A Model of Technology Diffusion

Linyi Cao* and Cheng Wang†

October, 2019

Abstract

Many new technologies, instead of being adopted simultaneously by all producers in the same area or industry, display a long and lagged diffusion process, with an S-shaped adoption curve. This even applies to technologies which were later proven to improve productivity significantly. We construct and test a model for explaining this observation. In the model, agents who are heterogeneous in beliefs choose their optimal stopping/adopting time, while they are learning from the output of others. As the population of agents who are experimenting on the new technology grows up, the learning process accelerates. Part of the incentives for them to wait is to free-ride on a larger experimenting group in the future. Our model can explain various technology diffusion data, such as the hybrid corn adoption in the U.S., or the adoption of the 12 industrial innovations.

*Washington University in St. Louis, Department of Economics, l.cao@wustl.edu.
†Fudan University, School of Economics, wangcheng@fudan.edu.cn.

1 Model

Let time be discrete and denoted $t = 0, 1, ...$ and there is one good in the model economy. An industry consists of a continuum of producers, normalized to be of measure one. These producers have been in the industry for long time, and up to time $t = 0$ all of them have used an incumbent technology to produce θ_0 (> 0) units of the good each period, with certainty. But at $t = 0$ a new technology has come to existence, ready to be adopted by all producers in the industry. This new technology is supposed to be more efficient relative to the old but, being new, also brings with it uncertainty about how reliable it is.

The new technology may be reliable as desired, in which case it produces a high output of θ_H ($> \theta_0$) units of the good with certainty. It could also be unreliable, in which case it produces either the output θ_H or an output θ_L (< 0).

The agents do not know, at $t = 0$ when the economy starts, whether the new technology is reliable or not. But if the technology is not reliable, it could be detected. Specifically, let $\nu \in [0, 1]$ be the measure (fraction) of the agents who are using the new technology in a given period – the adoption rate in that period. If the technology is *not* reliable, then with probability $1 - P(\nu)$ the true state of the technology would be detected in the period. Assume $1 - P(\nu)$ is strictly increasing in ν. That is, the more producers are using the technology, the more likely an unreliable technology is detected in a given period.

The values of θ_0, θ_L, θ_H and the function $P(\nu)$ are all known to the producers, while the value of ν evolves endogenously. Assume the value of $P(\nu)$ is close to one for all values of ν – the technology, before being brought into the market, must have been tested many times, in the lab and by real users, each time producing the desired output θ_H. But θ_L is so bad that once the technology is found out to be not reliable – i.e., in case θ_L does occur – then it should be abandoned.

What happens at $t = 0$ and the periods after is for each individual producer to decide whether to adopt the new technology, and at what time. Those who have switched to the new technology could also make a decision to switch back. Switching to the new technology imposes a one time cost z on the producer, and the cost may differ between individual producers. Let the producers be distributed in z with a distribution density $g(z)$, where $z \in [0, \bar{z}]$, with $\bar{z} > 0$, and $g(z) \geq 0$ for all z. This distribution can be degenerate in which case the switching costs are a constant for all producers. Lastly, switching back from the new to the old technology imposes a common one time cost z_* on any producer.

At the beginning of $t = 0$ when the story starts, producers hold individual beliefs about whether or not the new technology is reliable. Let $\pi(\in [0, 1])$ denote an individual producer's belief in a given period of the probability with which the new technology is reliable. At $t = 0$ the producers are distributed over the interval $[0, 1]$, in their initial value of π. Let the density of the initial beliefs be denoted $f^0(\pi)$, $\pi \in [0, 1]$.

For technical reasons which become clear later, we denote the logit transformation of π as

$$\rho \equiv log\Big(\frac{\pi}{1-\pi}\Big), \tag{1}$$

and, to abuse notation a little bit, we let the distribution of the initial ρs be denoted also as $f^0(\rho)$, $\rho \in \mathbb{R}$. In the following, we use the word "belief" to refer to the producer's π or ρ.

As time unfolds and the new technology gets used by more producers, new information comes in, producers update their beliefs about the new technology, using a Bayesian updating rule. That is, suppose an individual producer starts with an initial belief ρ, and suppose the high output θ_H is produced by all users of the new technology in the current period, with ν being the current adoption rate. Then the producer's updated belief, ρ', is given by

$$\rho' = \rho - logP(\nu). \tag{2}$$

But if a low output θ_L is produced in any period, then all producers' new belief ρ' drops to $-\infty$ immediately. This implies that an unreliable new technology will be detected almost for sure in the long run. And once detected, the economy will convert permanently back to the old technology.

At the beginning of a period t, the economy is characterized by two aggregate states. The first is a measure of producers who are still with the old technology, $f_O^t(\rho)$, $\forall\rho \in \mathbb{R}$. Specifically, $f_O^t(\rho)$ is the measure of producers who are still with the old technology and hold the belief ρ at the beginning of period t. The second is a measure of producers who adopted the new technology in a prior period: $f_N^t(\rho)$, $\forall\rho \in \mathbb{R}$. Of course, $f_O^0 = f^0$ and $f_N^0(\rho) = 0$, $\forall\rho \in \mathbb{R}$.

