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

Fernando Alvarez
University of Chicago

David Argente
Yale University

Francesco Lippi
LUISS and EIEF

Esteban Méndez
Central Bank of Costa Rica

Diana Van Patten
Yale University

January 28, 2025

Abstract

We develop a dynamic model of technology adoption featuring strategic complementarities: the benefits of the technology increase with the number of adopters. We show that complementarities give rise to gradual adoption, multiple equilibria, multiple steady states, and suboptimal allocations. We study the planner's problem and its implementation through adoption subsidies. We apply the theory to SINPE Móvil, an electronic peer-to-peer payment app developed by the Central Bank of Costa Rica, currently adopted by 60 percent of adults. Using transaction-level data and several administrative data we estimate sizable complementarities. In our calibrated model the optimal subsidy pushes the economy to universal adoption.

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

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

1 Introduction

We study the diffusion of a new technology in an economy characterized by strategic complementarities. These complementarities occur because the benefits that agents derive from the technology increase with the number of users –a phenomenon long recognized in the applied literature on technology diffusion (see Griliches (1957); Mansfield (1961)). Progress in this research area is hindered by the challenges that arise when modeling adoption dynamically –a large state space, non-linear decisions, multiple equilibria–, and by the lack of detailed data on technology diffusion. We develop a new tractable model of technology adoption and apply it to the diffusion of a new payment app, SINPE, a digital application created by the Central Bank of Costa Rica that allows instantaneous P2P transfers between bank account holders in the network. By its nature, the usefulness on this app depends on others joining the network. We aim to quantify the value of this complementarity using granular data from SINPE and other sources.[1] We use the model to discuss equilibrium existence, multiplicity of equilibrium paths, multiplicity of stationary equilibria, and the local stability of stationary equilibria (see e.g., Matsuyama (1991)). We characterize the planner's problem and its implementation through subsidies, and use a calibrated version of the model to analyze the optimal policy.

The model assumes that the benefits of the technology at time t depend on the number of agents who have adopted it, $N(t)$, and on an idiosyncratic persistent random component, $x(t)$. In particular, we assume that the flow benefit of the app is proportional to the product between these variables, $x(t)N(t)$, so that an agent is more likely to adopt if her private needs for it are high (a high x) and/or when more agents use the app (higher N). A single parameter, controlling the intensity of this interaction effect, measures the strength of the strategic complementarities. A high value of x also implies that an agent will use the technology more intensively, a feature that we leverage when calibrating the model to the data where we observe both adoption as well as the intensity of usage. Adoption entails a fixed (once and for all) cost and agents choose when to adopt taking the aggregate path of adoption as given. We show that when the idiosyncratic benefits are random the equilibrium features gradual adoption through a simple mechanism: agents wait for others to adopt.[2]

[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). The app was launched in May 2015 and over 60% of the adult population used it in 2021, with about 10% of the country's GDP transacted via SINPE. See Björkegren (2018) for a related network-goods analysis using data on mobile phones adoption in Rwanda.

[2] We also analyzed a model where x is heterogeneous across agents but fixed through time. This model features no interesting dynamics: agents with a sufficiently high x immediately adopt, and all others do not. It is not a model of gradual diffusion, but one of "jumps." Instead, the stochastic model features gradual adoption given the option value of waiting for a high draw of the idiosyncratic benefit. The SINPE technology

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 (arbitrary) measure of agents endowed with the technology, which serves as the initial condition of the equilibrium. Aggregation of the adoption decisions across agents yields a path for the fraction of agents that use the technology at each time t, $N(t)$. Given the initial and terminal conditions, the equilibrium has a classic fixed point structure: the optimal decision path ($\bar{x}$) depends on the aggregate path (N), and vice-versa.

The model yields three main results, each summarized by a theorem. We show that the optimal adoption rule for each agent, summarized by the threshold path $\bar{x}$, is a decreasing functional of the path of adoption N. The strength of this effect depends on the parameter that controls the strategic complementarity. Likewise, we show that the adoption path N is a decreasing functional of the path $\bar{x}$, for any initial distribution of adopters. An equilibrium is a fixed point given by the composition of these two functionals. The first theorem establishes the existence, and possibly the multiplicity, of dynamic equilibria. These equilibria form a non empty lattice, i.e., they are ordered so that there is a "largest one", N^H, and a "smallest" one, N^L. The adoption path of the largest equilibrium is above the smallest one at every point in time, $N^H(t) > N^L(t)$, for all t. More equilibria may exist and are bracketed between these ones (the paths of different equilibria do not cross). We establish these results using the monotone comparative statics logic by Milgrom and Shannon (1994), and Tarski's fixed point theorem. We show that there is a critical mass of adopters $\underline{N}_0$ such that, if the initial measure of adopters is below $\underline{N}_0$, then there is an equilibrium where no one will adopt eventually. We also study stationary equilibria, i.e. equilibria where N is constant through time, and show that, besides the stationary equilibrium with no adoption, the model has two additional interior stationary equilibria, which we label low- and high-adoption.

The second theorem characterizes the stability of the stationary equilibria by means of a perturbation analysis with respect to the initial condition, assumed to be one of the two interior equilibria. The analysis is non-trivial because it involves the linearization of an infinite dimensional system: the distribution of adopters. We handle the problem by leveraging techniques from the Mean Field Game (MFG) literature (e.g., Alvarez et al., 2023a; Auclert et al., 2022; Bilal, 2023), which determines the local stability by inspecting the eigenvalues of a linear operator. One novelty compared to the MFG problem studied in Alvarez et al. (2023a) is the possibility of *multiple* stationary equilibria. The stability condition then depends on the particular equilibrium that is chosen. We find that the high-adoption equilibrium is locally stable, while the low-adoption is unstable, a feature that leads us to discard it from the analysis.

was indeed adopted gradually, as shown in Figure 7.

We show how to characterize efficient allocations, taking into account the dynamics across the entire network. Equilibria are socially inefficient because agents do not internalize the fact that when they adopt they benefit all agents who already have the technology. The third theorem shows how to decentralize the planner's solution using a simple tool: a time-varying subsidy paid to those that use the technology.

We then leverage a comprehensive set of data collected since SINPE was created to analyze the dynamics of adoption and usage, to document the presence of strategic complementarities, and to discipline the parametrization of the model. Our baseline analysis links data on users –both receivers and senders– within their employer-employee network.[3] We identify the presence of strategic complementarities using arguably exogenous variations in the network size due to mass layoffs. We document a causal relation between the share of agents who have adopted (N) and usage of the app, both at the extensive margin as well as at the intensive margin: a sudden decrease of the network size lowers the probability of adoption and lowers the intensity of use.[4] This effect persists across a battery of ways to define usage and networks. It also emerges after using a leave-one-out instrument and following a balanced panel of adopters to address concerns regarding selection.

We match the theory with the data in a quantitative analysis where we calibrate the model using key moments from the data with the objective to compute the optimal adoption subsidy. To capture the initial gradual diffusion of the technology, observed in each network, we supplement the model with a layer of slow-information diffusion following the seminal work of Bass (1969). The strength of the strategic complementarities is chosen using the information retrieved from the mass layoffs described above. The calibrated model shows that the optimal subsidy speeds up adoption by the agents and ultimately pushes the economy towards universal adoption of the payment app.

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

[3]Individual-to-individual transactions account for over 95% of all transactions, regardless of the time period considered. We find that 44% of all SINPE transactions occur between coworkers. Family networks and spatial "neighborhood" networks are also considered for robustness.

[4]Namely, we focus on networks of coworkers and examine the effect of network changes on the intensity of the app's usage and its adoption for workers displaced by a mass layoff. By analyzing the usage intensity of workers who had already adopted the app prior to being displaced, we are able to isolate the influence of strategic complementarities rather than the effects of learning.

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

The paper also deals with technical issues of multiplicity and stability that have plagued the economic geography literature. Recent papers have developed algorithms that exploit the super- or sub-modularity of the objective function based on Tarski's theorem (Jia, 2008; Arkolakis et al., 2023). Our approach also leverages the monotonicity of our problem, but does so for an analysis of dynamic stability as a criterion to select an equilibrium and develops the planning problem to study efficiency.

The paper is organized as follows. The next section presents the model, Section 3 discusses the different types of equilibria that exist. Section 4 uses a perturbation method to inspect the stability of the stationary equilibria. Section 5 discusses the planning problem. Section 6 presents the data and documents the non-negligible role of strategic complementarities in the use and adoption of SINPE. A calibrated version of the model is used in Section 7 to discuss the optimal subsidy for the efficient adoption of SINPE.

2 The Model

This section presents a tractable model of technology adoption within a "network" of agents. The model fits alternative notions of network, later discussed in the empirical analysis, such as a group of co-workers, households living in the same neighborhood, or a (broad) notion of family members. The network is populated by a continuum of agents who differ in the potential benefits from adopting the technology. Let $N(t) \in [0, 1]$ be the fraction of agents who have adopted at time $t \in [0, T]$. The flow benefit at time t for an agent who has already adopted the technology is

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

where $\theta_0, \theta_n > 0$ are parameters. x is stochastic process, independent across agents, with variance σ^2 per unit of time, no drift, and reflecting barriers at $x = 0$ and $x = U$, so that

$dx = \sigma dW$ where W is a standardized Brownian motion. Later on –see equation (29)– we allow agents to choose the intensity of technology use in each period, in which case equation (1) gives the optimal value of such problem. We let $c > 0$ be the fixed cost of adopting the technology and $r > 0$ be the time discount rate. With probability ν per unit of time agents die, so that agents discount time at rate $\rho \equiv r + \nu$. Dead agents are replaced by newborns without the technology and an x drawn from the invariant density $f(x) = 1/U$ for $x \in [0, U]$, which is uniform because of the reflecting barriers assumption.

2.1 Individual Decisions, Aggregation, Equilibrium

We next describe the agent's optimal decision as a function of the whole path of aggregate adoption $N : [0, T] \to [0, 1]$, discuss how to aggregate individual decision to compute the aggregate path of adoption, and define the equilibrium.

Let $a(x, t)$ be the value function of an agent who has adopted 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 whole path N as given.

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

An agent with state x, who has not yet adopted at time t, has a value function $v(x, t)$ that solves the 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 and on calendar time t (the latter because of the dynamics of $N(t)$).

Discretized Model. For future use we introduce a discretized version of the model defined by positive integers I, J that determine the step size 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.[5] Next we state a preliminary

[5] See Definition 3 in Appendix A for a detailed definition.

result to establish that we can represent the optimal adoption rule at time t as a threshold rule, $\bar{x}(t)$.

LEMMA 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 model.

Let us denote $a_T(x) = a(x, T)$ and $v_T(x) = v(x, T)$, where both functions depend only on the constant $\bar{N}$, and 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 2.2. We denote the optimal threshold as $\bar{x} = \mathcal{X}(N; D_T)$, so that $\bar{x} : [0, T] \to [0, U]$.

Aggregation. Given the individual decision rule we can compute the implied path for the fraction of adopters, N. We start by defining the probability that an agent at s with state $x(s) = x$ survives until time t, while the value of her state remains below $\bar{x}$ during this period:

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

For an agent who at time s has $x \leq \bar{x}(s)$, the value of $P(x, s, t; \bar{x})$ gives the probability that the agent will survive up to t without adopting. Let $m_0(x)$ be the density of the agents at time $t = 0$ without the technology. Given the assumption about x, we require $0 \leq m(x) \leq 1/U$ for all $x \in [0, U]$. The fraction of agents who have adopted the technology at time t is

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

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 the discretized version of the model. We let $\mathcal{N}(\bar{x}; m_0)$ be the path of N as a function of $\bar{x}$ (the path of the adoption threshold) and of the initial condition m_0.

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 $D_T \equiv a_T - v_T$. To emphasize the backward looking nature of $\mathcal{N}$, note that it propagates the initial condition m_0. We then have the following definition.

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

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

and where $\bar{x}^* = \mathcal{X}\left(N^*; D_T\right)$.

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.

2.2 A Recursive Formulation of the Equilibrium

This section derives a recursive representation of the equilibrium that will be useful to study the local stability of the equilibrium and to study the planning problem.

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.[6] 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

[6] We derive these equations and their boundaries in Appendix F.

barrier. We are looking for a classical solution that satisfies:

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

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

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

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

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

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

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

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

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

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

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

3 Equilibria

In this section we establish equilibrium existence, the possibility of multiple equilibria and illustrate an iterative procedure to compute the equilibrium numerically. We also discuss equilibria with no adoption, i.e. situations in which given an initial condition m_0, no one will use the technology eventually. We conclude by discussing stationary equilibria.

3.1 Monotonicity and Existence of Equilibrium

The next lemma 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 lemma shows that an agent facing larger values of θ_0 and/or θ_n, will also adopt earlier.

LEMMA 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 $\bar{x}'(t) \leq \bar{x}(t)$ for all $t \in [0, T]$. Moreover, let $\theta \equiv (\theta_0, \theta_n)$ with the corresponding optimal threshold path $\bar{x}$. If $\theta' \geq \theta$ then $\bar{x}'(t) \leq \bar{x}(t)$ for all $t \in [0, T]$.

Lemma 2 also holds in the discretized version of the model.[7] The proof holds as 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 given the 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.

LEMMA 3. Fix m_0 and consider two paths for the thresholds $\bar{x}$ and $\bar{x}'$, 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 path $\bar{x}$, 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 $\mathcal{N}$, proven in Lemma 2 and Lemma 3, to establish through equation (6) that $\mathcal{F}$ is monotone. This allows us to use Tarski's theorem and establish the existence, and possibly the multiplicity, of equilibria. For technical reasons, the theorem applies to the finite-horizon discretized-version of the model introduced in

[7] For instance, it holds for a finite difference approximation, which we use for some computations, and which converges to the continuous-time version.

Section 2.1.[8]

THEOREM 1. Consider a finite horizon, discrete time - discrete state version of the model and $\theta_n \geq 0$. Fix an initial condition $m_0 \in \mathbb{R}^I_+$ and a terminal value function $D_T \in \mathbb{R}^I_+$.
(i) The equilibria of this model are a non-empty lattice. Hence the model has a smallest equilibrium, $\{\bar{x}^L, N^L\}$, and a largest one, $\{\bar{x}^H, N^H\}$, and any other equilibrium path $\{\bar{x}, N\}$ satisfies $N^L \leq N \leq N^H$ and $\bar{x}^L \geq \bar{x} \geq \bar{x}^H$ for all $t \in [0, T]$.
(ii) Let $\theta' \geq \theta$, and $m_0' \leq m_0$ for all $x \in [0, U]$. 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 $N' \geq N$ for all $t \in [0, T]$.

The first statement of the theorem establishes the 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 arbitrarily large finite horizon, T. An important consequence of the theorem is that the equilibrium set, given the initial distribution of non-adopters m_0 and the terminal valuation $D_T \equiv a_T - v_T$, is a lattice. We can compute the value of the extreme equilibria (i.e., the smallest and the largest) 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 focuses on the "high-adoption" equilibrium and establishes a useful comparative statics result: considering a larger θ, or a "smaller" m_0 (more agents endowed with the app at time zero), leads to more adoption.

3.2 No-Adoption Equilibrium

The setup may feature an equilibrium with zero adoption, i.e., $\bar{x}(t) = U$ for all t. For simplicity we focus on the case where $T = \infty$. This case is particularly easy because agents' decisions are in a corner. We characterize the basin of attraction for such equilibrium, i.e., we find a threshold for the number of adopters $\underline{N}$, such that a no-adoption equilibrium exists if and only if at $t = 0$ the mass of agents with the technology is smaller than $\underline{N}$.

PROPOSITION 1. 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

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

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

[8] See Definition 3 in Appendix A. The reason is the completeness of the lattice in which $\mathcal{F}$ is defined.

Note that $\underline{N} > 0$ if and only if $\frac{\rho c}{U} > \theta_0 [1 + g(\eta U)]$. 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 lower (upper) boundary is reached as $\sigma \to 0$ ($\sigma \to \infty$).
(ii) $\underline{N}$ is a decreasing function of θ_n.

An immediate corollary of this proposition is that $m_0(x) = 1/U$ is an invariant distribution provided that $\underline{N} \geq 0$, i.e., if the economy starts with no adoption, then it may remain in that equilibrium forever (no adoption is a stationary equilibrium). That $\underline{N} > 0$ requires θ_0 to be small is easily understood: if θ_0 is large agents with a high x will find it profitable to adopt regardless of what the others choose. Likewise, that $\underline{N} > 0$ is increasing in σ implies that if agents are hit by larges shocks the no-adoption equilibrium is more likely to occur. This result follows because, for a given U, a large σ makes the reversion to the mean faster, lowering the benefit of adoption. Finally, if θ_n is large then it is more profitable to coordinate on high N and the basin of attraction of the no-adoption equilibrium is smaller.

3.3 Stationary Equilibria

In this section we let $T = \infty$ and analyze the stationary equilibria of the model. We look for an initial condition m_0, such that the distribution is invariant, so that both $\bar{x}(t) = \bar{x}_{ss}$ and $N(t) = N_{ss}$ are constant through time. We will show that convergence to the stationary equilibrium must be gradual, i.e., that it is not possible to "jump" to the equilibrium given a generic initial condition in the model where $\sigma > 0$.[9]

A stationary equilibrium is given by two constant values of N_{ss} and $\bar{x}_{ss}$ that solve the time-invariant version of the partial differential equations presented in Section 2.2. From a mathematical point of view the equilibrium is a fixed point. Given N_{ss}, $\tilde{D}$ and $\bar{x}_{ss}$ solve:

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

[9] An immediate jump to the stationary equilibrium might instead occur in a model with $\sigma = 0$ (See the Online appendix J of Alvarez et al. (2023b)).

Conversely, given $\bar{x}_{ss}$, the density $\tilde{m}$ solves

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

Notice that the (stationary) equilibrium $\tilde{m}(x)$ and $\bar{x}_{ss}$ solve the fixed point

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

We begin by solving for $\tilde{D}(x)$ and $\bar{x}_{ss}$ given a value for N_{ss} (see Appendix B.1 for details). Using the solution for $\tilde{D}$ we can solve for $\mathcal{X}_{ss} : [0,1] \to [0,U]$, a function that gives the *optimal* stationary threshold as a function of a given N_{ss}. The monotonicity properties of the function $\tilde{D}$ on the parameters N_{ss}, θ_n, c and θ_0 give the following characterization of the threshold $\mathcal{X}_{ss}$.

Lemma 4. 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, $\mathcal{X}_{ss}^{-1}$, given by:

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

$$\bar{A}_1 \equiv \frac{1}{\eta} \frac{\left(1 - e^{-\eta U}\right)}{\left(e^{-\eta U} - e^{\eta U}\right)}, \; \bar{A}_2 \equiv \frac{1}{\eta} \frac{\left(1 - e^{\eta U}\right)}{\left(e^{-\eta U} - e^{\eta U}\right)} \text{ and } \eta \equiv \sqrt{2\rho/\sigma^2}. \tag{16}$$

Note that, from the expansion given in Lemma 4, 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 equation for $\tilde{m}(x)$, given a barrier $\bar{x}_{ss}$ subject to an exit point and to the boundary conditions coming from the reflecting barriers. We denote the corresponding value of the fraction that have adopted as $\mathcal{N}_{ss}(\bar{x}_{ss})$. Solving this equation we obtain

$$\mathcal{N}_{ss}(\bar{x}_{ss}) = 1 - \frac{\bar{x}_{ss}}{U} + \frac{\tanh(\gamma \bar{x}_{ss})}{U\gamma} \text{ where } \gamma \equiv \sqrt{2\nu/\sigma^2}. \tag{17}$$

Inspection of equation (17) yields the following characterization of $\mathcal{N}_{ss}$.

