	No	Yes

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 I10: Changes in Intensity of Usage and 2021 Network Changes

Dependent variable: %Δ Number of Transactions

	Logs (1)	Davis & Haltiwanger (2)	IHS (3)
Δ Share Adopters in 2021 Network	1.825***	1.870***	1.987***
	(0.007)	(0.007)	(0.007)
Observations	23,672,905	25,628,936	25,628,936
R-squared	0.022	0.029	0.026
Time 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. Standard errors, clustered by individual, are in parentheses.

Table I11: Intensity of Usage and Changes in Coworkers' Network After a Mass Layoff (within coworkers)

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

	(1)	(2)	(3)	(4)
$\Delta N_i^{coworkers}$	1.338***	1.001***	0.888***	0.831***
	(0.138)	(0.159)	(0.176)	(0.185)
$\Delta \ln wage_i$		0.363***	0.353***	0.370***
		(0.041)	(0.042)	(0.046)
$\Delta Covid_i$			0.034*	0.036
			(0.020)	(0.024)
Observations	1,554	1,554	1,554	1,554
Adjusted R^2	0.057	0.095	0.096	0.082
Time FE	No	Yes	Yes	Yes
Cohort FE	No	No	No	Yes

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, to December 2021. Standard errors, clustered by individual, are in parentheses.

J Details on Mass Layoffs

This section provides additional details on the choices made to construct the variables and sample used in Section 8.4

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

Table J12: 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.

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 J12.

Definition of variables We construct several variables that are used in equation (44). 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.

- $\Delta Covid_i$ 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 (45) 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.