It takes time to switch to the new technology from the old. In order to use the new technology in period t, the producer must make a decision to switch away from the old technology in period $t - 1$. Thus the adoption rate in a period t is determined in period $t - 1$ as

$$\nu_t = \int_{\mathbb{R}} f_N^t(\rho) d\rho. \tag{3}$$

(Note that this setup, though less intuitive, helps to simplify the value function.)

A producer who is still with the old technology then decides whether to switch to the new technology for the next period. A producer are using the new technology could also decide to switch back to the old. But in equilibrium, until the industry observes the first θ_L, all producers who had switched to the new technology will stay with the new technology.

Suppose a low output θ_L is observed in a period T. Then all producers using the new technology will cease production in $T+1$, and switch back to the old technology. This implies a value of V_* from period $T+2$ on, for all producers.

To close the model, we assume

$$u(\theta_0) \gg \lim_{\nu \to 0^+} P(\nu)u(\theta_H) + \big(1 - P(\nu)\big)u(\theta_L), \tag{4}$$

so that an unreliable technology, once detected, should be abandoned.

2 Equilibrium

At the start of a period, given that the new technology has not yet been detected as unreliable, individual producers are divided into two groups: those who are with the old technology, in state O, and those who are already with the new technology, in state N. Let $\nu \in [0,1]$ denote the fraction of the producers that are already with the new technology, that is in state N. That is, ν is the "adoption rate" and we assume that its value is commonly observed by all producers. At the start of the period producers hold beliefs about whether the new technology is reliable or not. Let ρ denote an individual producer's belief that the new technology is reliable. This ρ may differ across individual producers. In a rational expectations equilibrium which we are about to define, agents are assumed to perfectly perceive the initial distribution of ρ across agents and how that distribution evolve over time. Specifically, let the rationally perceived measure of the beliefs of the state O producers be denoted f_O, and the measure of the beliefs of the producers in state N be denoted f_N.

Let the value of an individual producer in state $i(= O$ or $N)$ and with belief ρ at the start of a period be denoted $V_i(\rho; f_O, f_N)$. This agent must make a decision in the current period about whether to switch from his current technology state to the other technology and we let his choice be denoted $x_i(\rho; f_O, f_N)$, where $x_i = 1$ indicates the decision of choosing the new technology and $x_i = 0$ indicates the old.

In a rational expectations equilibrium which we are describing, each agent takes as given his rationally perceived *equilibrium* distribution of the actions taken by all other producers in the current period, denoted $x_i^*(\rho; f_O, f_N)$.

Let us now describe the problem of an agent in state O, in the following Bellman equation:

$$\begin{aligned} V_O(\rho; f_O, f_N) &= \max_{x_O \in \{0,1\}} u(\theta_0) \\ &+ \beta \Bigg\{ (1-x_O)\Big[\big(\pi + (1-\pi)P(\nu)\big)V_O(\rho'; f'_O, f'_N) + \underbrace{(1-\pi)\big(1-P(\nu)\big)}_{\text{subjective prob of detection}} V_*\Big] \\ &+ x_O\Big[-\frac{z}{\beta} + \big(\pi + (1-\pi)P(\nu)\big)V_N(\rho'; f'_O, f'_N) + \underbrace{(1-\pi)\big(1-P(\nu)\big)}_{\text{subjective prob of detection}} V_{**}\Big]\Bigg\} \end{aligned}$$

subject to

$$\pi = \frac{exp(\rho)}{1 + exp(\rho)}, \tag{5}$$

$$\nu = \int_{\mathbb{R}} f_N(\rho) d\rho, \tag{6}$$

$$\rho' = \rho - logP(\nu), \tag{7}$$

and for all $\hat{\rho} \in \mathbb{R}$,

$$f'_O(\hat{\rho}') = \sum_{i=O,N} \big[1 - \tilde{x}_i(\hat{\rho}; f_O, f_N)\big] f_i(\hat{\rho}); \tag{8}$$

$$f'_N(\hat{\rho}') = \sum_{i=O,N} \tilde{x}_i(\hat{\rho}; f_O, f_N) f_i(\hat{\rho}). \tag{9}$$

where $\hat{\rho}' \equiv \hat{\rho} - logP(\nu)$.