Lemma 5. 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: $\mathcal{N}_{ss}(\bar{x}) = 1 - \frac{\bar{x}_{ss}}{U} + \frac{\sigma}{U\sqrt{2\nu}} + o(\sigma)$.

As is intuitive, the value of $\mathcal{N}_{ss}(\bar{x}_{ss})$ is *decreasing* in the level of the barrier $\bar{x}$. The system given by equation (16) and equation (17) determines $\bar{x}_{ss}$ and N_{ss}. In particular, a stationary equilibrum is described by the pair $\{\bar{x}_{ss}, N_{ss}\}$, which solves

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

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

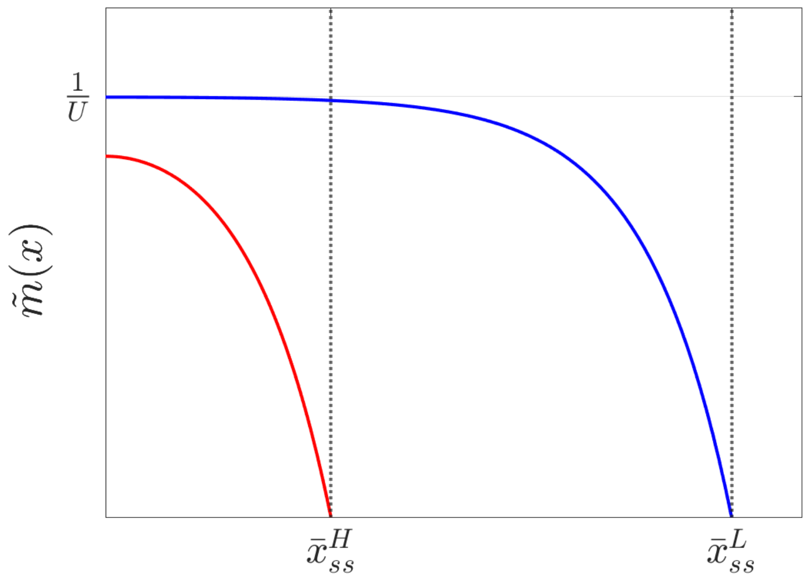
Proposition 2. Assume that $\nu > 0$ and that the parameters θ_0, θ_n, c and ρ are such that there are two interior stationary equilibria in the deterministic case of $\sigma = 0$, and label them as $\bar{x}_{ss}^H < \bar{x}_{ss}^L$. Then, (i) there exists a $\bar{\sigma} > 0$ such that for all $\sigma \in (0, \bar{\sigma})$ there are two interior stationary equilibria with $\bar{x}_{ss}^H < \bar{x}_{ss}^L$. (ii) The threshold for each stationary equilibria is continuous with respect to σ at $\sigma = 0$. (iii) The sign of the comparative static differs across stationary equilibria, with

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

The proposition shows that the high adoption stationary equilibrium behaves in an intuitive way, with more adoption (a lower $\bar{x}_{ss}^H$) associated with a smaller adoption cost (c), or with a larger intrinsic value of the technology (θ_0). The comparative statics for the low adoption stationary state are just the opposite: adoption is higher as the adoption cost increases. The latter (unrealistic) feature, and the unstable nature of the low adoption equilibrium (see the next section), will lead us to focus on the high adoption equilibrium in our quantitative analysis.

A notable feature of the stationary distribution of non-adopters is that in the invariant equilibrium there are agents with low benefits, namely with $x(t) < \bar{x}_{ss}$, who have the technology (provided $\sigma > 0$). These are agents who adopted the technology in the past (for some

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



High and Low Adoption Stationary Equilibria

$t' < t$ when $x(t') > \bar{x}(t')$, and whose x decreased over time. As a result, $m(x) < 1/U$ when $\sigma > 0$, and the density of non-adopters below $\bar{x}_{ss}$ is not uniform. Given that the density takes time to adjust, the stochastic model features dynamics in the adoption of a new technology: it takes time to change from the initial distribution to the invariant distribution, as agents adopt when $x(t) > \bar{x}(t)$ and it takes times for the $x's$ to crawl back below the stationary threshold. Figure 1 shows the densities of the invariant distribution of the high- and low-adoption equilibria, illustrating that both equilibria have adopters below the (respective) stationary threshold.

4 Stability of Stationary Equilibria

In this section we analyze the local stability of the stationary equilibria. We explore the question by perturbing the stationary distribution of adopters, using techniques from the Mean Field Game literature developed in Alvarez, Lippi and Souganidis (2023a). For this purpose, we use the equilibrium Definition 2. This dynamical system is infinite-dimensional because the state, at every time t, is given by the entire density $m(x, t)$.

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

We begin the analysis with the approximation of $\bar{x}(t) = \mathcal{X}(N)(t)$. That is, we study how perturbing the aggregate path of adoption N leads to adjusting the decision rule for threshold path $\bar{x}$. To do this, we take the directional derivative (Gateaux) with respect to an arbitrary perturbation n of a constant path N. In particular, we consider paths defined by $N(t) = N_{ss} + \epsilon\, n(t)$ around the stationary value N_{ss}. We denote this Gateaux derivative by $\bar{y}$, so that $\bar{x}(t) \approx \bar{x}_{ss} + \epsilon\bar{y}(t)$.

Lemma 6. Fix a stationary equilibrium with interior $\bar{x}_{ss}$, and its corresponding N_{ss}. Let D_T be equal to the stationary value function $\tilde{D}$ corresponding to that stationary equilibrium. Let $n : [0, T] \to \mathbb{R}$ be an arbitrary perturbation. Then

$$\bar{y}(t) \equiv \lim_{\epsilon \downarrow 0} \frac{\mathcal{X}(N_{ss} + \epsilon n; \tilde{D})(t) - \mathcal{X}(N_{ss}; \tilde{D})(t)}{\epsilon} = \frac{\theta_n}{\tilde{D}_{xx}(\bar{x}_{ss})} \int_t^T G(\tau - t) n(\tau) d\tau, \tag{18}$$

where

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

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

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

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

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

Lemma 7. Fix the interior threshold $\bar{x}_{ss}$ of a stationary equilibrium and its corresponding N_{ss}, and let $\tilde{m}$ be the corresponding invariant distribution of non-adopters. Let $\omega : [0, \bar{x}_{ss}] \to \mathbb{R}$ be an arbitrary perturbation to the distribution, and let $\bar{y} : [0, T] \to \mathbb{R}$ be an arbitrary perturbation of the threshold. Then

$$n(t) \equiv \lim_{\epsilon \downarrow 0} \frac{\mathcal{N}(\bar{x}_{ss} + \epsilon \bar{y}; \tilde{m} + \epsilon w)(t) - \mathcal{N}(\bar{x}_{ss}; \tilde{m})(t)}{\epsilon}$$
$$= n_0(\omega)(t) + \frac{\tilde{m}_x(\bar{x}_{ss})\sigma^2}{\bar{x}_{ss}} \int_0^t J(t-\tau)\bar{y}(\tau)d\tau \tag{19}$$

where

$$J(s) = \sum_{j=0}^{\infty} e^{-\mu_j s} \text{ with } \mu_j = \nu + \frac{1}{2}\sigma^2 \left(\frac{\pi(\frac{1}{2} + j)}{\bar{x}_{ss}} \right)^2 \tag{20}$$

$$n_0(\omega)(t) \equiv -\sum_{j=0}^{\infty} \frac{\bar{x}_{ss}}{\pi(\frac{1}{2} + j)} \frac{\langle \varphi_j, \omega \rangle}{\langle \varphi_j, \varphi_j \rangle} e^{-\mu_j t}, \tag{21}$$

$$\varphi_j(x) \equiv \sin\left(\left(\frac{1}{2} + j \right) \pi \left(1 - \frac{x}{\bar{x}_{ss}} \right) \right) \text{ for } x \in [0, \bar{x}_{ss}] \tag{22}$$
$$\frac{\langle \varphi_j, \omega \rangle}{\langle \varphi_j, \varphi_j \rangle} = \frac{2}{\bar{x}_{ss}} \int_0^{\bar{x}_{ss}} \varphi_j(x)\omega(x)dx \text{ and } \tilde{m}_x(\bar{x}_{ss}) = -\frac{\gamma}{U}\tanh(\gamma \bar{x}_{ss}).$$

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

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

The effect of the perturbation, $\bar{y}$, on the path of the threshold, $\bar{x}(s)$, is captured by the second term in equation (19). This term gives the effect at time t on the path $N(t)$ of a perturbation of the threshold rule $\bar{x}$, keeping the initial condition $\tilde{m}$ fixed. Also, consistent with our general result for $\mathcal{N}$, the effect of the threshold is negative, as $J > 0$ and $\tilde{m}_x(\bar{x}_{ss}) < 0$.

For future reference it is useful to understand the behavior of $n_0(t)$ as function of time. In particular, the rate at which the perturbation ω to the initial distribution converges back to the stationary distribution, while keeping $\bar{x}(t) = \bar{x}_{ss}$. This rate is given by the value of $\mu_0 = \nu + \frac{\sigma^2}{8}\left(\frac{\pi}{\bar{x}_{ss}}\right)^2$, i.e., the dominant eigenvalue.[10]

The next step is to use the last two lemmas 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 (18) and equation (19) to arrive to a single linear equation that $n(t)$ must solve as a function of ω.

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

$$n(t) = n_0(\omega)(t) + \Theta \int_0^T K(t,s)n(s)ds, \tag{23}$$

where $n_0(\omega)(t)$ is given in Lemma 7 and $\Theta \equiv \frac{\tilde{m}_x(\bar{x}_{ss})\sigma^2\theta_n}{\bar{x}_{ss}\tilde{D}_{xx}(\bar{x}_{ss})} > 0$. The kernel K is given by

$$K(t,s) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} c_j e^{-\mu_i t - \psi_j s}\left[\frac{e^{(\mu_i+\psi_j)\min\{t,s\}} - 1}{\mu_i + \psi_j}\right] > 0. \tag{24}$$

Moreover, $\text{Lip}_K \equiv \sup_t \int |K(t,s)|ds \leq \left(\frac{\bar{x}_{ss}^2}{\sigma^2}\right)^2$. Furthermore, if $\Theta\,\text{Lip}_K < 1$ there exists a unique bounded solution to equation (23) which is the limit of

$$n = \left[I + \Theta\mathcal{K} + \Theta^2\mathcal{K}^2 + \dots\right] n_0(\omega) \quad \text{where} \quad \mathcal{K}(g)(t) \equiv \int_0^T K(t,s)g(s)ds, \tag{25}$$

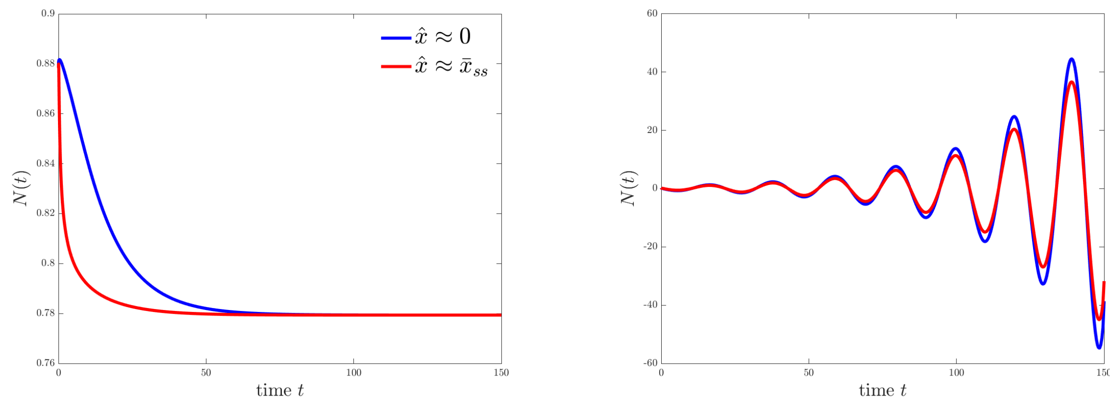
and where $\mathcal{K}^{j+1}(g)(t) \equiv \int_0^T K(t,s)\,\mathcal{K}^j(g)(s)\,ds$ for any bounded $g : [0,T] \to \mathbb{R}$.

A few remarks are in order. First, note that K depends on θ_n as μ_j, ψ_j are a function of $\bar{x}_{ss}$, which is itself a function of θ_n. The coefficient Θ depends on θ_n directly and indirectly via $\bar{x}_{ss}$. Hence equation (23) depends on which stationary equilibrium we focus on. Second, if

[10] The proof is in Appendix C.2 and resembles the one for the previous proposition.

we discretize time so that $t \in \{\Delta_t(j-1) : j = 1, \dots, J\}$ for $\Delta_t = \frac{T}{J-1}$, as done in Section 2.1, then the operator $\mathcal{K}$ is a $J \times J$ matrix with elements $K(t_i, t_j)$, and n_0, n are $J \times 1$ vectors, so that equation (23) becomes the linear equation $n = n_0 + \Theta\mathcal{K}\, n$. Third, the fact that $\Theta\mathcal{K} > 0$ implies that the terms $\Theta\mathcal{K} + \Theta^2\mathcal{K}^2 + \dots$ in equation (25) give the amplification over and above n_0, due to the time-varying path of the barrier $\bar{x}$.

Figure 2: Perturbation of Stationary Equlbria



(a) High Adoption Stat. Eqbm. (b) Low Adoption Stat. Eqbm.

Figure 2 illustrates the stability of the high and low adoption equilibria, respectively, in Panels (a) and (b). Each panel considers two shocks that displace a small mass of agents away from the invariant distribution of non-adopters and endows them with the app. The shocks differ in the direction in which the mass is displaced. The blue line depicts the case where the app is given to agents with low benefit, namely with $x \approx 0$, while the red line considers a perturbation where the app is given to agents with a high benefit, namely with $x \approx \bar{x}_{ss}$. Two remarks are due. First, the high adoption equilibrium is locally stable, as displayed in Panel (a): for all shocks considered, the system returns to its invariant distribution. We also note that the half life of the shock is much shorter when the perturbation assigns the app to agents with a high benefit ($x \approx \bar{x}_{ss}$), as these agents were going to get the app soon anyways. Second, Panel (b) reveals that the low adoption equilibrium is unstable: the dynamics of the system following a perturbation are explosive, i.e., the sequence in equation (25) does not converge so that the system does not return to the invariant distribution after the shock. To appreciate the explosive nature of the path nearby the low activity stationary equilibrium, notice the difference in the scales of the two panels.

5 The Planning Problem

This section sets up the planning problem, characterizes of its solution, and shows how it can be decentralized as an equilibrium with a subsidy.[11] The planner solves a non-trivial dynamic problem since the state of the economy is an entire distribution.

At time zero the planner solves:

$$\max_{\{\bar{x}(t)\}} \Bigg\{ \int_0^\infty e^{-rt} \int_0^U \underbrace{(1/U - m(s,t))}_{\text{Density of adopters}} \underbrace{s\,(\theta_0 + \theta_n N(t))}_{\text{Flow benefit}} ds\, dt$$

$$- \underbrace{\int_0^\infty e^{-rt} c\,(N_t(t) + \nu N(t))\, dt}_{\text{Flow of adoption cost: gross new adoptions}} \Bigg\}$$

subject to

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

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

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

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

$$m(x,0) = m_0(x) \text{ for all } x\,. \quad \text{Initial condition}$$

The objective function of the planner integrates the lifetime utility of agents using as a weight the discount factor e^{-rt} for the cohort born at t. The first term contains the utility flow of the agents who use the technology. The second term subtracts the cost of adoption, where $c(N_t(t) + \nu N(t))$ is the gross flow cost of adoption at time t. This flow cost is driven by the inflow of new adopters $N_t(t)$ and by the replacement of dead agents (who had adopted in the past) by newborns.[12] The first constraint defines $N(t)$, the second constraint is the KFE for the density of non-adopters, m. As before, the density is zero to the right of $\bar{x}(t)$, there is an exit point at $\bar{x}(t)$, and there is a boundary condition from the reflection at zero.

At each time t the planner decides a threshold $\bar{x}(t)$ that determines adoption, taking as given the initial condition $m_0(x)$ and the law of motion of the density m (affected by the choice of $\bar{x}$). To characterize the solution, we write the Lagrangian for this problem. We denote the

[11] Appendix D.1 characterizes the stationary solution of this problem. Appendix D.5 uses a linearized version of the problem to analyze dynamics around its invariant distribution, an exercise that is akin to the one of Section 4.

[12] At every moment there is an inflow ν of newborns without the app. A fraction $1 - \frac{\bar{x}(t)}{U}$ of the newborns immediately pays the cost c and adopts, see Appendix D.2 for details.

Lagrange multiplier of the KFE equation by $e^{-rt}\lambda(x,t)$ and replace $N(t)$ and $N_t(t)$ by the corresponding definition. To derive the p.d.e's for non-adopters, we first adapt the planning problem to a discrete-time discrete-state problem using a finite-difference approximation. In this set up, we allow for a more general policy, i.e., not necessarily a threshold rule. We obtain the first order conditions for a problem in finite dimensions and take limits to find the corresponding p.d.e's, summarized in the following proposition.[13]

LEMMA 8. A planner's problem is given by $\{\bar{x}(t), \lambda(x,t), m(x,t)\}$ such that adoption occurs for $x \geq \bar{x}(t)$, and the Lagrange multiplier λ, and the density of non-adopters m solve the p.d.e. for non-adopters:

$$\rho\lambda(x,t) = x\Big(\theta_0 + \theta_n[1 - \int_0^{\bar{x}(t)} m(s,t)ds]\Big) + \theta_n\Big(\tfrac{U}{2} - \int_0^{\bar{x}(t)} m(s,t)s\,ds\Big) \tag{26}$$
$$+ \tfrac{\sigma^2}{2}\lambda_{xx}(x,t) + \lambda_t(x,t) \text{ for } x \leq \bar{x}(t) \text{ and } t \geq 0$$
$$\lambda(x,t) = c \text{ for } x \geq \bar{x}(t) \text{ and } t \geq 0$$
$$\lambda_x(\bar{x}(t),t) = 0 \text{ for } t \geq 0 \tag{27}$$
$$\lambda_x(0,t) = 0 \text{ for } t \geq 0$$

and

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

This lemma has two important consequences. First, it allows us to compute the solution of the planning problem following similar steps as the ones used to compute the equilibrium in Section 3.1. Second, it indicates how to decentralize the optimal allocation as an equilibrium. Define $Z(t) \equiv \frac{U}{2} - \int_0^{\bar{x}(t)} m(s,t)s\,ds \geq 0$ and note that this non-negative magnitude is the difference between the average x in the population, $U/2$, and the average x among those who have not adopted the technology (the integral term). Comparing the p.d.e. for the Lagrange multiplier λ in equation (26) with the p.d.e. for D that characterizes the equilibrium in equation (9), we see that these equations only differ in the flow term $\theta_n Z(t)$. Thus, if agents who adopt the technology are given a flow subsidy $\theta_n Z(t)$ every period after they have adopted (independent of the app's usage), then the planner allocation is an equilibrium. Clearly, this is equivalent to a once and for all payment to agents adopting at t equal to $\theta_n \int_t^{\infty} e^{-\rho(s-t)} Z(s)ds$. Note that $\theta_n Z(t)$ contains the inframarginal valuation of the technology for those that use it,

[13] We provide details of this derivation in Appendix D.3.

so the subsidy's work by correcting the externality associated with the individual adoption. We summarize this discussion in the following theorem.

THEOREM 3. Fix an initial condition m_0 and the solution to the planner's problem $\{\bar{x}, \lambda, m\}$. The planner's allocation coincides with an equilibrium with the same initial conditions and a time-varying flow subsidy paid to adopters given by $\theta_n Z(t)$, where

$$Z(t) \equiv \frac{U}{2} - \int_0^{\bar{x}(t)} m(s,t)\, s\, ds \quad \text{for all } t \geq 0 \tag{28}$$

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

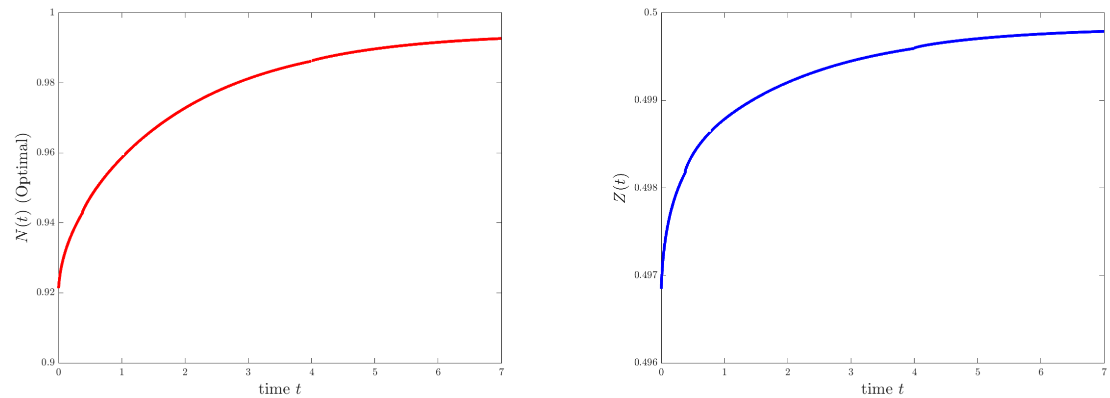
For future reference, we define $Z \equiv \mathcal{Z}(\bar{x}; m_0)$ as the solution of the path for Z defined in equation (28). In particular, given $\bar{x}$ and m_0, we solve for m using the KFE and then compute Z.

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 (27) given the paths of N and Z and terminal condition $\lambda(x,T) = \lambda_T(x)$. Let $\bar{x} = \mathcal{X}^P(N, Z; \lambda_T)$ denote the functional, defined as the $\mathcal{X}$ in Section 2.1, 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 analysis used in Section 3, based on monotonicity, to characterize the solution to this fixed point problem, and to compute it.

Figure 3 illustrates how the application of the optimal subsidy leads to a high adoption equilibrium. In Panel (a) of the figure, we plot the time path of the share of adopters, $N(t)$, for the planning problem, using the stationary equilibrium distribution of non-adopters as the initial distribution (i.e., $m_0(x) = \tilde{m}(x)$). Let denote by N_{ss} the value of the equilibrium steady state. In the planning problem, the path of $N(t)$ jumps immediately from N_{ss} (at the time the subsidy is introduced) and gradually converges to the stationary distribution for the planning problem.[14] Panel (b) shows the time path of the optimal subsidy to implement the optimal, $Z(t)$, which starts at the value $Z(0) = \frac{U}{2} - \int_0^{\bar{x}_H} \tilde{m}(s) s ds$ and increases over time. In this example, the high-adoption equilibrium has partial adoption, i.e. $N_{ss} < 1$, but the efficient allocation, as can be seen in panel (a), converges to almost full adoption of the technology.

[14]In this example, $N_{ss} = 0.42$.

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



(a) $N(t)$ (Optimal)

(b) $Z(t)$

6 Application: SINPE, A Digital Payments Platform

In May 2015, the Central Bank of Costa Rica (BCCR) launched SINPE Movil (hereafter, SINPE), a digital platform that enables users to make money transfers using their mobile phones.[15] To utilize SINPE, users must have a bank account at a financial institution and link it to their mobile number. According to the BCCR, the primary objective of SINPE 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 for relatively small transfers, which are not subject to any fee as long as they do not exceed a daily sum. The maximum daily amount transferred without a fee varies by bank; for most users, it is approximately \$310, although some banks have lower limits of \$233 and \$155.[16] The average transaction size in SINPE is about \$50, and has slowly decreased over time, as shown in Figure G2.

6.1 Data

This section describes the battery of administrative datasets used in the paper. First, we leverage data on Sinpe transactions. Our data on SINPE usage is comprehensive: For each user in the country, we have official records on the *exact date* when she adopted the technology, along with records on each transaction made. In particular, for each transaction, the data records the *amount transacted* along with the individual identifier of *the sender and the receiver* of the money. Records also include the sender's and the receiver's bank.

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

[16]Respectively, these limits in dollars correspond with approximately 200,000; 150,000; and 100,000 Costa Rican colones. These amounts correspond with 2021 limits and exchange rates.

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

We also leverage information on networks of coworkers for each formally employed individual, along with their income. Matched employer-employee data is obtained from the Registry of Economic Variables of the Central Bank of Costa Rica, which tracks the universe of formal employment and labor earnings. The data include *monthly* details on each employee, including her earnings and employment history spanning SINPE's lifetime (2015-2021).[17] The average number of coworkers in our sample is 4.7 (median 1). With this information, we can identify which people are working at the same firm in a given month to construct networks of coworkers which can be matched to SINPE records. Networks of coworkers vary at a monthly frequency as people change employers.

While our baseline analysis focuses on coworkers networks, we complement its statistics with those of other network types, namely, networks of neighbors and relatives. 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.[18] Data on nationwide family networks is available from the National Registry and makes it possible to reconstruct each person's family tree.[19] The data includes individual identifiers that can be linked to SINPE. The same data source provides details on individual demographics. Finally, we leverage data on corporate income tax returns from the Ministry of Finance for the universe of formal firms. The data contains typical balance sheet variables since Sinpe's inception, and includes details on each firm's sector and location.

6.2 From Model to Data

As described in the previous section, we obtained (i) transaction-level data including information on the senders and receivers who took part in each transaction since the app's inception, and (ii) individual-level data on networks from official sources. Further, crucially, we can link identifiers in (i) and (ii). We leverage this substantial data effort to construct measures of networks (N) for *each* individual and to obtain individual-level measures of adoption at the extensive and intensive margins. Our baseline analysis focuses on networks of coworkers—the network for which we can more credibly identify network changes that are plausibly orthogonal to changes in app usage. This will enable us to document evidence of selection (x) and

[17] 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).

[18] 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).

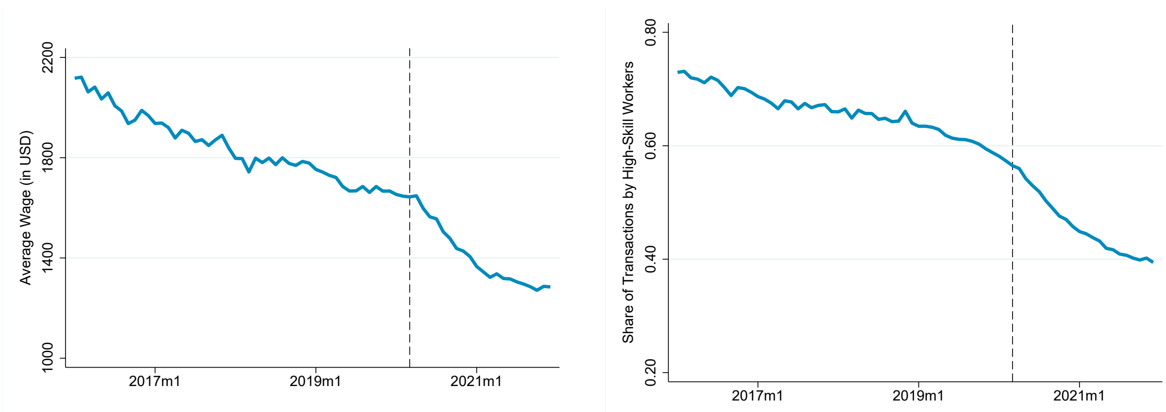
[19] We find that the average number of first-degree, second-degree, and third-degree relatives is 6.4 (median 5), 10.9 (median 9), and 22.0 (median 18), respectively.

cleanly identify θ_n, which governs the strength of the strategic complementarities and will be crucial for the policy analysis and the estimation of the optimal subsidy.

6.2.1 Evidence of Selection at Entry

Through the lens of our model, early adopters—who started using the technology even when the network was small—should be more intense users (with higher x). Consistent with this notion, we document that early adopters have distinct characteristics as compared with users who adopted later. For this exercise, and throughout the entire paper, we classify an individual as an adopter starting from the time when she first used the app. First, as shown in Figure 4, 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)).[20] Early adopters are also younger, on average, than later adopters, as shown in Figure G6.

Figure 4: Average Wage and Skill at the Time of Adoption



(a) Wage (b) High Skill

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

Second, we can more closely bring the model to the data by interpreting the flow benefit of agents who adopt the technology as being proportional to how intensively they use SINPE. Suppose SINPE users choose the intensity with which they use the app. Specifically, suppose

[20] We classify an occupation as high-skill if it requires education or training beyond a high-school diploma. The dashed vertical line in each figure denotes the beginning of the pandemic, which just as in Figure G1 did not have a major impact on overall trends.

ξ_t is the probability of a transaction per unit of time, maximizing the following expression:

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

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

$$[\theta_0 + \theta_n N_t]x_t = \max_{\xi_t} \frac{1+p}{p}\left[\beta(x_t, N_t)\xi_t - \frac{\xi_t^{1+p}}{1+p}\right] \text{ for all } x_t \in [0, U] \text{ and } N_t \in [0, 1]. \quad (29)$$

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

$$\ln \xi_t^* = \frac{1}{1+p} \ln [(\theta_0 + \theta_n N_t)] + \frac{1}{1+p} \ln x_t. \quad (30)$$

Given the discreteness of the number of transactions in the data, ξ_t^* is interpreted as the mean of a Poisson distribution; transactions each period are drawn from a Poisson probability distribution with mean ξ_t^* (i.e. $\mathtt{T}_t \sim \text{Poisson}(\xi_t)$). Thus, if we were to remove the network×time variation from the logarithm of the number of transactions, then they would proxy for $\ln x_t$, as through the lens of the model only the idiosyncratic variation would remain. The model also predicts that users with a higher x would adopt the technology earlier. Thus, we can obtain a relation between intensity of usage ($\mathtt{T}_{it}^n$) and the share of user i's network who had adopted the technology at the time *when she first used the app* ($N_{i,entry}^n$):

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

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

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

Table 1: Number of Transactions and Size of Network at Entry

	Dependent variable: Number of Transactions (IHS)		
	(1)	(2)	(3)
Size of Coworkers' Network at Entry	-1.300***		
	(0.043)		
Size of Neighbors' Network at Entry		-2.730***	
		(0.025)	
Size of Family Network at Entry			-1.181***
			(0.006)
Observations	16,138,736	34,409,818	14,700,288
Network×Time/Cohort FE	Yes	Yes	Yes
Adjusted R-squared	0.304	0.234	0.199

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

6.2.2 Estimating the Strength of the Strategic Complementarities

The core idea behind strategic complementarities is that usage benefits increase with the size of a user's network. Recall the expression in equation (30). Under this interpretation of the model, the intensity with which the application is used, which is observable in the data (e.g., number or value of transactions), is proportional in logs to the flow benefit of adopting the application as described in the model. After taking the first order Taylor expansion of $\ln(\theta_0 + \theta_n N_t)$ around N^* and plugging it into equation (30), we obtain:

$$\ln \mathtt{T}_t \approx \ln(\theta_0 + \theta_n N^*) + \frac{1}{1+p}\frac{\theta_n(N_t - N^*)}{\theta_0 + \theta_n N^*} + \frac{1}{1+p}\ln x_t. \tag{31}$$

Moreover, taking first differences, it follows that:

$$\Delta \ln \mathtt{T}_t = \beta \Delta N_t + \nu_t, \tag{32}$$

where $\beta \equiv \frac{1}{1+p}\frac{\vartheta}{1+\vartheta N^*}$, $\vartheta \equiv \frac{\theta_n}{\theta_0}$, and $\nu_t \equiv \frac{1}{1+p}\Delta \ln x_t$. Further, if $p \approx 0$, then $\vartheta = \frac{\beta}{1-N^*\beta}$. Thus, throughout all the tables in this section, we can evaluate N^* at its mean value to recover ϑ from each β; these are our coefficients of interest since strategic complementarities in the adoption of the technology exist if $\beta > 0 \iff \vartheta > 0 \iff \theta_0 > 0$ and $\theta_n > 0$. Note that equation (32) is in *differences*, therefore, *any individual or network characteristics which are time invariant would cancel out.*

With these expressions, one can first naively run an OLS specification. We do so in Appendix **??** and find a significant correlation between the intensity of app usage and the share of individuals in the user's network who have adopted it. This correlation remains robust across various network definitions, usage intensity measures, and specifications. Then, we show that the impact of network size on usage intensity persists even after employing a leave-one-out instrument to address potential endogeneity concerns and measurement errors. Additionally, this relationship is unaffected when accounting for selection through a balanced panel of adopters. However, to quantify the model, one ultimately needs to take a stand on the causal impact of changes in the number of adopters; we do so by focusing on mass layoffs.

Usage After a Mass Layoff (Intensive Margin of Adoption). This strategy focuses on the network of coworkers and implements both (i) a mover design, where we follow workers displaced during mass layoffs to examine the effect of network changes on the intensive and extensive margins of adoption and (ii) an analysis of stayers, in which we instead focus on workers who remained at a firm after a mass layoff.[21] The main hypothesis of the movers exercise is that workers, who were displaced during a mass layoff and who ended up at firms where a larger share of colleagues had SINPE (larger N), have more incentives to use the app via the effect of strategic complementarities. Similarly, the idea behind the analysis of stayers is that workers who remain at a firm that, for instance, laid off most of its SINPE-using employees (smaller N), have now less incentives to use the app.

We first analyze the impact of a mass layoff on movers' usage. To do so, we focus on workers who were fired during a mass layoff and 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 before, it is possible to derive the relationship in equation (**??**) from our theoretical model, which speaks to the technology's usage intensity. Thus, we consider:

$$\begin{aligned}\Delta \ln \mathtt{T}_i =& \alpha + \zeta \Delta N_i^{coworkers} + \gamma \Delta \ln wage_i + \psi \Delta \ln size_i + \varphi\, date\, hired_i + \\ & \omega \Delta Covid_i + \delta \lambda_{ic} + \nu \ln \sum_{t=0}^{move} \mathtt{T}_{ti} + \nu \sum_{t=0}^{move} (\ln \mathtt{T}_{t,\text{ new firm}} - \ln \mathtt{T}_{t,\text{ old firm}}) + \epsilon_i, \end{aligned} \tag{33}$$

[21] To define a mass layoff, we follow Davis and Von Wachter (2011) and identify establishments with at least 50 workers that contracted their monthly employment by at least 30% *and* had a stable workforce before this episode and did not recover in the following 12 months. Given we also analyze stayers, we implement a few additional refinements. Details are provided in Appendix G.2.1.

where $\Delta \ln \mathtt{T}_i$ refers to the change in monthly intensity with which individual i used SINPE within 6 months *after* arriving at her new firm compared with 6 months *before* being fired; $\Delta N_i^{coworkers}$ 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 at each firm; $date\, hired_i$ controls for the date 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; λ_{ic} controls for cohort (i.e., the date when individual i adopted SINPE); $\ln \sum_{t=0}^{move} \mathtt{T}_i$ is the sum of all historical transactions made by agent i since she adopted the app, and $\sum_{t=0}^{move} (\ln \mathtt{T}_{t,\text{ new firm}} - \ln \mathtt{T}_{t,\text{ old firm}})$ is the difference in the (log) historical transactions made by workers at the new firm and the old firm up until the move occurred, which aims to control for factors—other than strategic complementarities—which might facilitate adoption at the new vs. the old firm.

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

Panel (a) of Table 2 displays our results using the number of transactions per user as our dependent variable. Changes in the intensity of usage depend positively and significantly on the change in the share of adopters at the old and new firm. Panel (a1) of Figure 5 displays the marginal effect of these network changes following the specification described by Column (2) of Table 2. As this panel shows, not only is the relationship between usage and network

Table 2: Intensity of Usage and Changes in Coworkers' Network After a Mass Layoff

Dependent Variable: Δ *Number of transactions (IHS)*

	(a) Movers			(b) Stayers		
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta N_i^{coworkers}$	2.646***	1.406***	1.283***	3.284***	0.952**	0.971**
	(0.203)	(0.268)	(0.294)	(0.237)	(0.443)	(0.435)
$\Delta \ln wage_i$		0.383***	0.385***		0.203**	0.132
		(0.070)	(0.077)		(0.087)	(0.103)
$\Delta Covid_i$		0.168**	0.167***		-0.010	-0.012
		(0.027)	(0.032)		(0.025)	(0.024)
Observations	917	917	917	2,236	2,236	2,236
Time FE	No	Yes	Yes	No	Yes	Yes
Cohort FE/Historical T	No	No	Yes	No	No	Yes
Adjusted R-squared	0.153	0.244	0.262	0.093	0.122	0.184

Notes: The unit of observation is the individual. We run regressions using data on mass layoffs which occurred between May 2015, when the technology was introduced, until December 2021. While time and cohort fixed-effects' inclusion varies across columns, all other independent variables in equation (33) are present across columns. Standard errors are in parentheses.

changes positive, but also whenever a worker moves to a firm with a lower adoption rate, her usage decreases (i.e., the change on the vertical axis is negative), a relationship that would be hard to reconcile with a pure learning story.[22]

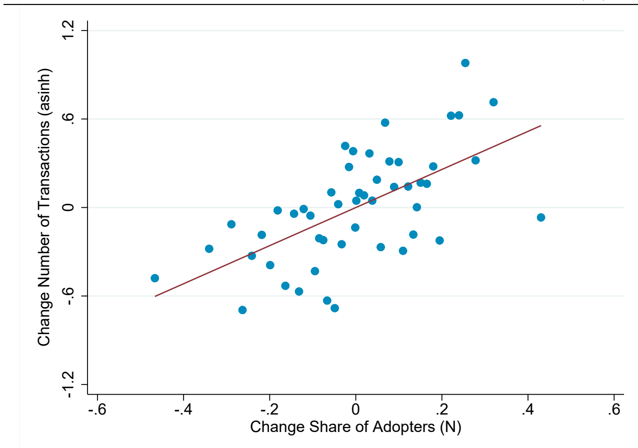
Column (3) controls for cohort, i.e., date of adoption, which aims to mitigate any effect of more experienced users behaving differently than beginners. Column (3) also controls for the total historical transactions made, which in a similar spirit as cohort, intends to mitigate any effect resulting from learning how to use the app from others. Interestingly, as compared with Column (3), adding these controls does not change the coefficient of interest almost at all. This result aligns with the following intuition: at the intensive margin—once users have already adopted and used the app—a learning story is less plausible, as reflected by ζ not changing after controlling for cohort and historical usage.

The analysis can be taken to an even more detailed level if, instead of considering all transactions in the left-hand-side variable, we focus only on those which had a coworker as a counterpart. This subsample allows us to better identify changes in usage intensity which are a direct consequence of the arguably exogenous changes in the network of coworkers. Reassuringly, results are remarkably similar to those using all transactions, as shown in Panel (a2) of Figure 5.