For a producer in state N, he takes as given $\tilde{x}_i(\rho; f_O, f_N),\ i = O, N$ and solves:

$$\begin{aligned} V_N(\rho; f_O, f_N) &= \max_{x_N \in \{0,1\}} \big(\pi + (1-\pi)P(\nu)\big)u(\theta_H) + (1-\pi)\big(1-P(\nu)\big)u(\theta_L) \\ &+ \beta\Bigg\{(1-x_N)\Big[-\frac{z_*}{\beta} + \big(\pi + (1-\pi)P(\nu)\big)V_O(\rho'; f'_O, f'_N) + (1-\pi)\big(1-P(\nu)\big)V_*\Big] \\ &+ x_N\Big[\big(\pi + (1-\pi)P(\nu)\big)V_N(\rho'; f'_O, f'_N) + (1-\pi)\big(1-P(\nu)\big)V_{**}\Big]\Bigg\} \end{aligned}$$

subject to (5)-(9)

To make a distinction, we call i the individual agent's "technology state", and $\rho,\ f_O,\ f_N$ the "belief states".

Notice that (7)-(9) describe the evolution of the belief states given that θ_L doesn't occur. The individual subjective probability of detecting a θ_L is given by $(1-\pi)\big(1-P(\nu)\big)$. If θ_L is ever observed, all producers' belief about the reliability of the new technology will converge to $\pi = 0$ or $\rho \to -\infty$ and stay permanently there in an absorbing state. Given this, and given (4), we have two terminal values for the absorbing state:

$$V_* = \lim_{\rho \to -\infty} V_O = \frac{u(\theta_0)}{1-\beta}; \tag{10}$$

$$V_{**} = \lim_{\rho \to -\infty} V_N = -z_* + \beta V_*. \tag{11}$$

Definition 1. *A rational expectations equilibrium of the model consists of a set of value and decision functions for the individual agents,*

$$\big\{V_i(\rho; f_O, f_N),\, x_i^*(\rho; f_O, f_N) : i = O, N\big\},$$

and a law of motion for the economy's aggregate states $\{f_O, f_N\}$, given in equations (8)-(9), such that

(i) Given the evolution of the belief states, $\big\{V_i(\rho; f_O, f_N),\, x_i^*(\rho; f_O, f_N)\big\}$ solves agent i's Bellman equation, $i = O,\, N$.

(ii) The evolution of the aggregate states ν, f_O and f_N, is consistent with the initial ν, f_O and f_N, the individual producer's optimal decisions and the Bayesian rule that the agents use for updating beliefs.

Following from (iii) of the definition, if a low output (black swan event) is ever observed, all belief states degenerate to $\pi = 0$ ($\rho \to -\infty$). Otherwise, ρ evolves according to (7), f_i evolves according to (8) and (9).

3 Characterizing the Equilibrium

Assume $\lim_{\nu \to 0^+} P(\nu) < 1$. That is, an unreliable new technology has a baseline probability of being detected, even if there are a very small measure of producers who have adopted it.

Proposition 1. *Suppose that the new technology is reliable. Suppose the initial distribution of beliefs doesn't have a mass point at $\pi = 0$. Then the equilibrium adoption rate will converge to 1 as time goes to infinity.*

But characterizing the dynamics on the equilibrium path is more challenging. In this section, we prove several properties of the equilibrium policy functions. With those properties, we can write an equivalent maximization problem, in which f_O, f_N are replaced by a cutoff belief $\bar{\rho}$ and the mean of the distribution of beliefs m_ρ.

Proposition 2. *In a rational expectations equilibrium of the model, $x_i(\rho; f_O, f_N)$ has the following cutoff property: For any given aggregate state (f_O, f_N), and $i = O, N$, there exists a cutoff belief $\bar{\rho}_i$ such that*

$$x_i(\rho; f_O, f_N) = \begin{cases} 1 & , \ \rho \geq \bar{\rho}_i \\ 0 & , \ \rho < \bar{\rho}_i \end{cases}.$$

Furthermore, we always have $\bar{\rho}_O \geq \bar{\rho}_N$.

The proof can be found in Appendix 5.4. The cutoff property of the policy functions is rather intuitive. It basically says that, given all other conditions equal, if a producer find it optimal to use the new technology, then all producers who hold higher beliefs - that the new technology is reliable - should also find it optimal to use the new technology. And $\bar{\rho}_O \geq \bar{\rho}_N$ says that, it is always easier to stick with the current technology, due to the switching costs.

Our goal is to find a tractable replacement of f_O, f_N on the equilibrium path. Knowing that the policy functions exhibit cutoff property, an immediate guess would be that the two measures of beliefs f_O, f_N are always separating the population distribution of beliefs f by some cutoff $\bar{\rho}$. The following proposition of cutoff rule verifies this guess.

Proposition 3. *Suppose the current aggregate state f_O, f_N are separating the population distribution of beliefs f by a cutoff $\bar{\rho}$. Following the policy functions, the next period f_O', f_N' must also be separating f' by a cutoff $\bar{\rho}'$.*

The proof of the proposition is in Appendix 5.5.