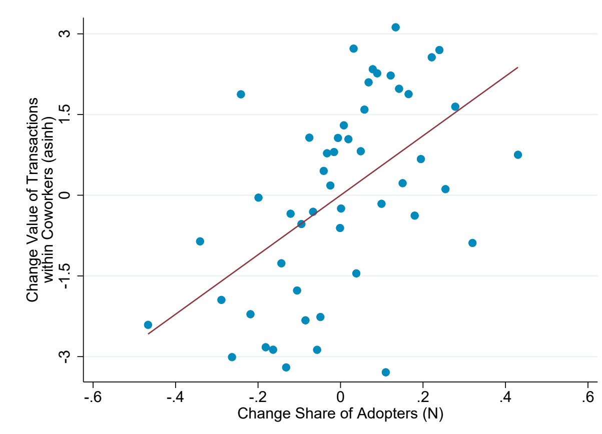
[22] The marginal effect considering the value of transactions as dependent variable, as opposed to the number of transactions, is reported in Figure ??.

Figure 5: Marginal Effect of Network Changes on Usage Intensity

(a) Movers

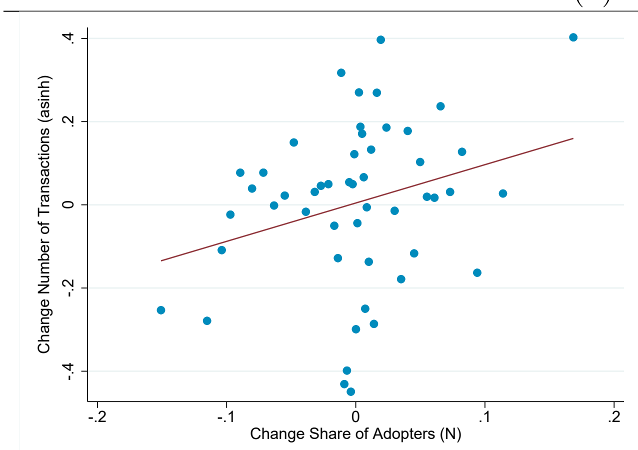


(a1) All transactions

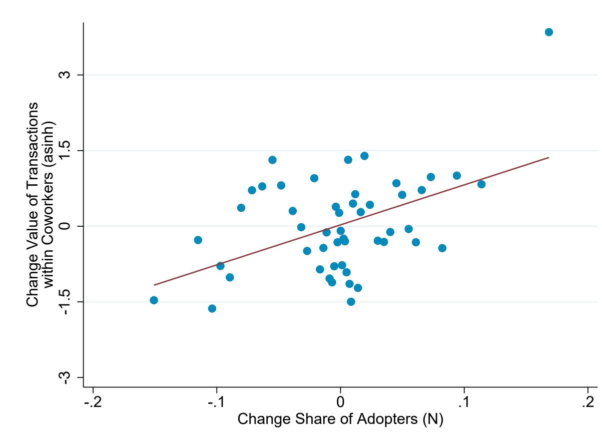


(a2) Transactions with coworkers only

(b) Stayers



(b1) All transactions



(b2) Transactions with coworkers only

Notes: Panel (a1) plots the marginal effect of $\Delta N_i^{coworkers}$ in the specification described by Column (3) of Table 2, while Panel (b1) plots the marginal effect of $\Delta N_i^{coworkers}$ in the specification described by Column (6) of Table 2. 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. Panels (a2) and (b2) are similar, but differ as the dependent variable in these estimations is the number of transactions *which have a coworker as a counterpart* (transformed using the inverse hyperbolic sine function) on each period for each user.

A similar analysis can be conducted based on *stayers*. Namely, we focus on workers who remain at a firm after it experienced a mass layoff. Their change in N will therefore depend on how the composition of SINPE adopters changed after the mass layoff. We then consider a regression similar to equation (33), except for the last regressor which would be zero in this case.[23] Results based on stayers are reported in Panel (b) of Table 2 and Panel of Figure 5. Remarkably, although the movers design is based on a very different sample than the analysis based on stayers, the estimated coefficients in our preferred specifications, in columns (3) and (6) of Table 2, are statistically equal.

Adoption After a Mass Layoff. Lastly, we analyze changes in the extensive margin of adoption. For movers, we consider the change in the probability of adoption for displaced workers *who had not adopted the app by the time they were rehired*, and how it depends on the change in the share of coworkers who had SINPE at their old and new firm. We consider:

$$\begin{aligned} Adopt_i = &\alpha + \zeta \Delta N_i^{coworkers} + \gamma \Delta \ln wage_i + \psi \Delta \ln size_i + \varphi \, date\, hired_i + \\ &\omega \Delta Covid_i + \nu \sum_{t=0}^{move} (\ln \mathtt{T}_{t,\text{ new firm}} - \ln \mathtt{T}_{t,\text{ old firm}}) + \epsilon_i, \end{aligned} \tag{34}$$

where $Adopt_i$ equals one if individual i adopted SINPE within 6 months after arriving at her new firm, and zero otherwise. Other variables are defined in the same way as in equation (33). For stayers, we instead consider the probability of adoption for workers who were not fired by a firm which underwent a mass layoff and how it depends on the change in the composition of workers who had SINPE, before and after the mass layoff took place. We then use a regression similar to equation (34), except for the last regressor which would be zero.

Panels (a1) and (b1) of Figure 6 estimate equation (34) using a logit model. The marginal effects of changes in network adoption are reported in brackets. The analysis of movers in panel (a1) consistently finds 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. Reassuringly, panel (b1) also finds that workers who experienced an increase in the share of adopters among their peers were more likely to adopt SINPE themselves. The marginal effect of $\Delta N_i^{coworkers}$, under the specification described by Column (3) in each table, is shown in panels (a2) and (b2). These marginal effects are monotonous and, as expected, are present only when the change in the share of adopters is positive, regardless of the subsample considered.

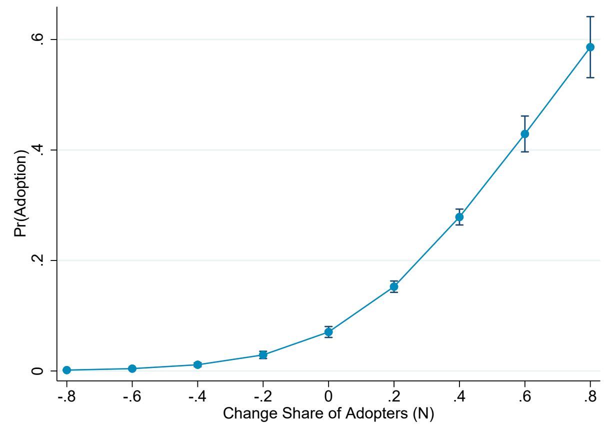
[23] An additional control equal to the change in the average wage at the workers' firm delivers statistically equal results, both for the intensive and extensive margin analyses.

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

(a) Movers

	(1)	(2)	(3)
$\Delta N_i^{coworkers}$	8.321***	5.197***	5.018***
	(0.161)	(0.207)	(0.208)
	[0.450]	[0.376]	[0.350]
$\Delta \ln wage_i$		0.076	0.040
		(0.055)	(0.056)
$\Delta Covid_i$			0.097***
			(0.026)
Observations	10,176	8,035	8,035
Time/Cohort FE	No	Yes	Yes
Pseudo R2	0.507	0.529	0.530

(a1) Changes in Adoption Probability

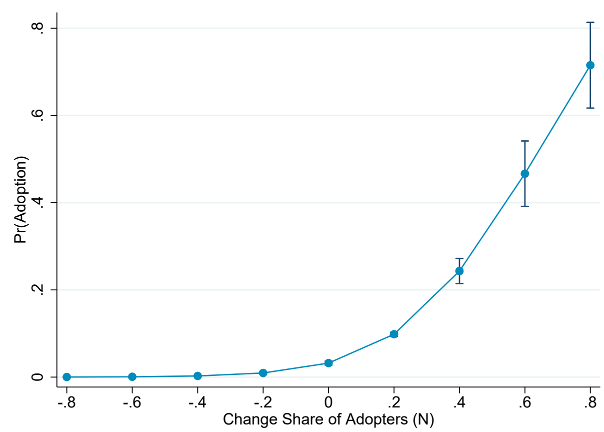


(a2) Marginal Effect of Network Changes

(b) Stayers

	(1)	(2)	(3)
$\Delta N_i^{coworkers}$	11.249***	6.489***	6.418***
	(0.172)	(0.356)	(0.357)
	[0.716]	[0.442]	[0.437]
$\Delta \ln wage_i$		-0.104	-0.098
		(0.089)	(0.089)
$\Delta Covid_i$		(0.089)	(0.089
			(0.016)
Observations	8.32124,329	23,005	23,005
Time/Cohort FE	No	Yes	Yes
Pseudo R2	0.330	0.356	0.357

(b1) Changes in Adoption Probability



(b2) Marginal Effect of Network Changes

Notes: Panels (a1) and (b1): 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. Panels (a2) and (b2): The figures plot the marginal effect of $\Delta N_i^{coworkers}$ in the specification described by column (3) of panels (a1) and (b1), respectively, in this figure. Vertical bars denote 95% confidence intervals.

7 Quantitative Performance and Optimal Subsidy

In this section, we calibrate our model and evaluate its performance relative to SINPE data. We begin by describing an extension of the model that combines the model of strategic complementarities with a learning model. The model with only strategic complementarities assumes that all individuals are informed about the technology at all times. However, according to the 2017 Survey of Payment Methods conducted by the Central Bank of Costa Rica, only about 5% of adults reported knowing about SINPE Móvil two years after its launch. The hybrid model helps align the model with this fact and, as a result, with the smooth and relatively flat path for $N(t)$ during the first few years after the launch displayed in Figure 7. Next, we describe our calibration procedure in detail.

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

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

where $I(t)$ denotes the fraction of the population informed about the technology and the parameter β_0, which gives the number of meetings per unit of time between those informed, $I(t)$, and those uninformed, $1-I(t)$. The term $\frac{\beta_0}{U} I(t)(1-I(t))$ is the flow of agents per unit of time that learn about the app. See Appendix E for details.[24]

Calibration. We interpret the flow benefit of agents who adopt the technology as being proportional to how many transactions they conduct, and assuming a convex adjustment cost (i.e., $p > 0$). U can be normalized without loss of generality (we use the normalization $U = 1$), so the problem features seven independent parameters: $\nu, r, \theta_n, \theta_0, \sigma$, p, and c. The model with learning has an additional parameter, β_0, and an initial condition for the informed population, $I(0)$.

The parameters ν, r, β_0 and are calibrated externally. We set ν to 0.0278 to match the

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

rate at which agents stop using SINPE; namely, the average fraction of agents in 2019-2021 who had adopted SINPE but did not conduct a single transaction in the app within a year. We use the last three years of the data, when the adoption rate is higher, to focus on periods closer to a stationary equilibrium. We set the discount factor r to be consistent with a 5 percent annual interest rate. This value is a lower bound for r, which can admit higher values if we assume agents expect new technologies to arrive in the future and replace SINPE. The values of ν and r imply $\rho = r + \nu = 0.0778$. Lastly, we set $\beta_0 = 1.33$ to match the share of people informed about the app (approximately 5% two years after the launch) for an initial condition of $I(0) = 0.001$ (i.e. 0.1 percent of the workers are informed about SINPE at the time it was launched).

The parameters $\theta_n, \theta_0, \sigma, p$ are calibrated using simulated methods of moments (SMM). Intuitively, we aim to choose parameters that make the model consistent with the distribution of transactions in the data and the mass layoff exercise. To achieve this, in the data, we focus on workers at firms active from 2019 to 2021 with more than 5 employees. We take advantage of having closed form solutions for the steady state. Thus, we concentrate on firms close to a stationary equilibrium, specifically those whose N (fraction of employees with the app) changed by less than 0.1 percentage points in 2021. We then compute moments from the empirical distribution of transactions over the years 2020-2021 and simulate the model, replicating the same characteristics as our empirical sample. In addition, we simulate a sample of firms that replicates the characteristics of those subject to a mass layoff. We do this to run the same estimation, presented in Section 6.2.2, in the simulated data to obtain information on the parameters governing the strength of the strategic complementarities. We then choose the parameters that minimize the distance between the moments in the data and the model. We provide more details of our strategy below.

Simulation. We begin by simulating the model for a monthly panel of agents. Our simulation takes as given the values of ν, r, and β_0, since they are calibrated externally, and $N_{ss} = 0.90$, which is obtained from our sample of firms close to a stationary equilibrium. Initial conditions $x(0)$ are drawn from the stationary distribution of adopters. To find this distribution, we first find $\bar{x}_{ss}$ using the following equation:

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

Then, given $\bar{x}_{ss}$, we find the distribution of adopters using the stationary distribution of non-adopters:

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

using that $N_{ss} = I_{ss} - M_{ss}$ and $I_{ss} = (1 - \frac{\nu}{\beta_0})$. We simulate a panel of 5,000 individuals.[25] In the simulation, agents die at rate ν and they become inactive in the application just as in the data. The process of x follows a Brownian motion, independent across agents, with variance per unit of time σ, no drift, and reflecting barriers at $x = 0$ and $x = U$. Since x is unobserved and what is observed are transactions, as before, we interpret the flow benefit of agents who adopt the technology as being proportional to how intensively they use SINPE. Thus, we compute: $\xi_t = [\theta_0(1+\vartheta N_{ss})x_t]^{\frac{1}{1+p}}$, where $\vartheta \equiv \frac{\theta_n}{\theta_0}$, to find the number of transactions $\mathtt{T}_t$ by drawing them from a Poisson probability distribution $\mathtt{T}_t \sim \text{Poisson}(\xi_t)$.

Mass Layoff. We also simulate a panel of workers at firms with the same characteristics as those experiencing mass layoffs in the data. Specifically, as presented in Table G3, we simulate a sample of 292 firms with 94 employees each. We focus on workers who remain at a firm after it has experienced a mass layoff (i.e., stayers).[26] 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 after a mass layoff. The change in N depends on how the composition of adopters changes after the mass layoff, which involves randomly choosing and removing a fraction of workers from each firm undergoing a mass layoff. We choose the magnitude of these mass layoffs to match the average size of these events in the data (i.e., 57%). We then run the same regression that is implemented in Table 2. First, we calculate the number of transactions before and after the mass layoff event. Then, we regress the change in monthly transactions within six months of the mass layoff event on the change in the share of coworkers who had adopted the app before and after the event. The estimated coefficient is a moment that we target in our calibration

Cost of Adopting. The adoption cost, c, can be obtained from the solution of the stationary problem for adoption, given a value of $\bar{x}_{ss}$ and the parameters $\theta_n, \theta_0, \sigma$. In particular, we use the following equation:[27]

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

[26] Table 2 shows that the estimated impact of a mass layoff on usage is statistically equal for movers and stayers.

[27] All details on the derivation of this equation can be found in Appendix B.1.

$$\frac{\bar{x}_{ss}+\bar{A}_1 e^{\eta\bar{x}_{ss}}+\bar{A}_2 e^{-\eta\bar{x}_{ss}}-c/\bar{\theta}_{ss}}{1+\eta\left(\bar{A}_1 e^{\eta\bar{x}_{ss}}-\bar{A}_2 e^{-\eta\bar{x}_{ss}}\right)}=\frac{1}{\eta}\frac{e^{\eta\bar{x}_{ss}}+e^{-\eta\bar{x}_{ss}}}{\left(e^{\eta\bar{x}_{ss}}-e^{-\eta\bar{x}_{ss}}\right)} \quad \text{where } \eta=\sqrt{2\rho}/\sigma$$

where $\bar{A}_1=\frac{1}{\eta}\frac{\left(1-e^{-\eta U}\right)}{\left(e^{-\eta U}-e^{\eta U}\right)}$, $\bar{A}_2=\frac{1}{\eta}\frac{\left(1-e^{\eta U}\right)}{\left(e^{-\eta U}-e^{\eta U}\right)}$ and $\bar{\theta}_{ss}\equiv\frac{\theta_0+\theta_n N_{ss}}{\rho}$.

Calibrated Moments. We target the following five moments: the mean number of transactions, the median number of transactions, the absolute value of changes in transactions, the coefficient of the mass layoffs regression, and the autocorrelation of the number of transactions. As done throughout the paper, all the targeted data moments are calculated after controlling for COVID-19 cases. Parameters $\theta_n, \theta_0, \sigma$, and p are chosen to minimize the sum of the norms of the percent deviations of simulated moments from target moments.[28] Table 3 reports the empirical and simulated moments.[29]

Table 3: Moments: Distribution of Transactions

Parameter	Value	Std. Dev.	Moment	Data	Model
σ	0.034	0.005	Mean Number of Transactions	6.88	6.86
θ_0	26.31	3.195	Median Number of Transactions	6.08	6.66
p	0.0056	0.0003	Absolute Value Changes in Transactions	3.48	2.77
$\vartheta\equiv\frac{\theta_n}{\theta_0}$	5.722	1.141	Coefficient Mass Layoffs Regression	0.97	0.97
			Autocorrelation of Transactions	0.97	0.95

Intuitively, the mean and median number of transactions provide information about θ_0 and p, as shown by equation (30). The dispersion in the changes of transactions and the autocorrelation of transaction provide relevant information to pin down σ; a lower variance decreases the absolute value of the changes in transactions but increases the autocorrelation coefficient. Lastly, equation (32) shows that the coefficient of the mass layoffs regression informs the estimation of θ_n.[30] Targeting this moment allows us to leverage the rich variation across networks to inform the model estimation. Importantly, by running the same

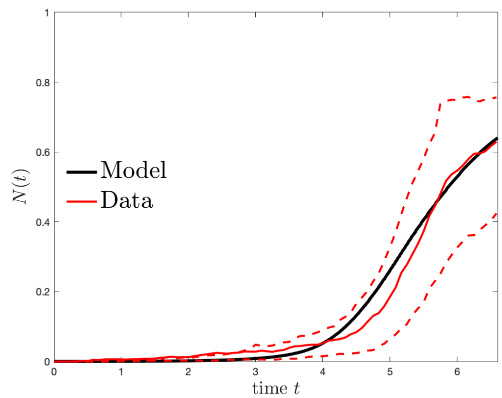
[28] This is, $\min\sum_i^5\omega(i)\frac{|\text{Model}(i)-\text{Data}(i)|}{\text{Data}(i)}$, where Model($i$) is a simulated i-th moment and Data(i) is a target value of i-th moment. We assign half of the weight to each of the mean and median of transactions since they provide similar information for the calibration. We assign twice as much weight to the coefficient of the mass layoff regression, as this moment provides information about the strength of the strategic complementarities.

[29] We simulate the model 200 times and use the average values of the moments from the simulated data. In the model and the data, we calculate the autocorrelation of the average transactions over two years to minimize the impact of measurement error in the autocorrelation coefficient. The standard deviation of the parameters is obtained from the the SMM variance-covariance matrix, which is obtained from calculating the derivative of the criterion function with respect to each parameter.

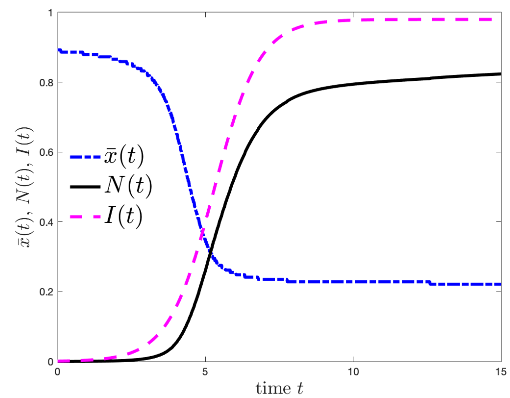
[30] The learning model in Appendix E cannot capture the patterns observed after mass layoffs, as it features random diffusion of the technology. Consequently, after adoption, an agent's flow benefit does not depend on the network size N (i.e., $\theta_n=0$).

regression in both the data and the model, we do not rely on approximating the relationship between the change in transactions and the change in the share of users who have adopted the technology around the stationary equilibrium to obtain information about ϑ, as done in equation (32). Overall, Table 3 shows that the model is quantitatively consistent with the empirical distribution of transactions.[31]

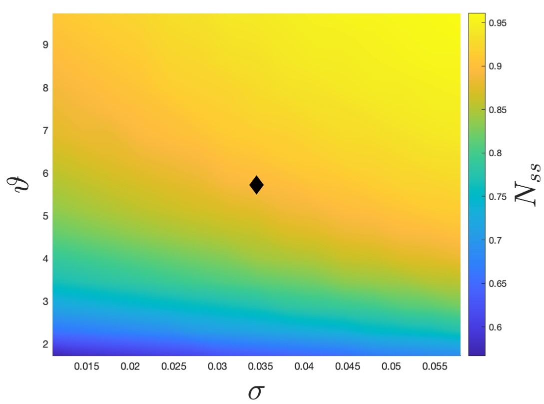
Figure 7: Path of Adopters (Short-Run and Long-Run)



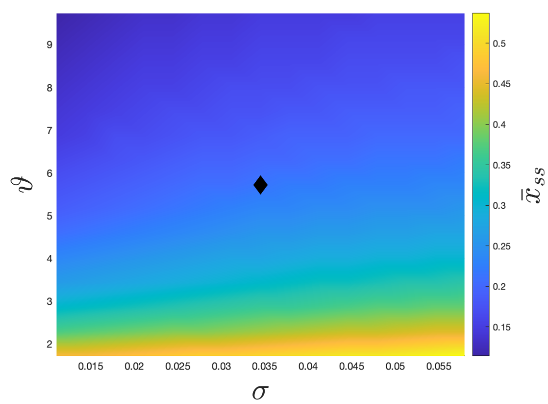
(a) Model vs Data



(b) Long-Run Path



(c) Comparative Statics: N_{ss}



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

Notes: Panel (a) compares the path of adopters in the model and in the data. The solid red line shows the patterns of diffusion of the technology in the median firm, 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 10^{th} and 90^{th} percentiles. Panel (b) shows the share of informed agents, $I(t)$, the share of adopters, $N(t)$, and the levels of $\bar{x}(t)$ predicted by the model under our baseline calibration. Panel (c) and (d) show how N_{ss} and $\bar{x}_{ss}$ change with ϑ and σ, keeping the rest of the parameters constant. The black diamonds indicate the levels of ϑ and σ in our baseline calibration.

Results. Using the estimated parameters, we simulate the dynamic model to obtain the adoption path predicted by the model. Panel (a) of Figure 7 compares the path of adoption in

[31] A sensitivity analysis of the relevant parameters can be found in Appendix H.

the model and in the data. The solid red line indicates the diffusion of the technology in the median firm and the dashed lines represent the 10^{th} and 90^{th} percentiles after controlling for COVID-19 cases.[32] The figure shows that both the speed and the level of adoption generated by the model are consistent with those in the data. Panel (b) shows the path of $I(t)$, $N(t)$ and $\bar{x}(t)$. The path of $I(t)$ shows that most people are informed about the technology within the first 7 years; in the stationary distribution, approximately 98% of the population knows about the application and 90% of the workers the median firm adopt the application as shown by the path of $N(t)$. Importantly, the declining path of $\bar{x}(t)$ indicates that, consistent with our empirical evidence, the model features selection: agents that benefit the most from the technology adopt first. This contrasts with the model that features only learning, which shows no selection in the adoption of the technology.[33]

Panels (c) and (d) of Figure 7 display the values of N_{ss} and $\bar{x}_{ss}$ in the stationary equilibrium as we vary ϑ and σ, while holding others constant. These panels illustrate the comparative statics of the stationary equilibrium derived in Section 3.3. Panel (c) shows how the stationary level of adoption changes with ϑ and σ (a black diamond denotes N_{ss}'s level in the baseline calibration). As ϑ increases, so does the strength of the strategic complementarities, and not surprisingly, N_{ss} increases as ϑ rises. The effect of σ is more subtle and results from two opposing forces. On the one hand, higher σ decreases N_{ss} since agents have a higher option value of waiting to adopt. On the other hand, higher σ increases N_{ss}, since it implies a smaller density of non-adopters below $\bar{x}_{ss}$. In our calibration the latter effect dominates and N_{ss} increases with σ. Panel (d) displays a similar exercise for $\bar{x}_{ss}$. It shows that strategic complementarities ϑ play an important role in decreasing the adoption threshold. Moreover, given ϑ, a higher σ increases $\bar{x}_{ss}$

Variation Across Networks. The model is consistent with both high and low adoption networks of firms, each implying a different path of adopters in equilibrium. Specifically, we calibrate the model by targeting moments from individuals at firms whose level of adoption is either above the median (high adoption) or below the median (low adoption).[34] We target the same data moments computed for different samples of workers, specifically those working at firms whose average level of adoption is either above the median, $N_{ss}^{high} = 0.96$, or below

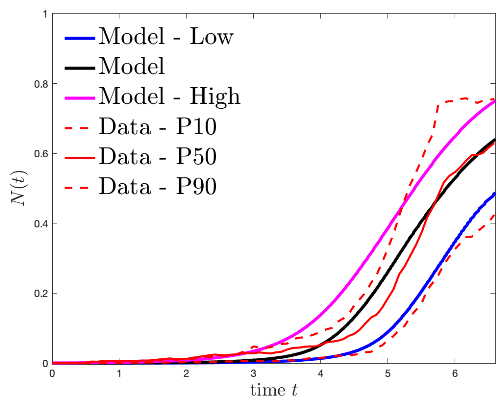
[32] We adjust the adoption path in the data to control for the pandemic. To do so, we estimate the impact of COVID-19 cases on the number of new users and subtract the predicted number of pandemic-driven new users from the cumulative number of users.

[33] Appendix H.3 presents a version of the model without strategic complementarities and only learning (i.e., $\vartheta = 0$). In this case, the path of $\bar{x}(t)$ is completely flat. Figure H7 shows the paths of $N(t)$ and $\bar{x}(t)$ for different speeds of information diffusion; namely, different values of β_0. It shows that selection occurs in the model even when the speed of information diffusion is very high.

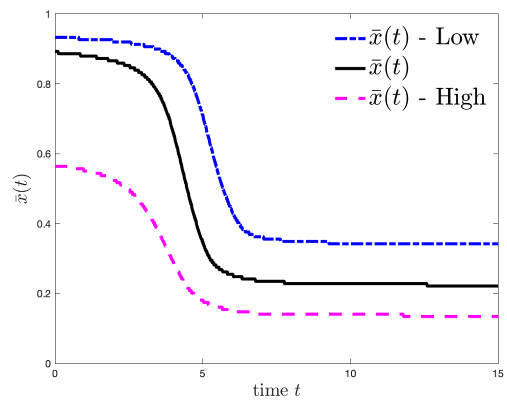
[34] The details of the calibration can be found in Appendix H.2.

the median, $N_{ss}^{low} = 0.73$, and we assume the same coefficient for the mass layoffs regressions in both calibrations. We estimate a higher level of strategic complementarities (i.e., higher ϑ) in networks with high adoption and a higher convexity in the cost of conducting transactions in low adoption networks (i.e., higher p). Panel (a) of Figure 8 shows the path of adopters in the two calibrated networks (high and low adoption) relative to the data, indicating that these calibrated versions of the model are consistent with the 10^{th} and 90^{th} percentiles of adoption in the data. Panel (b) show the path of $\bar{x}(t)$, which indicates the strength of the strategic complementarities in each of the calibrated networks. In the high adoption network, 96% of the population adopts the application. In the low adoption network, only 73% of the population adopts in the stationary equilibrium.

Figure 8: Variation Across Networks: Path of Adopters



(a) Model vs Data

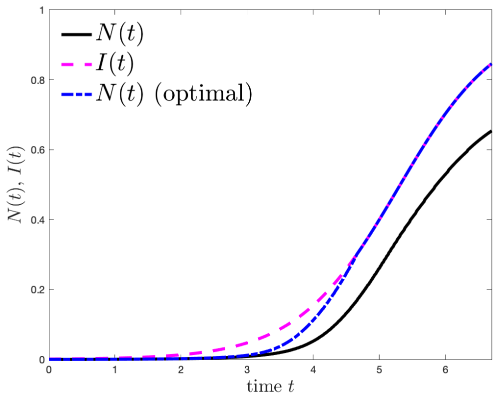


(b) Path of $\bar{x}(t)$

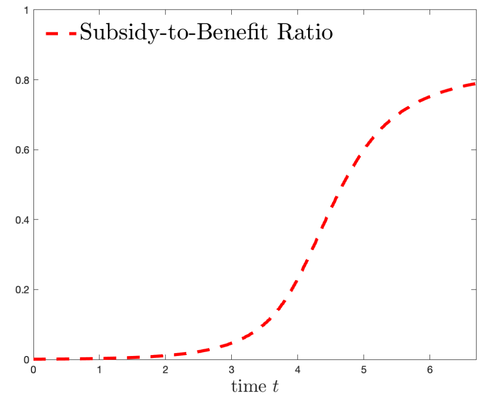
Notes: Panel (a) compares the path of adopters in the model and in the data. The solid red line shows the diffusion patterns of the technology in the median firm, and the solid black line shows the diffusion patterns in the benchmark calibration of the model. The dashed red lines indicate the 10^{th} and 90^{th} percentiles of adoption in the data. The solid magenta line shows the path of adopters in the model calibrated for high adoption, and the solid blue line shows the path of adopters in the low adoption calibration. Panel (d) shows the levels of $\bar{x}(t)$ under each of the calibrations, respectively.

Optimal Subsidy. Panel (a) of Figure 9 shows the optimal adoption path in the model with complementarities (blue line) relative to the high-adoption equilibrium (black line). During the first three years after the launch of the technology, the optimal level of adoption is similar to that of the equilibrium without subsidy. Afterward, the optimal path of adopters from the planning problem is higher. In fact, by the beginning of 2020, it is equal to the total number of informed agents in the economy—over 13 percentage points higher than the levels of adoption observed in the data—and by the end of 2021, it is over 15 percentage points

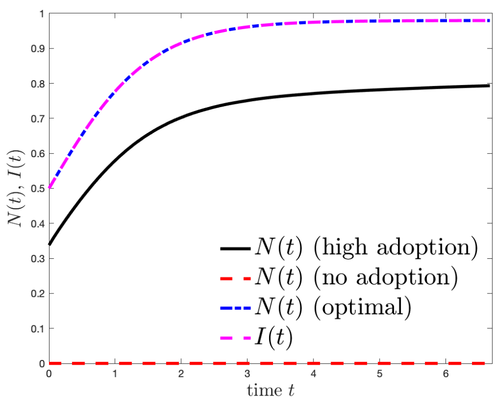
Figure 9: Planning Problem: Solution and Optimal Subsidy



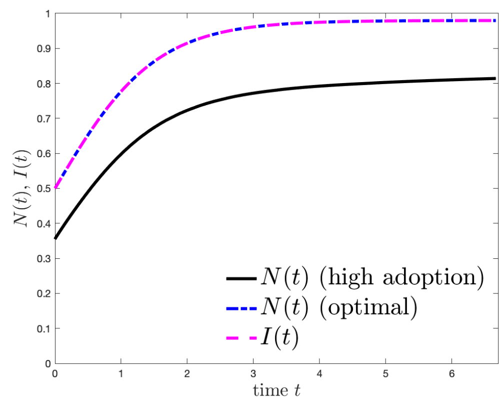
(a) Optimal Path of Adopters



(b) Optimal Subsidy



(c) Multiplicity: High Cost



(d) No Multiplicity: Low Cost

Notes: Panel (a) shows the share of informed agents, $I(t)$, the share of adopters, $N(t)$, and the optimal levels of adoption, $N(t)$ (optimal), according to the solution of the planning problem. Panel (b) shows the path of the ratio between the optimal subsidy $\theta_n Z(t)$ and the flow benefit of the average adopter, $Z(t)(\theta_0 + \theta_n N(t))$. Panel (c) shows the share of informed agents, $I(t)$, the share of adopters, $N(t)$, and the optimal levels of adoption, $N(t)$ (optimal), according to the solution of the planning problem for a high adoption cost and 70% of the population informed 7 months after the launch of the technology. Panel (b) shows the same variables for a lower adoption cost and 70% of the population informed 7 months after the initial launch. The initial distribution in both panels is $m_0(t) = 1/U$.

higher. Panel (b) shows the path of the optimal subsidy.[35] As the share of adopters increases, so does the externality. Thus, the optimal subsidy, which is the same across agents, increases over time. To see why, notice the optimal flow subsidy in equation (28) can be written as

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

[35] Figure 9 shows the subsidy $\theta_n Z(t)$ as a ratio of the net flow benefits (i.e., $(\theta_0 + \theta_n N(t))\mathbb{E}(x|adopted)$). In the invariant distribution, the subsidy-to-benefit ratio is approximately 0.84.

where the expectation over x is taken over the set of agents that have adopted the technology (see Theorem 3). The first term $\theta_n N$ captures the size of the adoption externality, i.e., the additional benefits for agents that adopt the technology. Thus, the subsidy increases as more agents adopt. Conversely, $\mathbb{E}(x|adopted)$ decreases as more agents adopt, since the marginal adopter has lower idiosyncratic benefits from adopting the technology. Intuitively, the planner internalizes that subsidizing agents with low x also benefits the rest of the agents, even if the subsidy to incentivize these agents to adopt is large. The first component of the optimal subsidy dominates and eventually pushes the economy to universal adoption. The optimal subsidy contrasts with that of a pure learning model, which is constrained efficient and where the optimal subsidy to adopt the technology is zero. Importantly, *the planner is also constrained by the share of people who are informed*; otherwise, while the subsidy would still be increasing and the same across agents, there would be a "jump" in the subsidy level as soon as the application is launched, as depicted in panel (b) of Figure 3.[36]

In Appendix H.2 we estimate the model using variation across different networks. Our findings indicate that the model aligns with both high and low adoption networks of firms, each implying different paths of adopters in equilibrium and different optimal adoption paths in the planning problem. Consistent with our benchmark calibration, all versions of the model show that the optimal subsidy pushes the economy toward universal adoption. Figure H6 shows that only for lower levels of ϑ does the planner prescribe lower adoption levels.

Multiplicity. Our model can be used to study economies with higher adoption costs featuring multiple equilibria. We consider an economy with higher adoption cost c and higher fraction of the population informed about the technology at launch. This example is motivated by a recent experience in El Salvador, where 70% of the population knew about a payment app introduced by the government (i.e., Chivo Wallet) 7 months after its initial launch.[37] Panel (c) shows the possible paths of adopters $N(t)$ for this economy. It shows that, when the adoption cost is larger (in this case 10% higher than in Costa Rica), the low-adoption equilibrium where nobody adopts the technology is not ruled out; for the same initial conditions, there is an equilibrium with high adoption and one with no adoption. Panel (d) shows the same paths for a lower adoption cost. Our model allows for the study and quantification of policies that eliminate the no-adoption equilibrium even if the optimal subsidy is not implemented. In this case, a large enough permanent subsidy can lower the adoption cost, solve the coordination failure, and send the economy to the high adoption

[36] Figure H8 shows the optimal adoption paths and the respective subsidy-to-benefit ratios for different speeds of information diffusion.

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

equilibrium, i.e., from Panel (c) to (d).[38]

8 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 that allows us to model this transition. The model provides a framework to generate gradual adoption through a novel mechanism—waiting for others to adopt—and allows us to derive predictions that can be tested empirically. We solve for the social planner's problem. The planner in our setup controls the entire distribution of adopters across time. The presence of strategic complementarities enriches the problem and allows us to link our results to the "big push" literature, as they imply that small subsidies can lead to large changes in adoption given the multiplicity of equilibria. In our framework, the optimal subsidy increases over time but it is flat across people, thus, easily implementable. The methodology we develop can be useful for a wide set of multidimensional dynamic problems, and can be applied to the study of any technology that features strategic complementarities, learning, or both.

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

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

References

Alvarez, F., Lippi, F., Souganidis, P., 2023a. Price setting with strategic complementarities as a mean field game. Econometrica 91, 2005–2039.

Alvarez, F.E., Argente, D., Lippi, F., Mendez, E., Van Patten, D., 2023b. Strategic Complementarities in a Dynamic Model of Technology Adoption: P2P Digital Payments. Working Paper 31280. National Bureau of Economic Research.

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

Arkolakis, C., Eckert, F., Shi, R., 2023. Combinatorial discrete choice: A quantitative model of multinational location decisions. Technical Report. National Bureau of Economic Research.

Auclert, A., Rigato, R.D., Rognlie, M., Straub, L., 2022. New Pricing Models, Same Old Phillips Curves? Technical Report. National Bureau of Economic Research.

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.

Bilal, A., 2023. Solving heterogeneous agent models with the master equation. Technical Report. National Bureau of Economic Research.

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.

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

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

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

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.

Jia, P., 2008. What happens when wal-mart comes to town: An empirical analysis of the discount retailing industry. Econometrica 76, 1263–1316.