We start the economy with $f_O^0 = f^0$ and $f_N = 0$, which satisfy the cutoff rule. So we can replace the equilibrium dynamics of f_O, f_N by the equilibrium dynamics of $\bar{\rho}$ and f. Remember that from (2) we know $f^t(\rho)$ is always a shifting from $f^0(\rho)$. For a large set of distributions, this implies a change of mean, and we can replace f by its mean m_ρ. In conclusion, on the equilibrium path, we can use $\bar{\rho}$ and m_ρ as state variables instead.

One more thing to be noted before we go to the new maximization problem. The cutoff property of f_O, f_N could be violated in off-equilibrium situations. Think about the case where producers who hold belief ρ are in state N, producers who hold belief $\hat{\rho}$ are in state O, but $\bar{\rho}_O > \hat{\rho} > \rho > \bar{\rho}_N$. In the next period, this violation situation will continue, and presumably for longer periods. This is why we claim that the new maximization problem we are going to establish, is equivalent to the original one, only in the sense that they generate exactly the same equilibrium dynamics.

We rewrite the value functions and policy functions under three real-valued state variables: the individual belief ρ; The cutoff belief $\bar{\rho}$; And finally, the mean of the population distribution of beliefs m_ρ.

$$V_i(\rho; \bar{\rho}, m_\rho): \ \mathbb{R}^3 \to \mathbb{R};$$
$$x_i(\rho; \bar{\rho}, m_\rho): \ \mathbb{R}^3 \to \{0, 1\};$$

for $i = O, N$.

Because of the cutoff rule, we can also replace the rational perception of other producers' actions $\tilde{x}_i$, by a rational perception of the future cutoff $\tilde{\rho}'(\bar{\rho}, m_\rho)$. It can also be regarded as the aggregate policy function. We specify a functional form for

$$P(\nu) = \alpha^{\nu+c},$$

where $\alpha \in (0, 1)$ and $c > 0$.[1] And write the following maximization problem.

Taking $\tilde{\rho}'(\bar{\rho}, m_\rho)$ as given, a producer with the old technology solves

$$\begin{aligned} V_O(\rho; \bar{\rho}, m_\rho) \quad = \quad & \max_{x_O \in \{0,1\}} u(\theta_0) + \beta \Bigg\{ (1 - x_O) \Big[\big(\pi + (1-\pi)\alpha^{\nu+c}\big) V_O(\rho'; \bar{\rho}', m'_\rho) + (1-\pi)(1-\alpha^{\nu+c}) V_* \Big] \\ & + x_O \Big[-\frac{z}{\beta} + \big(\pi + (1-\pi)\alpha^{\nu+c}\big) V_N(\rho'; \bar{\rho}', m'_\rho) + (1-\pi)(1-\alpha^{\nu+c}) V_{**} \Big] \Bigg\} \end{aligned}$$

s.t.

$$\pi = \frac{exp(\rho)}{1 + exp(\rho)}; \tag{12}$$

$$\nu = \int_{\rho \geq \bar{\rho}} f(\rho | m_\rho) d\rho; \tag{13}$$

[1] See a discussion of why we pick this particular functional form in Appendix 1.

$$\rho' = \rho + (-log\alpha)(\nu + c); \tag{14}$$

$$\bar{\rho}' = \tilde{\rho}'(\bar{\rho}, m_\rho); \tag{15}$$

$$m'_\rho = m_\rho + (-log\alpha)(\nu + c); \tag{16}$$

And taking as given $\tilde{\rho}'(\bar{\rho}, m_\rho)$, a producer who is currently with the new technology solves

$$\begin{aligned} V_N(\rho;\bar{\rho},m_\rho) \quad = \quad & \max_{x_N\in\{0,1\}} \Big(\pi + (1-\pi)\alpha^{\nu+c}\Big)u(\theta_H) + (1-\pi)(1-\alpha^{\nu+c})u(\theta_L) \\ & \beta\Bigg\{(1-x_N)\Bigg[-\frac{z^*}{\beta} + \Big(\pi + (1-\pi)\alpha^{\nu+c}\Big)V_O(\rho';\bar{\rho}',m'_\rho) + (1-\pi)(1-\alpha^{\nu+c})V_*\Bigg] \\ & +x_N\Bigg[\Big(\pi + (1-\pi)\alpha^{\nu+c}\Big)V_N(\rho';\bar{\rho}',m'_\rho) + (1-\pi)(1-\alpha^{\nu+c})V_{**}\Bigg]\Bigg\} \end{aligned}$$

s.t.

$$(13) - (17)$$

$$V_* = \frac{u(\theta_0)}{1-\beta}; \tag{17}$$

$$V_{**} = -z_* + \beta V_*. \tag{18}$$

I claim that the maximization problem above is equivalent to the original problem, in the sense that they generate exactly the same equilibrium dynamics. At last, we shall describe the consistency condition for the rational expectation equilibrium. Define

$$\bar{\rho}_O(\bar{\rho}, m_\rho) = min\{\rho \in R \mid x_O(\rho;\bar{\rho},m_\rho) = 1\},$$

$$\bar{\rho}_N(\bar{\rho}, m_\rho) = min\{\rho \in R \mid x_N(\rho;\bar{\rho},m_\rho) = 1\},$$

then we can write the consistency condition as

$$\tilde{\rho}'(\bar{\rho}, m_\rho) = max\Big\{min\big\{\bar{\rho}, \bar{\rho}_O(\bar{\rho}, m_\rho)\big\}, \bar{\rho}_N(\bar{\rho}, m_\rho)\Big\} + (-log\alpha)(\nu + c). \tag{19}$$

where ν is derived from (13).