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

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

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

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

Online Appendix for

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

December 2024

Fernando Alvarez
University of Chicago and NBER

David Argente
Yale University and NBER

Francesco Lippi
LUISS and EIEF

Esteban Méndez
Central Bank of Costa Rica

Diana Van Patten
Yale University and NBER

Table of Contents

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

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

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

A.3 Computing the Equilibrium Set 5

Appendix B. Proofs and ancillary results **6**

B.1 Solution for $\tilde{a}(x)$ and $\tilde{v}(x)$ 12

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

Appendix C. Perturbation of the Stationary Equilibrium **15**

C.1 Linearization and Solution of the KB Equation 16

C.2 Linearization and Solution of the KF Equation 17

C.3 Equilibrium in the Perturbed MFG 19

Appendix D. Planning Problem **22**

D.1 Planning Problem, stationary case 22

D.2 Dynamics of N and Flow of Adoption Cost 23

D.3 Derivation of the PDE's for the Planner's Problem 24

D.4 Solution of the Stationary Planning Problem 31

D.5 Perturbation and Stability of Invariant Distribution 33

D.6 Perturbation of the Planning Problem 35

Appendix E. A "Pure" Learning Model **39**

E.1 Proofs for the Learning Model 45

Appendix F. HJB Equations for $a(x, t)$ and $v(x, t)$ **50**

Appendix G. Empirical Appendix **51**

G.1 Descriptive Figures and Summary Statistics: SINPE 51

G.2 Evidence on Selection at Entry: Robustness 55

Appendix H. Quantitative Exercises **58**

H.1 Elements of Identification 58
H.2 Variation Across Networks 58
H.3 Only Learning: $\vartheta = 0$ 61
H.4 Comparative Statics 62

Appendix for Online Publication

A Discretization and Computation of Equilibrium

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

We define the discretization of the model as follows:

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

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

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

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

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

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

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

Thus we define:

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

and write:

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

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

$$\begin{aligned} a_{1,j-1} &= \Delta_t \, x_1 \, (\theta_0 + \theta_n N_j) + (1 - \rho \Delta_t) \left[(1-p) a_{1,j} + p a_{2,j}\right] \\ a_{I,j-1} &= \Delta_t \, x_I \, (\theta_0 + \theta_n N_j) + (1 - \rho \Delta_t) \left[p a_{I-1,j} + (1-p) a_{I,j}\right] \end{aligned}$$

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

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

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

$$\begin{aligned} &\frac{1}{\Delta_t} = \frac{J-1}{T} > \rho \text{ and} \\ &\sigma \frac{\sqrt{\Delta_t}}{\sqrt{1 - \rho \Delta_t}} = \sigma \frac{\sqrt{T}}{\sqrt{J - 1 - \rho T}} < \Delta_x = \frac{U}{I-1} \end{aligned}$$

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

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

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

$$\begin{aligned} \tilde{a}_1 &= \Delta_t \, x_1 \, (\theta_0 + \theta_n N_{ss}) + (1 - \rho \Delta_t) \left[(1-p) \tilde{a}_1 + p \tilde{a}_2\right] \\ \tilde{a}_I &= \Delta_t \, x_I \, (\theta_0 + \theta_n N_{ss}) + (1 - \rho \Delta_t) \left[p \tilde{a}_{I-1} + (1-p) \tilde{a}_I\right] \end{aligned}$$

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

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

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

$$\begin{aligned}
v_{1,j-1} &= \max\left\{-c + a_{1,j}\,,\; (1-\rho\Delta_t)\left[(1-p) v_{1,j} + p v_{2,j}\right]\right\} \\
v_{I,j-1} &= \max\left\{-c + a_{I,j}\,, (1-\rho\Delta_t)\left[p v_{I-1,j} + (1-p) v_{I,j}\right]\right\}
\end{aligned}$$

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

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

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

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

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

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

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

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

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

$$\begin{aligned}
\tilde{v}_1 &= \max\left\{-c + \tilde{a}_1\,,\; (1-\rho\Delta_t)\left[(1-p)\tilde{v}_1 + p\tilde{v}_2\right]\right\} \\
\tilde{v}_I &= \max\left\{-c + \tilde{a}_I\,,\; (1-\rho\Delta_t)\left[p\tilde{v}_{I-1} + (1-p)\tilde{v}_I\right]\right\}
\end{aligned}$$

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

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

we can write it as:

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

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

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

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

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

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

There is also the corresponding stationary distribution 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 the 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 stationary equilibruim. 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

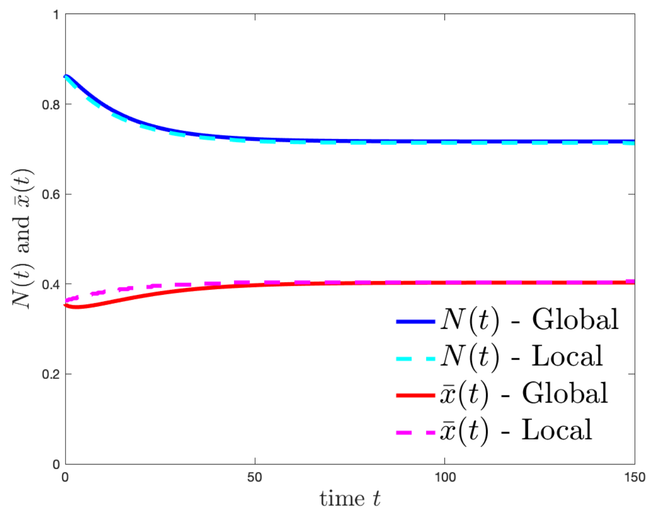
$$N^{k+1} = \mathcal{F}\left(N^k; m_0, D_T\right) \equiv \mathcal{N}\left(\mathcal{X}\left(N^k; D_T\right); m_0\right) \text{ for } k = 0, 1, 2, \ldots$$

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 corresponding to the stationary value function 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 corresponding stationary distribution (recall that $\tilde{m}$ is the density of the stationary distribution of agents without the technology). Second, the shock (deviation from the stationary distribution) 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 stationary distribution. The other feature apparent with these computations is that the stationary equilibrium is stable even starting far away from the stationary distribution.

Figure A1: Global vs Local Solutions



B Proofs and ancillary results

Proof. (of Lemma 2). 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:

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

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

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

$$-c + \mathbb{E}\left[\int_0^\infty e^{-\rho t} x(t)\left(\theta_0 + \theta_n N(t)\right) dt \,|\, x(0) = x_1\right] \geq \\ \mathbb{E}\left[-ce^{-\rho\tau_1} + \int_{\tau_1}^\infty e^{-\rho t} x(t)\left(\theta_0 + \theta_n N(t)\right) dt \,|\, x(0) = x_1\right]. \tag{35}$$

or equivalently that

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

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

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

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

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

Using this strict inequality we get a contradiction with equation (35), and hence we establish the desired result. □

We give a normalization of the primal problem that is useful in the empirical application. The lemma shows that the problem features five independent parameters, as U and θ_0 can be normalized without affecting the dynamics of the technology diffusion.

Lemma 9. The problem with parameters $\{c, \rho, \nu, \sigma, \theta_0, \theta_n, U\}$, initial condition m_0, $f(x) = \frac{1}{U}$ 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 $\hat{v}\,(z,t) \equiv \theta_0 v\,(zU, t)$.

Proof. (of Lemma 9). 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 Lemma 2). For this proof we set up the problem as a stopping time problem. We first prove a useful result in Lemma 10, 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 Lemma 2. 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 10. □

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 10. 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 10). 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:

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

which follows because for each sample path ω we have:

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

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

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

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

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

and for each ω

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

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

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

Thus

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

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 Lemma 3) The fraction of agents that have not adopted at time t can be written as

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

where

$$P(x,s,t) = Pr\left[X(r) \leq \bar{x}(r),\ \text{for all } r \in [s,t] \mid X(s) = x\right] e^{-\nu(t-s)} \tag{36}$$

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 (36) is the fraction of those that have not adopted at in the initial distribution, and still have not adopted, and survive, at time t. The second term keeps tract of those cohort that have died at time s, and replaced by new agents, and themselves survive and not adopt up to time t.

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

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

□

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

We note that $\{N : \{0,\Delta_t,\ldots,T\} \to [0,1]\} = [0,1]^J$ where J is the integer that defines Δ_t. This set is a complete lattice. This function is monotone by virtue of Lemma 2 and Lemma 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 Lemma 2 and Lemma 3. □

Proof. (of Proposition 1) 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 Lemma 4) 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 B.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 B.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. □

Proof. (of Lemma 5) 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 $-(\tanh(\bar{x}\gamma) - \bar{x}\gamma \operatorname{sech}^2(\bar{x}\gamma)) = -(\tanh(\bar{x}\gamma) - \bar{x}\gamma \tanh'(\bar{x}\gamma)) < 0$, where we used that $\tanh(z)$ is strictly concave for $z > 0$. □

Proof. (of Proposition 2). In the deterministic case, i.e., when $\sigma = 0$, there are at most two interior stationary equilibrium (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 stationary equilibrium we write

$$N^a\left(\bar{x}^j(c)\right) = N^o\left(\bar{x}^j(c), c\right) \tag{37}$$

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 stationary equilibrium 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 (37) 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 stationary equilibrium:

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

Following exactly the same steps we get:

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

□

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

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

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

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

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

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

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

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

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

which has solution:

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

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

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

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

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

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

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

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

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

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

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

which gives equation (16) 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 $\sigma = 0$.

B.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 $\bar{\theta}$.

C Perturbation of the Stationary Equilibrium

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

$$m_0(x) = \tilde{m}(x) + \epsilon\, \omega(x) \ . \tag{38}$$

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:

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

C.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

$$
\rho d(x,t) = x\theta_n n(t) + \frac{\sigma^2}{2} d_{xx}(x,t) + d_t(x,t) \tag{39}
$$

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:

$$
\begin{aligned}
d(\bar{x}_{ss},t) &= 0 \\
\tilde{D}_{xx}(\bar{x}_{ss})\bar{y}(t) + d_x(\bar{x}_{ss},t) &= 0 \\
d_x(0,t) &= 0
\end{aligned} \tag{40}
$$

for $t \in [0,T]$ and $d(x,T) = 0$ for $x \in [0,\bar{x}_{ss}]$. Note that equation (40) 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 (39) with respect to x and combining it with equation (40) we find

$$
\bar{y}(t) = \frac{\theta_n}{\tilde{D}_{xx}(\bar{x}_{ss})} \int_t^T G(\tau - t) n(\tau) d\tau \tag{41}
$$

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 11. The solution for the KBE equation for d, satisfying the p.d.e. in equation (39), and the boundary conditions in equation (40), 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:

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

$$\hat{d}_j(t) \equiv \int_t^T e^{-\psi_j(\tau - t)} \hat{z}_j(\tau) d\tau \qquad \text{for } t \in [0,T]$$

$$\hat{z}_j(t) \equiv \theta_n n(t) \frac{\langle \varphi_j, x\rangle}{\langle \varphi_j, \varphi_j\rangle} = \theta_n n(t) \frac{2\bar{x}_{ss}}{\left(\frac{1}{2}+j\right)\pi}\left(1 - \frac{\cos(\pi j)}{\pi(j+\frac{1}{2})}\right) \qquad \text{for } t \in [0,T]$$

$$\text{where } \psi_j \equiv \rho + \frac{\sigma^2}{2}\left(\frac{\pi(\frac{1}{2}+j)}{\bar{x}_{ss}}\right)^2 \quad \text{and} \quad \hat{d}_j(T) = 0$$

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 (39) 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 (40), is

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

C.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:

$$p_t(x,t) = \frac{\sigma^2}{2} p_{xx}(x,t) - \nu p(x,t) \tag{42}$$

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

$$\tilde{m}_x(\bar{x}_{ss})\bar{y}(t) + p(\bar{x}_{ss}, t) = 0 \tag{43}$$
$$p_x(0, t) = 0$$

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

$$p(0, x) = \omega(x) \tag{44}$$

The solution for p satisfies the p.d.e given in equation (42), its boundary conditions in equation (43), and the initial condition in equation (44). We have

LEMMA 12. The solution for the KFE equation for p, satisfying the p.d.e given in equation (42), the boundary conditions in equation (43), and the initial condition in equation (44), is given by

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

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:

$$\begin{aligned} n(t) &= -\int_0^{\bar{x}_{ss}} p(x,t)dx \\ &= n_0(t) + \frac{\tilde{m}_x(\bar{x}_{ss})\sigma^2}{\bar{x}_{ss}} \int_0^t J(t-\tau)\bar{y}(\tau)d\tau \end{aligned} \tag{45}$$

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

C.3 Equilibrium in the Perturbed MFG

Recall that from equation (41), $\bar{y}(t)$ 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 (45) 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 (41) and equation (45) we get

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

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

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

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

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

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

so that

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

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

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

and as $T \to \infty$

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

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

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

A sufficient condition for the existence and uniqueness of the equilibrium IRF, i.e., of the uniqueness and existence of a solution to equation (23) 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 (17), and we calculate $\tilde{m}_x(\bar{x}_{ss})$ from $\tilde{m}(x) = \frac{1}{U}\left(1 - \frac{\cosh(\gamma x)}{\cosh(\gamma\bar{x}_{ss})}\right)$.

D Planning Problem

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

D.1 Planning Problem, stationary case

A stationary equilibrium is characterized by two constants N_{ss} and $\bar{x}_{ss}$ that solve the time invariant version of the p.d.e. stated in Section 5. The p.d.e. for non-adopters in the stationary case becomes the following o.d.e.:

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

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 for the stationary case. The details of the solution can be found in Appendix D.4. The following proposition summarizes the solution of planning problem at a stationary distribution.

Proposition 3. 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 3 indicates that the solution of the stochastic version of the planning problem also has the option value present in the equilibrium. This proposition can be used to show that the stationary level of adoption in the planning problem is higher than the adoption level of the the high-activity stationary equilibrium.

D.2 Dynamics of N and Flow of Adoption Cost

Recall that

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

Taking the derivative with respect to time

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

where the first term is zero because of 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.

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

To derive the problem in continuous time, we write the adoption problem in a discrete-time discrete-state setup. We do so by using the finite-difference approximation and then considering 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 is 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:

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

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 the measure $m_{i,t}$ to the measure $g_{i,t}$. Note that $m_{i,t} + g_{i,t} = \frac{1}{I}$ since the sum of the two is the invariant (uniform) distribution. The initial conditions are $g_{i,0} = 0 \; \forall i$ and $m_{i,0} = \frac{1}{I}$ all non-adopters. The law of motion of the state is then:

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

which can be written in vector notation as:

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

where L is an $I \times I$ stochastic matrix which depends on I, ν, σ^2, Δ and Δ_x. We assume that $\Delta\left(\nu + (\sigma/\Delta_x)^2\right) < 1$ so that all implied probabilities are positive.

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

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:

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

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:

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

The derivative of Lagrangian with respect to α_{it} gives:

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

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

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

where

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

We derive the **smooth pasting** condition 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:

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

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

In summary, a planner's problem is given by $\{\bar{x}(t), \lambda(x,t), m(x,t)\}$, namely the path of the 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 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 $+\infty$).

D.4 Solution of the Stationary 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

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

D.5 Perturbation and Stability of Invariant Distribution

In this section we analyze the linearization of the planning problem around its stationary distribution. This linearization is analogous to the one for the equilibrium in Section 4.

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 stationary value N_{ss} and Z_{ss}. We will denote this Gateaux derivative by $\bar{y}$.

Proposition 4. Let λ_T be equal to the stationary value function $\tilde{\lambda}$ corresponding to that invariant distribution. Let $n : [0,T] \to \mathbb{R}$ and $z : [0,T] \to \mathbb{R}$ be two arbitrary perturbations. Then

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

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 Lemma 6.

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 invariant distribution $\tilde{m}$. In particular, we consider paths defined by $\bar{x}(t) = \bar{x}_{ss} + \epsilon\, \bar{y}(t)$ around the invariant threshold x_{ss}, and around the invariant distribution $m_0(x) = \tilde{m}(x) + \epsilon \omega(x)$. We will denote this Gateaux derivative by z.

Proposition 5. Let $\tilde{m}$ be the corresponding invariant 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

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

where

$$z_0(\omega)(t) \equiv -\sum_{j=0}^{\infty} \frac{\bar{x}_{ss}^2(\pi j + \frac{1}{2} - \cos(j\pi))}{\pi(\frac{1}{2} + j)} \frac{\langle \varphi_j, \omega \rangle}{\langle \varphi_j, \varphi_j \rangle} e^{-\mu_j t} \text{ and}$$
$$H_{zy}(q) = \tilde{m}_x(\bar{x}_{ss})\sigma^2 \sum_{j=0}^{\infty} \eta_j e^{-\mu_j q}$$

where $\varphi_j, \tilde{m}_x, \mu_j$ and γ are defined as in Lemma 7.

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 D.6.1.

Theorem 4. Let $\bar{x}_{ss}$ be the invariant threshold of the planner problem, with its corresponding N_{ss}, Z_{ss}, and let $\tilde{m}$ be the corresponding invariant distribution of non-adopters. Let $m_0(x) = \tilde{m}(x) + \epsilon\omega(x)$. Let λ_T be equal to the stationary value function $\tilde{\lambda}$. The linearized equilibrium must solve

$$\bar{y}(t) = \bar{y}_0(t) + \tilde{\Theta} \int_0^T \tilde{K}(t,s)\bar{y}(s)ds \text{ where} \tag{46}$$
$$\bar{y}_0(\omega)(t) \equiv \int_t^T G_{yn}(\tau - t)n_0(\omega)(\tau)d\tau + \int_t^T G_{yz}(\tau - t)z_0(\omega)(\tau)d\tau$$

where n_0 is derived in Lemma 7, z_0 is derived in Proposition 5, $\tilde{\Theta} \equiv \frac{\theta_n \tilde{m}_x(\bar{x}_{ss})\sigma^2}{\tilde{\lambda}_{xx}(\bar{x}_{ss})\bar{x}_{ss}}$ and where the kernel $\tilde{K}$ is given by

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

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 (46) 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 invariant density of non-adopters. In this case, we let $m_0(x)$ be the invariant distribution of no-adopters of the problem, so that the shock resembles starting an equilibrium with lower adoption than that prescribed by the planning solution.

D.6 Perturbation of the Planning Problem

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

$$\rho\ell(x,t) = x\theta_n n(t) + \theta_n z(t) + \frac{\sigma^2}{2}\ell_{xx}(x,t) + \ell_t(x,t) \tag{47}$$

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:

$$\begin{aligned} \ell(x,T) &= 0 \\ \ell_x(0,t) &= 0 \\ \ell(\bar{x}_{ss},t) &= 0 \\ \tilde{\lambda}_{xx}(\bar{x}_{ss})\bar{y}(t) + \ell_x(\bar{x}_{ss},t) &= 0 \end{aligned} \tag{48}$$

Proposition 6. 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 $c_j = 2\left(1 - \frac{\cos(\pi j)}{\pi(j+\frac{1}{2})}\right)$.

D.6.1 Perturbation Analysis of the Planning Problem

Recall that from equation (48), $\bar{y}(t)$ is equal to

$$\begin{aligned}
\bar{y}(t) &= \frac{-\ell_x(\bar{x}_{ss}, t)}{\tilde{\lambda}_{xx}(\bar{x}_{ss})} \\
&= \int_t^T \frac{\theta_n}{\tilde{\lambda}_{xx}(\bar{x}_{ss})}\sum_{j=0}^{\infty} c_j e^{-\psi_j(\tau-t)}n(\tau)d\tau + \int_t^T \frac{2\theta_n}{\tilde{\lambda}_{xx}(\bar{x}_{ss})\bar{x}_{ss}}\sum_{j=0}^{\infty} c_j e^{-\psi_j(\tau-t)}z(\tau)d\tau \\
&= \int_t^T G_{yn}(\tau-t)n(\tau)d\tau + \int_t^T G_{yz}(\tau-t)z(\tau)d\tau
\end{aligned} \tag{49}$$

The expression for $n(t)$ is given by equation (45) 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

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

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 (49) can be written as

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

where

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

and

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

$$= \int\limits_0^T \int\limits_{\max\{t,s\}}^T G_{yn}(\tau-t)H_{ny}(\tau-s)\bar{y}(s)ds\,d\tau + \int\limits_0^T \int\limits_{\max\{t,s\}}^T G_{yz}(\tau-t)H_{zy}(\tau-s)\bar{y}(s)ds\,d\tau$$

with

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

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

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

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

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

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

Then, we find

$$\int_0^T M(t,s)\bar{y}(s)ds = \int\limits_0^T \left(\int\limits_{\max\{t,s\}}^T G_{yn}(\tau - t)H_{ny}(\tau - s)\bar{y}(s)d\tau + \int\limits_{\max\{t,s\}}^T G_{yz}(\tau - t)H_{zy}(\tau - s)\bar{y}(s)d\tau \right) ds$$
$$= \tilde{\Theta}(\bar{x}_{ss}) \int_0^T \int_{\max\{t,s\}}^T \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} c_j e^{-\psi_j(\tau-t)} e^{-\mu_i(\tau-t)}\bar{y}(s)d\tau\, ds$$
$$+ \tilde{\Theta}(\bar{x}_{ss}) \int_0^T \int_{\max\{t,s\}}^T \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} c_i e^{-\psi_j(\tau-t)} e^{-\mu_i(\tau-t)}\bar{y}(s)d\tau\, ds$$

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 (49) 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 $e^{-rt}M(t,s) = e^{-rs}M(t,s)$

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

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

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

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

Optimal Adoption. Now we turn to the decision of agents. The uninformed agents have no decision to make. The decision problem of those that are informed is similar to the stationary 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 stationary 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 7. 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 $\bar{x}$:

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

The expression in the right hand side of $N(t)$ in Proposition 7 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 8. 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

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

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

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

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

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

$$N(t) = I(t) - \tfrac{\bar{x}}{U}\left(H(t)I(0) + \beta_0\int_0^t H(t-\tau)I(\tau)\,(1-I(\tau))\,d\tau\right) \text{ where}$$

$$H(z) \equiv \sum_{j=0}^{\infty}\omega_j e^{-\mu_j z} \text{ with } \omega_j \equiv \frac{2}{\left(\pi(j+\frac{1}{2})\right)^2} > 0 \text{ and } \sum_{j=0}^{\infty}\omega_j = 1\,.$$

Combining the expression for $N(t)$ in Proposition 9 with the path of $I(t)$ solved for in Proposition 8 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 the distribution $\tilde{m}$, as well as the stationary number of adopters N_{ss}.

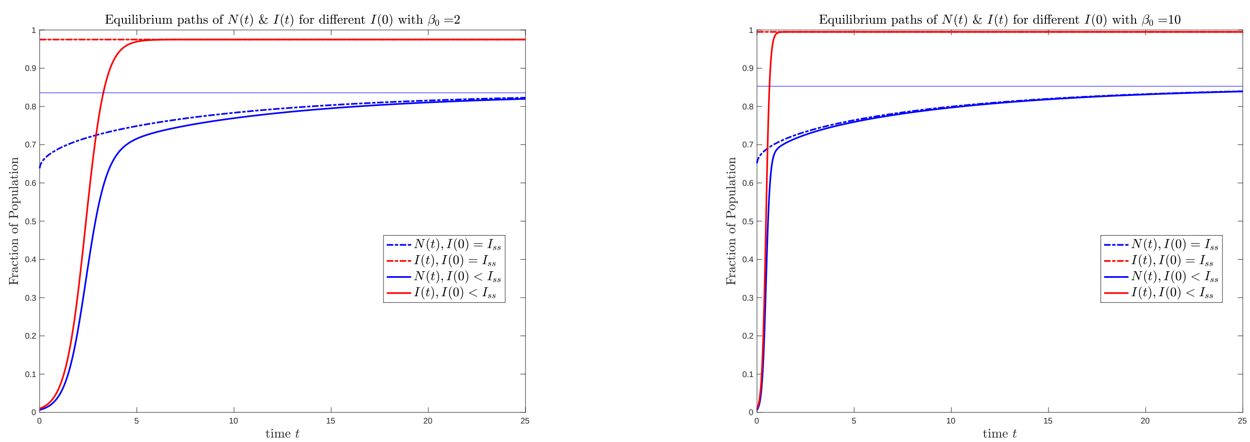
Proposition 10. Assume that $s_0(x) = S_0/U$ for all $x \in [0, U]$, that $\bar{x} < U$, $\beta(S) = \beta_0(1 - S)$, and that $\beta_0 > \nu > 0$. Then the invariant density $\tilde{m}$ is given by:

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

$$N_{ss} = I_{ss} - \int_0^{\bar{x}} \tilde{m}(x) dx = (1 - \tfrac{\nu}{\beta_0}) \left[1 - \tfrac{\bar{x}}{U} \left(1 - \frac{\tanh(\gamma \bar{x})}{\gamma \bar{x}} \right) \right] \tag{50}$$

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 (50) 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 7) 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:

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

□

Proof. (of Proposition 8) 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 $x = U$:

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

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

$$S' = -(\nu + \beta_0(1-S))\,S + \nu = (1-S)(\nu - \beta_0 S)$$
$$= \nu(1-S)\left(1 - \frac{S}{S^*}\right)$$

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:

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

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 9) Given the path $\{S(t)\}$ define

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

We start with

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

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

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

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

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

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

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

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 $\hat{b}_j(t)$:

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

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

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

So we can write:

$$N(t) = I(t) - \tfrac{\bar{x}}{U}\left(\sum_{j=0}^{\infty}\omega_j e^{-\mu_j t}I(0) + \beta_0\int_0^t \sum_{j=0}^{\infty}\omega_j e^{-\mu_j(t-s)}I(s)\left(1-I(s)\right)ds \right) \text{ where}$$

$$\omega_j \equiv \frac{2}{\left(\pi(j+\frac{1}{2})\right)^2} > 0 \text{ and } \sum_{j=0}^{\infty}\omega_j = 1\,.$$

Defining

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

we can write:

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

□

Proof. (of Proposition 10) 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 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 $x = 0$.

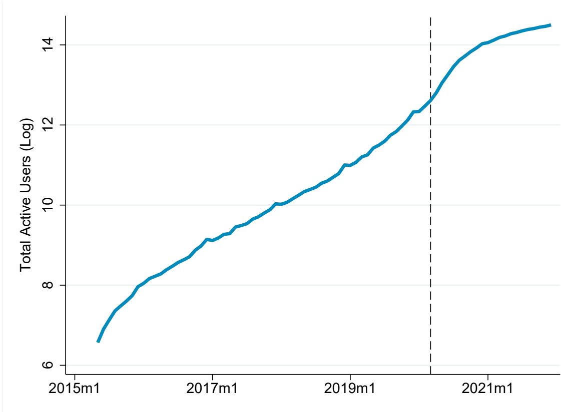
G Empirical Appendix

G.1 Descriptive Figures and Summary Statistics: SINPE

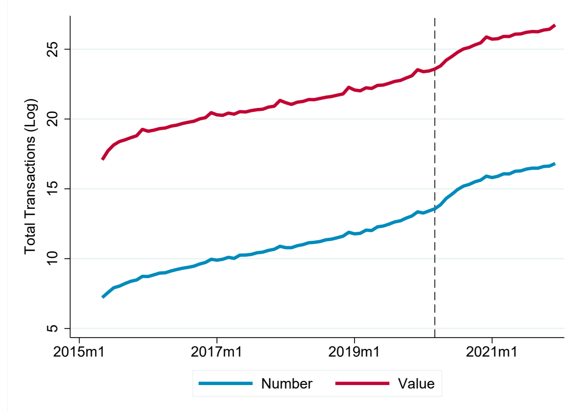
The technology diffused gradually. The aggregate adoption of SINPE has grown at a constant rate over time since its inception in 2015, as shown in Figure G1 using monthly data on the total number of adopters.[39] By 2021, close to 79% of the adult population in the country owned a bank account, and over 60% of adults were SINPE subscribers who had not deactivated their account. Moreover, the value of annual transactions in SINPE is approximately 10% of GDP. Thus, this setting has the unique feature of allowing us to study the adoption of mobile payments in the entire population of the country, across many years since the inception of the technology, and until it reached almost the universe of the country's adult population. The fact that adoption occurs gradually coincides with the dynamics of our dynamic stochastic model, and rules out the deterministic case in which adoption happens on impact.

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

Figure G1: Users, Transactions, and Value of Transactions



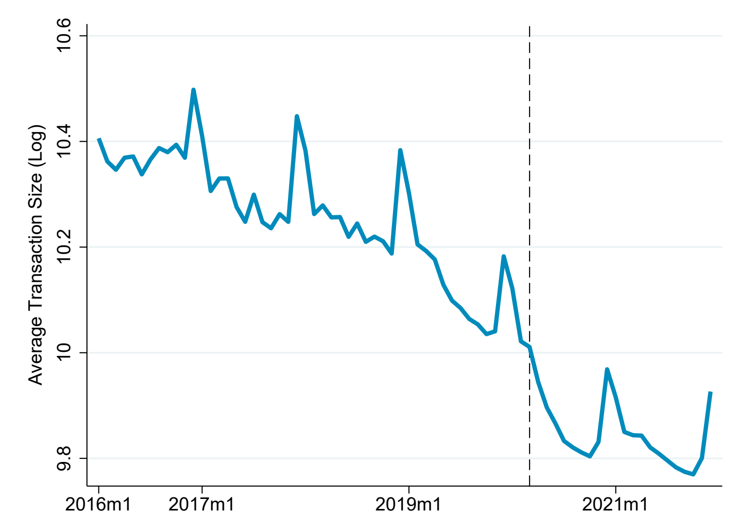
(a) Total Users



(b) Number and Value of Transactions

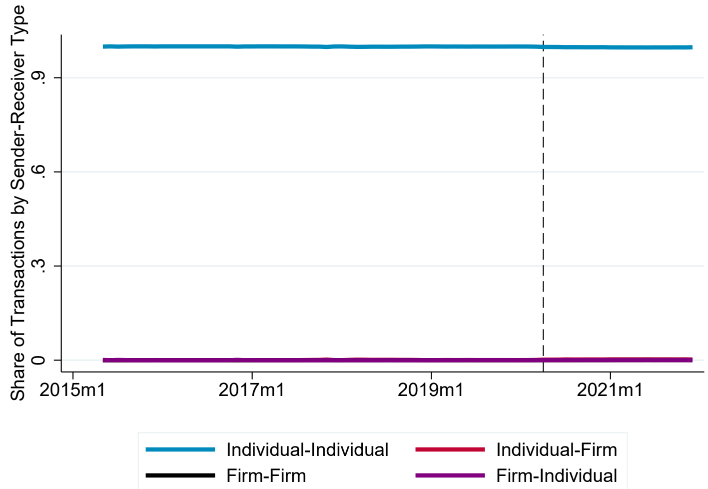
Notes: Panel (a) shows total active SINPE users. We include only active subscriptions by individuals, as users have the option of deactivating their account. Panel (b) shows both total transactions in the application and total value of transactions by individuals. Both figures include a vertical dashed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

Figure G2: Average Transaction Size



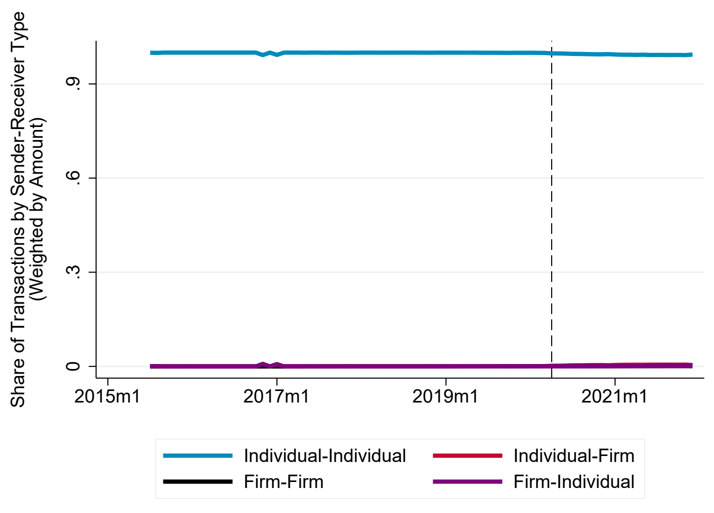
Notes: The figure shows the evolution of the average transaction size in SINPE.

Figure G3: Transactions by Sender-Receiver Type



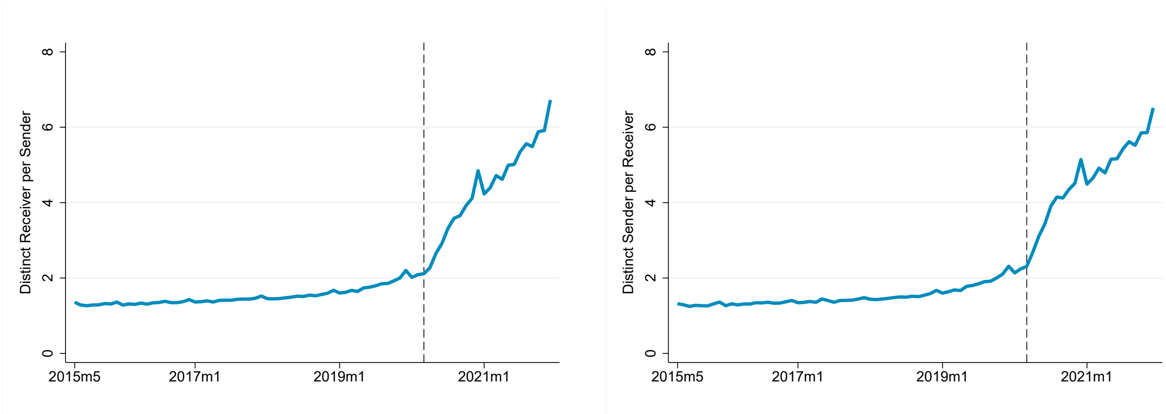
Notes: Transactions are classified according to the type of user. Individuals correspond with Costa Rican adult citizens. Firms correspond with formal enterprises.

Figure G4: 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 G5: Mean Number of Connections per User



(a) Receivers (b) Senders

Figure G6: Average Age at the Time of Adoption

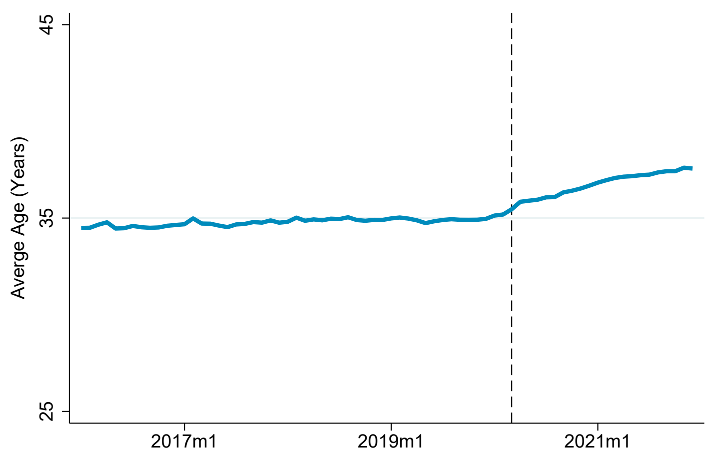


Table G1: Mean Share of Transactions Within Network (2015-2021)

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

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

G.2 Evidence on Selection at Entry: Robustness

Table G2: Amount Transacted and Size of Network at Entry

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

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

G.2.1 Details on Mass Layoffs

This section provides additional details on the choices made to construct the variables and sample used for the analysis of mass layoffs in Section 6.2.

Definition of a Mass Layoff To define a mass layoff, we follow Davis and Von Wachter (2011) and identify establishments with at least 50 workers that contracted their monthly employment by at least 30% *and* which did not recover in the following 12 months. We define a recovery as a firm which went back to its initial size (or above) within the following 12 months. Given this definition, the descriptive statistics of firms and workers impacted by a mass layoff are reported in Table G3.

Exercises Based on Stayers We also conduct several exercises based on stayers, i.e., workers who remain at a firm after it experiences a mass layoff. Therefore, we refine the definition above to exclude odd cases. Namely, the stayer must remain at the firm at least 6 months after the mass layoff (this applies to cases in which there was more than one wave of layoffs) and we exclude cases where the firm had experienced an increase of at least 30% in its workforce within the 6 months prior to the mass layoff (this applies to a few instances of, for instance, seasonal hiring or project-specific hiring). It is worth noting that (i) these refinements do not have any significant effect on the results for movers, and (ii) if anything, these refinements lead to smaller results for stayers.

Table G3: Mass Layoffs: Descriptive Statistics

Number of firms	292	
Number of displaced workers who had not adopted SINPE when fired	10,176	
Number of displaced workers who had adopted SINPE when fired	917	
Average firm size	264	(989)
Median firm size	94	
Average monthly wage pre-layoff, laid-off workers	$688	($732)
Average monthly wage pre-layoff, all workers	$848	($1,133)

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. We also exclude cases where the firm had experienced an increase of at least 30% in its workforce within the 6 months prior to the mass layoff. Wages were calculated based on an exchange rate of 634 colones per dollar and the last month in which workers were employed. We include mass layoffs which occurred between May 2015, when the technology was introduced, and December 2021. The last row includes the average monthly wage pre-layoff for all workers who were employed at those firms at the time of the mass layoff.

Definition of Variables We construct several variables that are used in equation (34). We now provide more details on each of them.

- $Adopt_i$ equals one if individual i adopted SINPE within 6 months after arriving at 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.

- $\sum_{t=0}^{move} \left(\ln \tilde{\mathrm{T}}_{t,\text{ new firm}} - \ln \tilde{\mathrm{T}}_{t,\text{ old firm}}\right)$ is the difference in the (log) historical transactions made by workers at the new firm and the old firm prior to the move, which aims to control for factors, other than strategic complementarities, which might facilitate adoption at the new vs. the old firm.

- $\Delta Covid_i$ controls for the change in the cumulative COVID-19 cases (transformed using the inverse hyperbolic sine function) in the individual's neighborhood across the 6 months before the layoff and after the rehiring. This change is zero for pre-pandemic years, thus, this variable is introduced using an inverse hyperbolic sine transformation, as opposed to a logarithm.

The regression described in equation (33) relies on the same variables that we described above, but also includes additional ones which we now describe.

- $\Delta \ln \tilde{\mathtt{T}}_i$ refers to the change in monthly intensity with which individual i used SINPE within 6 months *after* arriving at 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 resulting from a "learning curve." We transform $\tilde{\mathtt{T}}_i$ 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.

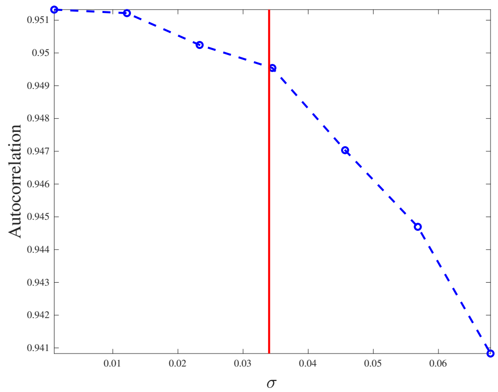
- $cohort_i$ controls for the month when individual i adopted SINPE. We include this variable to attenuate any effect resulting from learning how to better use the app.

- $\ln \sum^t \tilde{\mathtt{T}}_i$ 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.

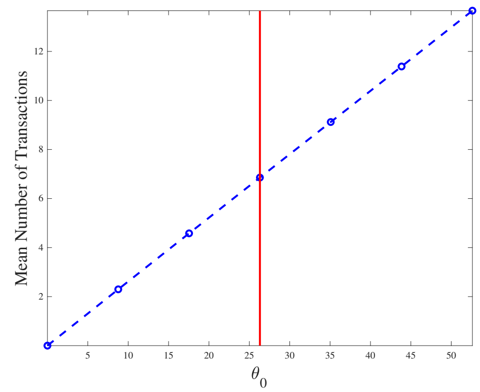
H Quantitative Exercises

H.1 Elements of Identification

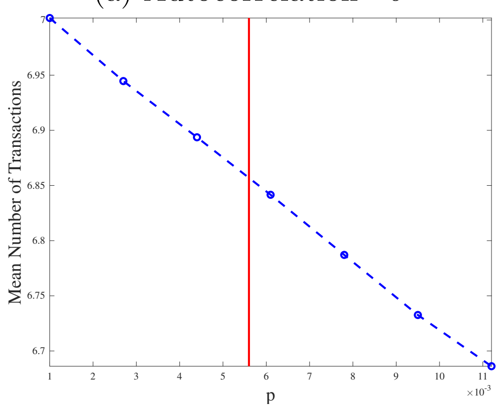
Figure H1: Sensitivity Analysis



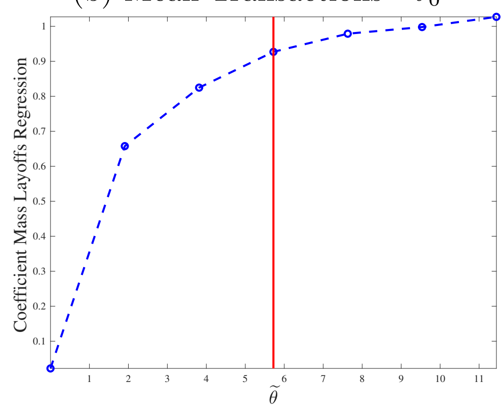
(a) Autocorrelation - σ



(b) Mean Transactions - θ_0



(c) Mean Transactions - p



(b) Coefficient of Mass Layoffs - ϑ

Notes: The graphs plot the relationship between the estimated parameters and several moments relevant for identification. We set all parameters to their baseline estimates reported in Table 3 and represented by the vertical red line. Then, we move each parameter around its estimated value holding the others constant. In panel (a) the moment reported is the autocorrelation and the parameter is σ. In panels (b) and (c) the moment reported is the average number of transactions and the parameters are θ_o and p respectively. In panel (d) the moment reported is the coefficient of the mass layoffs regression and the parameter is ϑ.