4 Numerical Simulations

We have solved the equilibrium numerically, you can find a detailed discussion of the algorithm in Appendix 5. We choose the following specification for the initial distribution of beliefs

$$f^0(\rho) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(- \frac{(\rho - \mu)^2}{2\sigma^2} \right),$$

that is, a normal distribution with mean μ and variance σ^2. In the benchmark parameterization, we use the following values:

Table 4.1: Benchmark Parameterization

Utility Levels	$u(\theta_0) = 0$, $u(\theta_H) = 1$, $u(\theta_L) = -15$
Probability	$\alpha = 0.7$, $c = 0.2$
Initial Distribution	$\mu = -2.2$, $\sigma = 1$
Grid Points	$N = 101$, $\pi_{min} = 10^{-7}$, $\pi_{max} = 1 - \pi_{min}$
Other	$\beta = 0.9$, $z = z^* = 1$

We choose the normal distribution in the ρ-space because it has a very good property: the initial mean of the distribution does not affect the equilibrium, only the initial variance does. The initial mean will come to play in the simulation part though. We have also tried with uniform distribution in the π-space, but find no qualitative difference.

Using the solved aggregate policy function $\tilde{\rho}'(\bar{\rho}, m_\rho)$, we have simulated the equilibrium path of adoption rate, starting from different initial means. The basic procedure for simulation is: we start from some given initial state $(\bar{\rho}^0, m_\rho^0)$, calculate the next period state from

$$\bar{\rho}^1 = \tilde{\rho}'(\bar{\rho}^0, m_\rho^0),$$

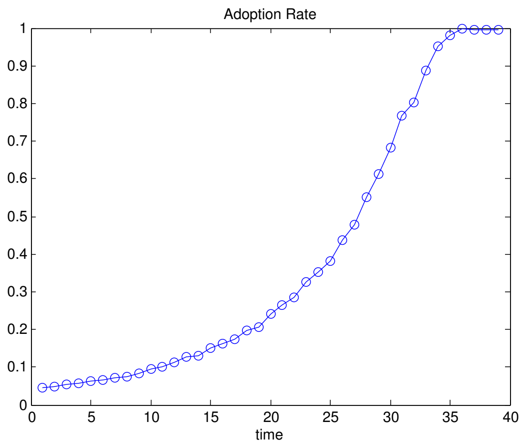
$$m_\rho^1 = m_\rho^0 + (-log\alpha)(\nu^0 + c),$$

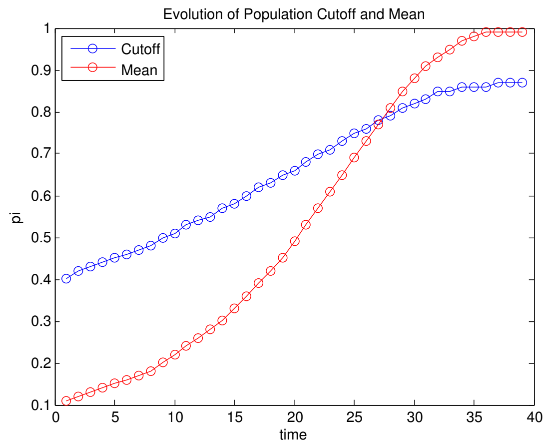
and iterate for T many times. Notice that once $(\bar{\rho}^t, m_\rho^t)$ are given, the adoption rate ν^t can be calculated from

$$\nu^t = 1 - \Phi\left(\frac{\bar{\rho}^t - m_\rho^t}{\sigma}\right),$$

where Φ is the CDF of a standard normal distribution.

Graph 4.2: Adoption Rate and Evolution of Aggregate State





The adoption rate curve is clearly S-shaped. And we have this pattern for a large range of different initial means.

5 Appendix

5.1 The Extensive Form Problem

Though we are not going to solve it directly, let's write the extensive form maximization problem of a producer, and the corresponding rational expectations equilibrium(REE).

Starting with some initial belief π_0, technology $i = O, N$ and aggregate state $\left(f_O^0, f_N^0\right)$, the producer makes a history-contingent plan of all future choices of technology $\left\{x_t\right\}_{t=0}^{\infty}$. Let $x_t = 1$ indicates the choice of the new technology, and remember that the technology has a time-to-built, so x_t will become effective in $t+1$.

The history, which in general is a record of all past events the producer could observe, contains two components. The first is the history of all past outputs $\left\{\theta_s\right\}_{s=0}^{t-1}$. After knowing $\left\{\theta_s\right\}_{s=0}^{t-1}$, we can pin down the producer's past beliefs $\left\{\pi_s\right\}_{s=0}^{t}$ too. The second is the history of all producers' past choices of technology. Since we are interested in a symmetric equilibrium in which producers in the same state always make the same choice, this history is summarized by $\left\{f_O^s, f_N^s\right\}_{s=0}^{t}$. Define

$$h^t \equiv \left\{\left\{\theta_s\right\}_{s=0}^{t-1} ;\ \left\{f_O^s, f_N^s\right\}_{s=0}^{t}\right\} \tag{20}$$

as a history, and H^t as the set of all possible histories, up to the beginning of period t. We can write the history-contingent plan of choices of technology as $\left\{x_t(h^t)\right\}_{t=0}^{\infty}$, the beliefs as $\left\{\pi_t(h^t)\right\}_{t=0}^{\infty}$, where

$$x_t(h^t) :\ H^t \to \{0,1\};$$
$$\pi_t(h^t) :\ H^t \to [0,1].$$

In an REE, the producer should be able to form a rational expectation of future dynamics of belief measures, which is history-contingent too. Define $\left\{\tilde{f}_O^{t+1}(\rho|h^t), \tilde{f}_N^{t+1}(\rho|h^t)\right\}_{t=0}^{\infty}$ as the producer's rational expectation of future dynamics of belief measures, where

$$\tilde{f}_i^{t+1}(\rho|h^t) :\ H^t \to \mathcal{F},\ i = O, N.$$

Now we can write the maximization problem, for an individual producer who holds initial belief π_0 or equivalently ρ_0, and in initial aggregate state (f_O^0, f_N^0). For notation consistency, we define $x_{-1} = 0$ $(= 1)$ if the producer starts at state O (N), and $h^0 = \emptyset$.

The producer takes as given a rational expectation $\left\{\tilde{f}_O^{t+1}(\rho|h^t), \tilde{f}_N^{t+1}(\rho|h^t)\right\}_{t=0}^{\infty}$, and solves

$$\max_{\{x_t(h^t)\}_{t=0}^{\infty}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \mathbb{E}_t \Bigg\{ -max\big\{x_t(h^t) - x_{t-1}, 0\big\} z - max\big\{x_{t-1} - x_t(h^t), 0\big\} z^*$$
$$+\beta\Bigg[\big(1 - x_t(h^t)\big)u(\theta_0) + x_t(h^t)\Bigg[\Big(\pi_{t+1}(h^{t+1}) + \big(1 - \pi_{t+1}(h^{t+1})\big)P\big(\tilde{\nu}_{t+1}(h^t)\big)\Big)u(\theta_H)$$
$$+\Big(1 - \pi_{t+1}(h^{t+1})\Big)\Big(1 - P\big(\tilde{\nu}_{t+1}(h^t)\big)\Big)u(\theta_L)\Bigg]\Bigg] \;\Bigg|\; h^t \Bigg\}$$

s.t. $\forall\ t$, conditional on $h^t \in H^t$

$$\tilde{\nu}_{t+1}(h^t) = \int_{\mathbb{R}} \tilde{f}_N^{t+1}(\rho|h^t)d\rho$$

$$\nu_t = \int_{\mathbb{R}} f_N^t(\rho)d\rho$$

$$\rho_{t+1}(h^{t+1}) = \begin{cases} \rho_t - logP(\nu_t) & ,\ if\ \theta_t = \theta_H \\ -\infty & ,\ if\ \theta_t = \theta_L \end{cases} \tag{21}$$

$$\pi_{t+1}(h^{t+1}) = \frac{exp\big(\rho_{t+1}(h^{t+1})\big)}{exp\big(\rho_{t+1}(h^{t+1})\big) + 1}$$

The first constraint says that, the producer forms a rational expectation of the next period adoption rate, which directly affects the current choice of x_t. The second constraint says that given h^t, which contains f_N^t, the producer knows for sure the current adoption rate, which affects belief updating. The third constraint is Bayesian updating, if a high output is observed, the producer's belief ρ moves up by $-logP(\nu_t)$. If a low output is observed, it drops to $-\infty$, and stays there. The last constraint is to transform belief back to a probability.