H.2 Variation Across Networks

In this section, we provide more details on the estimation of the model using variations across networks. We show that the model is consistent with both high and low adoption networks of firms, each implying a different path of adopters in equilibrium and a different optimal path of adoption in the planning problem.

Specifically, we calibrate the model by targeting moments from individuals at firms whose level of adoption is either above the median (high adoption) or below the median (low adoption). To do this, first, we assume that the externally calibrated parameters are the same in both low and high adoption networks (i.e., ν, r, β_0). We set these parameters to the same values as in the benchmark calibration. Second, although the speed of information diffusion is the same in the two networks (i.e., β_0), we assume that more people knew about the technology at entry in the high adoption networks than in the low adoption networks. Specifically, we assume that in high adoption networks, 0.133 percent of workers were informed about SINPE at its launch, while in low adoption networks, only 0.066 percent were informed. This means we add (or subtract) a third of the people in each network relative to the benchmark calibration.

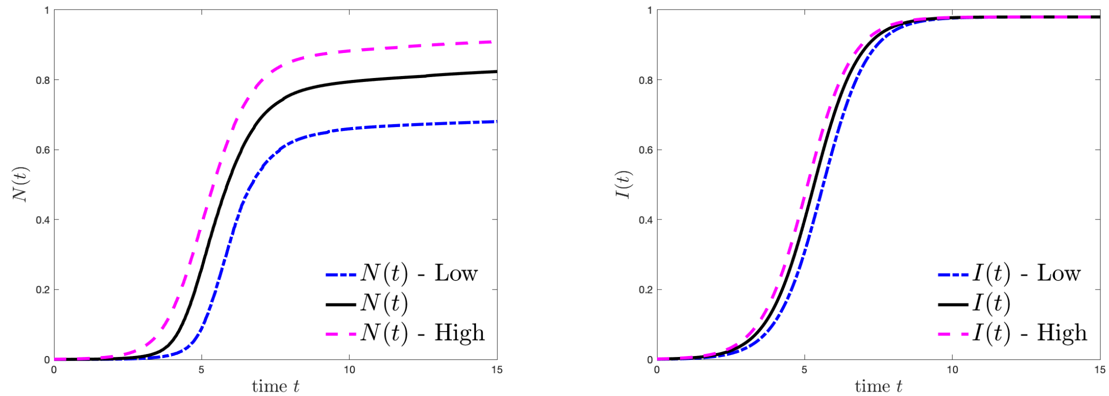
We then calibrate θ_n, θ_0, σ, and p using the simulated method of moments (SMM) for both high and low adoption networks. As before, we choose the parameters to make the model consistent with the distribution of transactions in the data and the mass layoff exercise. We follow the same procedure as in our benchmark calibration. In particular, we target the same data moments computed for different samples of workers, specifically those working at firms whose average level of adoption is either above the median, $N_{ss}^{high} = 0.96$, or below the median, $N_{ss}^{low} = 0.73$, and we assume the same coefficient for the mass layoffs regressions in both calibrations.

Table H1: Moments: Distribution of Transactions

Parameter	Value Low	Value High	Moment	Data Low	Model Low	Data High	Model High
σ	0.021	0.033	Mean Transactions	7.13	7.13	7.11	6.66
θ_0	48.29	23.01	Median Transactions	6.34	6.94	6.38	6.46
p	0.0259	0.0029	Absolute Changes	3.83	2.90	3.53	2.70
$\vartheta \equiv \frac{\theta_n}{\theta_0}$	3.16	6.38	Coefficient Mass Layoffs	0.97	0.97	0.97	0.96
			Autocorrelation Transactions	1.00	0.93	0.92	0.95

Table H1 shows the estimated parameters for both calibrations. We estimate a higher level of strategic complementarities (i.e., higher ϑ) in networks with high adoption and a higher convexity in the cost of conducting transactions in low adoption networks (i.e., higher p). The estimated variance in both calibrations is slightly higher in networks with high adoption. The table also shows that the calibrated models are quantitatively consistent with the empirical distribution of transactions for high and low adoption samples.

Figure H2: Path of Adopters (Long-Run)

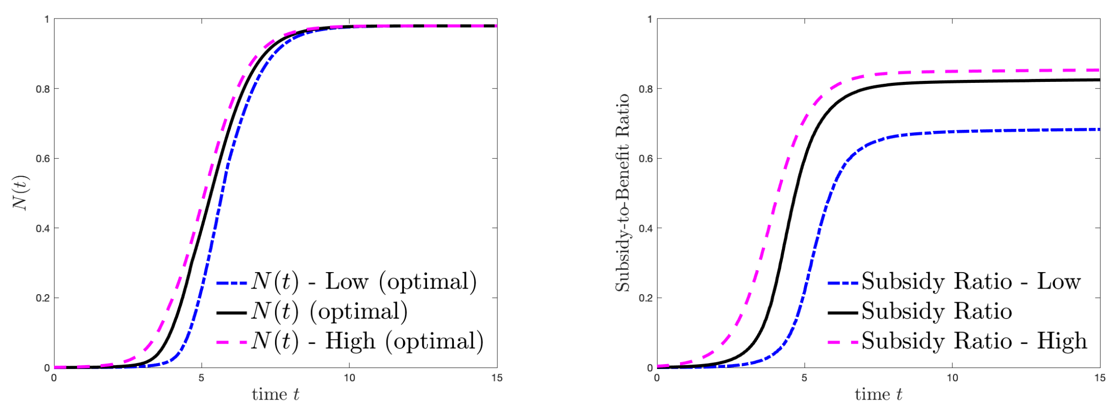


(a) Path of $N(t)$ (b) Path of $I(t)$

Notes: Panel (a) shows the share of adopters, $N(t)$, predicted by the model under our baseline calibration, high adoption, and low adoption calibrations. Panel (b) shows the share of informed agents, $I(t)$, under each of the calibrations, respectively.

Panels (a) and (b) of Figure H2 show the paths of $N(t)$ and $I(t)$. Each panel includes the results for the high and low adoption versions of the model relative to the results of our benchmark calibration. In the high adoption network, 96% of the population adopts the application. In the low adoption network, only 73% of the population adopts in the stationary equilibrium. As before, most people are informed about the technology within the first 7 years, and in the stationary distribution, approximately 98% of the population knows about the application.

Figure H3: Planning Problem: Solution and Optimal Subsidy



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

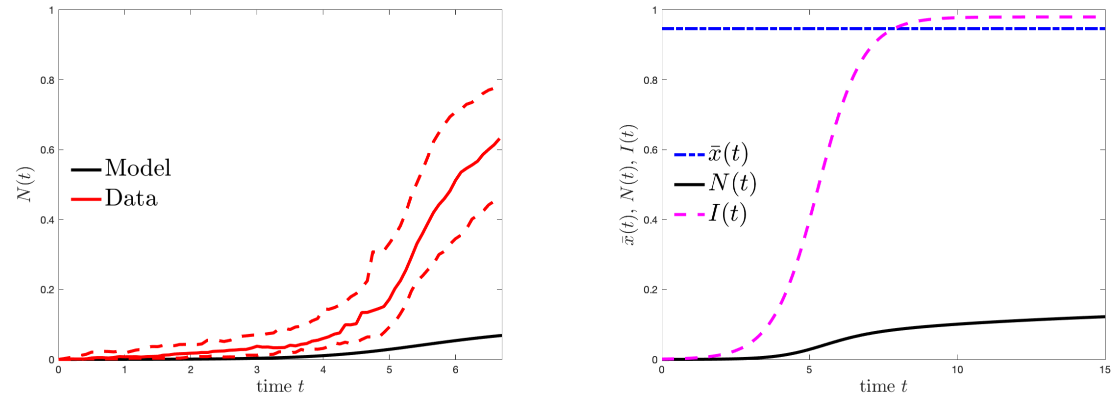
Notes: Panel (a) shows the optimal levels of adoption, $N(t)$ (optimal), according to the solution of the planning problem under the benchmark calibration, high adoption calibration, and low adoption calibration. Panel (b) shows the path of the ratio between the optimal subsidy $\theta_n Z(t)$ and the flow benefit of the average adopter, $Z(t)(\theta_0 + \theta_n N(t))$, under the benchmark calibration, high adoption calibration, and low adoption calibration.

Panel (a) of Figure H3 shows the optimal adoption path for high and low adoption networks, relative to our benchmark calibration. As before, during the first three years after the launch of the technology, the optimal level of adoption is similar to that of the high-adoption equilibrium. Afterward, the optimal path of adopters from the planning problem is higher. The optimal subsidy, shown in panel (b), increases over time as the externality increases. Since $\theta_n N$ is higher in the high adoption network, the subsidy-to-benefit ratio in this case is also higher. Nonetheless, in all versions of the model, the optimal subsidy pushes the economy toward universal adoption.

H.3 Only Learning: $\vartheta = 0$

In this section, we examine the behavior of a model without strategic complementarities. Not surprisingly, if we keep all parameter at their baseline value and set $\theta_n = 0$, the model predicts much lower adoption at its stationary equilibrium, $N_{ss} = 0.19$. The adoption in this model is purely determined by the idiosyncratic benefits of the technology. Panel (a) of Figure H4 shows that convergence to the stationary equilibrium takes longer in a model without complementarities. Recall that the model matches the fraction of agents informed about the technology three years after it was launched. Panel (b) suggests that in a pure learning model, adoption would be much slower than that observed in the data. Panel (b) also shows that the path of $\bar{x}(t)$ in the model with only learning is flat, which indicates there is no selection in the adoption of the technology in contrast to what is observed in the data. Importantly, this version of the model is constrained efficient: the optimal subsidy to use the technology is zero.

Figure H4: Path of Adopters - Only Learning (Short-Run and Long-Run)



(a) Model vs Data

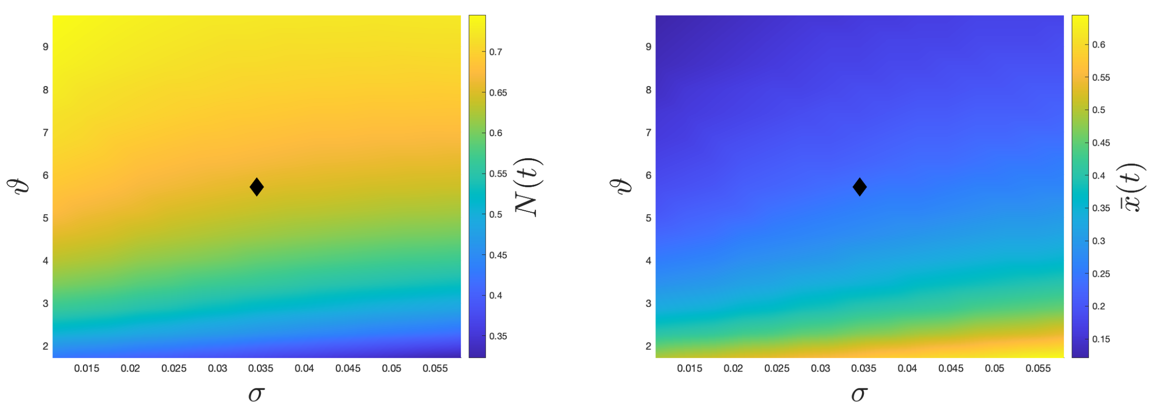
(b) Long-Run Path

Notes: Panel(a) compares the path of adopters in the model with $\theta_n = 0$ and in the data. The solid red line shows the patterns of diffusion of the technology in the median firm, 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 10^{th} and 90^{th} percentiles. Panel (b) shows the share of informed agents, $I(t)$, the share of adopters, $N(t)$, and the levels of $\bar{x}(t)$ predicted by the model under our baseline calibration but setting $\theta_n = 0$.

H.4 Comparative Statics

H.4.1 Stochastic Model: Short-Run

Figure H5: Adoption: $N(t)$ and $\bar{x}(t)$



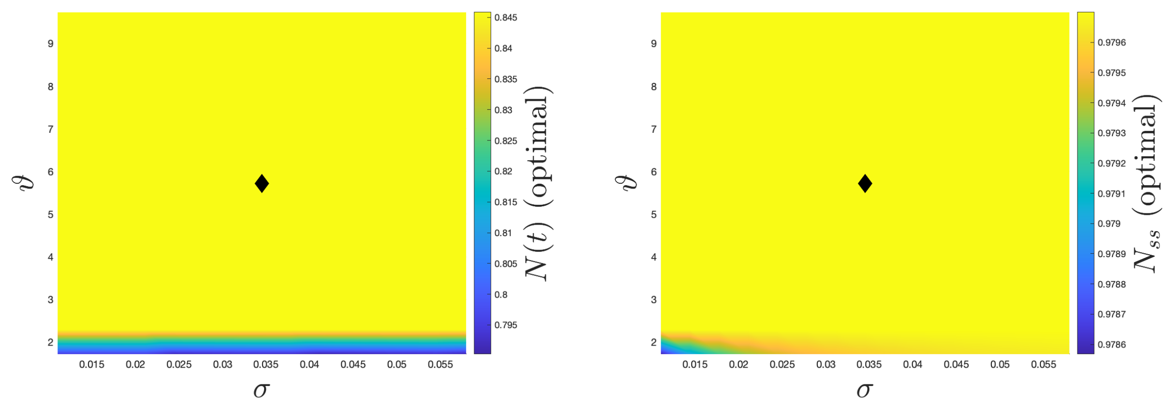
(a) Comparative Statics: $N(t)$

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

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

H.4.2 Stochastic Model: Planning Problem

Figure H6: Optimal Adoption: $N(t)$ and N_{ss}

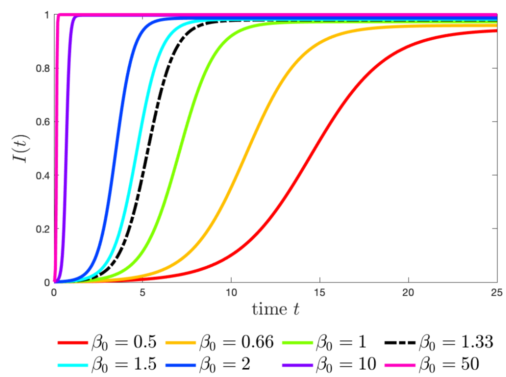


(a) Comparative Statics: $N(t)$ (b) Comparative Statics: N_{ss}

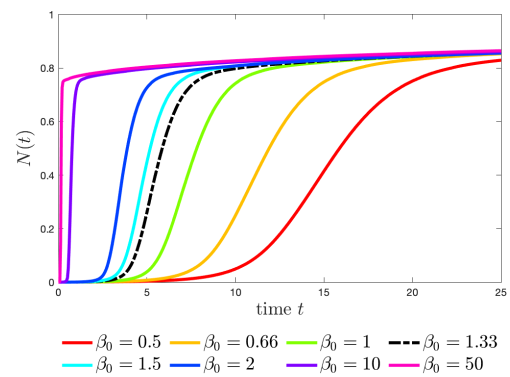
Notes: Panel (a) shows how $N(t)$ changes 7 years after the technology was launched with ϑ and σ, keeping the rest of the parameters constant. The black diamonds indicate the levels of ϑ and σ in our baseline calibration. Panel (b) shows the same comparative static for N_{ss}.

H.4.3 Learning

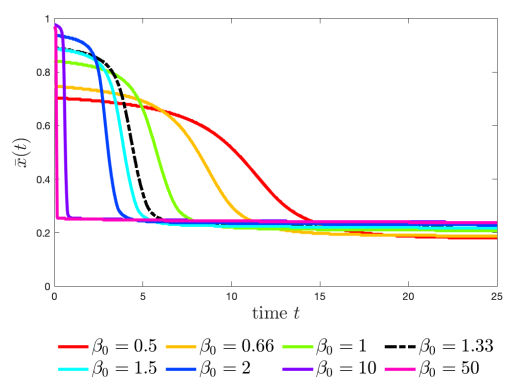
Figure H7: Heterogeneity: β_0



(a) Path of $I(t)$



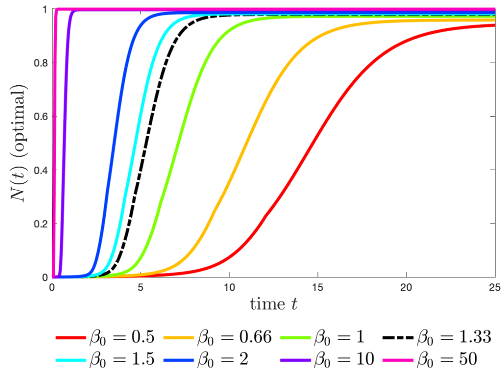
(b) Path of $N(t)$



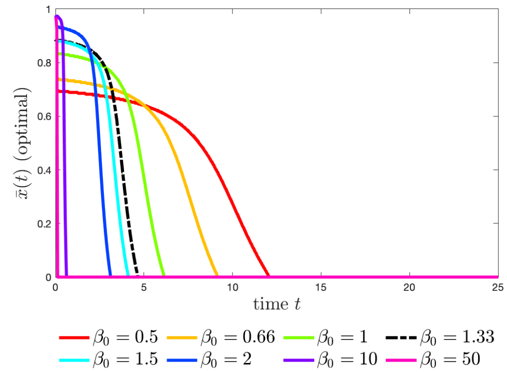
(c) Path of $\bar{x}(t)$

Notes: Panel (a) shows the share of informed agents, $I(t)$. Panel (b) shows the share of adopters, $N(t)$, and panel (c) shows the path of $\bar{x}(t)$ in the model. Each panel shows paths for different values of β_0.

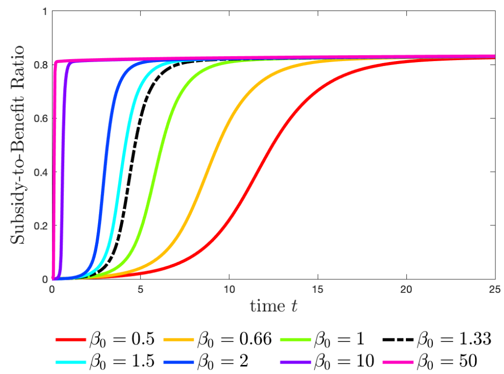
Figure H8: Heterogeneity β_0: Planning Problem Solution and Optimal Subsidy



(a) Optimal Path of $N(t)$



(b) Optimal Path of $\bar{x}(t)$



(c) Optimal Path Subsidy

Notes: Panel (a) shows the optimal levels of adoption, $N(t)$ (optimal) and panel (b) shows the path of $\bar{x}(t)$ (optimal) according to the solution of the planning problem. Panel (c) shows the ratio between the optimal subsidy $\theta_n Z(t)$ and the flow benefit of the average adopter, $Z(t)(\theta_0 + \theta_n N(t))$. Each panel shows paths for different values of β_0.