In equilibrium, the evolution of aggregate state must be consistent with individual's optimal choices, that is, conditional on h^t

$$f_O^{t+1}(\rho_{t+1}|h^t) = \sum_{i=O,N} \big(1 - x_t(h^t)\big) f_i^t(\rho_t); \tag{22}$$
$$f_N^{t+1}(\rho_{t+1}|h^t) = \sum_{i=O,N} x_t(h^t) f_i^t(\rho_t); \tag{23}$$

where each ρ_{t+1} corresponds to a ρ_t based on (21). And at last, for the equilibrium to be a REE, we require that for $\forall\ t$, $\forall\ h^t \in H^t$

$$\tilde{f}_i^{t+1}(\rho|h^t) = f_i^{t+1}(\rho|h^t),\ \ \rho \in \mathbb{R};\ \ i = O, N.$$

5.2 The $P(\nu)$ Function

If we think of the experiment as a hypothesis test, $1 - P(\nu)$ would be the power function, i.e. the probability of rejecting the null (the new technology is reliable) when the alternative (the new technology is unreliable) is true. We want it to satisfy

$$-P'(\nu) > 0, \ \lim_{\nu \to \infty} 1 - P(\nu) = 1.$$

The power of the test should be increasing with its size, and approaches a detection for sure in the limit.

An extra property we want this power function to have is that, if we run two experiments, one is with two stages of size ν_1 and ν_2, the other is with only one stage of size $\nu_1 + \nu_2$. Then the final belief we get should be the same. In other words, the size of the experiment should be accumulative.

One can prove that $P(\nu) = \alpha^{\nu}$ is the only continuous function satisfying the property. To see this, in the first experiment, we update belief in the following manner,

$$\begin{aligned}
\rho' &= \rho - logP(\nu_1), \\
\rho'' &= \rho' - logP(\nu_2) \\
&= \rho - logP(\nu_1) - logP(\nu_2),
\end{aligned}$$

while in the second experiment, we have

$$\hat{\rho}'' = \rho - logP(\nu_1 + \nu_2).$$

To such that $\rho'' = \hat{\rho}''$, we need

$$P(\nu_1 + \nu_2) = P(\nu_1)P(\nu_2), \ \forall \ \nu_1, \nu_2.$$

This uniquely pins down the functional form for $P(\nu) = \alpha^{\nu}$, as long as we assume that $P(\nu)$ is continuous. The detection probability is $1 - \alpha^{\nu}$ then.

Interestingly, this form is an analogue of the detection probability in the canonical Bayesian learning. Suppose we have the following signal structure, if the new technology is reliable (unreliable), it produces the bad outcome with probability 0 $(1 - a)$. Now we run $n \in \mathbb{Z}$ many i.i.d. experiments, one bad outcome is enough for us to conclude that the new technology is unreliable, so the detection probability is $1 - a^n$.

However, there are technical difficulties going directly from finitely many producers, to a continuum of them. Denote N as the total number of producers in the finite case, our ν resembles $\underset{n,N\to\infty}{lim} n/N$, and we can rewrite $1-a^n$ as $1-(a^N)^{n/N}$. What we want is $\underset{N\to\infty}{lim} a^N = \alpha$, but there is no fixed a making it happen. We need a to go up to 1 as N goes to infinity. In other words, we need the informativeness of the individual signal, which is determined by the difference between 0 and $1-a$ in our case, to diminish when the number of signals goes up. In the literature, it is common to directly assume $P(\nu)$ with a continuum of producers, rather than to derive it as the limit of some finite cases.

At last, for technical reasons involving the definition of equilibrium, we can't allow $\underset{\nu\to 0^+}{lim} P(\nu) = 1$. Thus we choose $P(\nu) = \alpha^{\nu+c}$, where $c > 0$ is a constant.[2] The functional form doesn't follow the accumulative property exactly, but it is the closest we can get.

5.3 An Alternative Signal Distribution

Other signal forms could also work. For example, suppose the per capita output

$$\bar{Y} \sim N\big(\theta_H, \frac{\sigma^2}{\nu+c}\big)$$

when the new technology is reliable. And

$$\bar{Y} \sim N\big(\theta_L, \frac{\sigma^2}{\nu+c}\big)$$

when the new technology is unreliable.

We would have

$$\rho' = \rho + \frac{(\nu+c)\left(\theta_L^2 - \theta_H^2\right)}{2\sigma^2} + \frac{(\nu+c)\left(\theta_H - \theta_L\right)}{\sigma^2}\bar{Y}.$$

It is still nice and clean, but in this case the belief could move up and down repeatedly, since there is no absorbing state. We leave it for future research.

[2]It is equivalent to $P(\nu) = c\alpha^\nu$, where $c \in (0, 1)$.

5.4 Proof of Proposition 2

The proof consists of two steps. First we are going to prove that $V_N(\rho; f_O, f_N) - V_O(\rho; f_O, f_N)$ is non-decreasing in ρ. Then we will show that $\bar{\rho}_i,\ i = O, N$ exist, and $\bar{\rho}_O \geq \bar{\rho}_N$.

Step1: For any given (f_O, f_N), the difference term $\Delta V(\rho; f_O, f_N) \equiv V_N(\rho; f_O, f_N) - V_O(\rho; f_O, f_N)$ is non-decreasing in $\rho \in \mathbb{R}$.

To start with, notice that both $V_i(\rho; f_O, f_N),\ i = O, N$ are continuous in $\rho \in \mathbb{R}$, and differentiable in $\rho \in \mathbb{R}$ except for some measure 0 set of points.

In a state $i = O$ producer's maximization problem, the producer essentially compares

$$EV_O \equiv \big(\pi + (1-\pi)P(\nu)\big)V_O(\rho'; f'_O, f'_N) + (1-\pi)\big(1 - P(\nu)\big)V_*;$$
$$EV_N \equiv -\frac{z}{\beta} + \big(\pi + (1-\pi)P(\nu)\big)V_N(\rho'; f'_O, f'_N) + (1-\pi)\big(1 - P(\nu)\big)V_{**}.$$

Since we know that

$$\underset{\rho\to\infty}{lim} V_O(\rho; f_O, f_N) = u(\theta_0) - z + \frac{\beta u(\theta_H)}{1-\beta};$$
$$\underset{\rho\to\infty}{lim} V_N(\rho; f_O, f_N) = \frac{u(\theta_H)}{1-\beta};$$

and

$$u(\theta_H) - u(\theta_0) + z - \frac{z}{\beta} > 0.$$

We can conclude that $EV_N > EV_O$, as $\rho \to \infty$. In other words, there exists a $\hat{\rho}_O$ such that for $\forall\ \rho \geq \hat{\rho}_O$, $x_O(\rho; f_O, f_N) = 1$. Similarly, we can argue this for $x_N(\rho; f_O, f_N)$ and a $\hat{\rho}_N$.

Define $\hat{\rho} = max\big\{\hat{\rho}_O, \hat{\rho}_N\big\}$, then for $\forall\ \rho \geq \hat{\rho}$, we have $x_O(\rho; f_O, f_N) = x_N(\rho; f_O, f_N) = 1$ and

$$\Delta V(\rho; f_O, f_N) = \big(\pi + (1-\pi)P(\nu)\big)u(\theta_H) + (1-\pi)\big(1 - P(\nu)\big)u(\theta_L) - u(\theta_0) + z,$$

which is non-decreasing in π (and ρ). Thus $\Delta V(\rho; f_O, f_N)$ is non-decreasing in ρ on $[\hat{\rho}, \infty)$.

As for $\rho \in (-\infty, \hat{\rho})$, because of the differentiability, we can discuss the slope of $\Delta V(\rho; f_O, f_N)$ at any point, except for a measure 0 set of points. For any particular ρ, there are three cases to consider.

Case1: $x_O(\rho; f_O, f_N) = x_N(\rho; f_O, f_N) = 0$. We have

$$\Delta V(\rho; f_O, f_N) = \big(\pi + (1-\pi)P(\nu)\big)u(\theta_H) + (1-\pi)\big(1 - P(\nu)\big)u(\theta_L) - u(\theta_0) - z_*.$$

Obviously it is non-decreasing at such ρ.

Case2: $x_O(\rho; f_O, f_N) = x_N(\rho; f_O, f_N) = 1$. Same as those $\rho \in [\hat{\rho}, \infty)$.

Case3: $x_O(\rho; f_O, f_N) = 0,\ x_N(\rho; f_O, f_N) = 1$. We have

$$\begin{aligned} \Delta V(\rho; f_O, f_N) &= \big(\pi + (1-\pi)P(\nu)\big)u(\theta_H) + (1-\pi)\big(1 - P(\nu)\big)u(\theta_L) - u(\theta_0) \\ &+ \big(\pi + (1-\pi)P(\nu)\big)\Delta V(\rho'; f'_O, f'_N) + (1-\pi)\big(1 - P(\nu)\big)(V_{**} - V_*), \end{aligned}$$

where $\rho' = \rho - logP(\nu)$. As long as we can prove that $\Delta V(\rho'; f'_O, f'_N)$ is non-decreasing at ρ', $\Delta V(\rho; f_O, f_N)$ would be non-decreasing at ρ. Also notice that ρ' is greater than ρ by $-logP(\nu) > 0$. We can apply the same logic to $\Delta V(\rho'; f'_O, f'_N)$ to get a ρ'', and iterate this process until we reach some $\rho^n \in [\hat{\rho}, \infty)$. Then, use backward induction, we can show that $\Delta V(\rho; f_O, f_N)$ is non-decreasing at this ρ.

Note that $x_O(\rho; f_O, f_N) = 1,\ x_N(\rho; f_O, f_N) = 0$ is trivially impossible. So by its continuity, we can conclude that $\Delta V(\rho; f_O, f_N)$ is non-decreasing in $\rho \in \mathbb{R}$.

Step2: For any given (f_O, f_N), the cutoff beliefs $\bar{\rho}_i,\ i = O, N$ exist, and $\bar{\rho}_O \geq \bar{\rho}_N$.

Let's prove it for the state $i = O$ producers first. To show that $\bar{\rho}_O$ exists, it suffices to show that suppose $x_O(\rho_1; f_O, f_N) = 1$ and $\rho_2 > \rho_1$, then $x_O(\rho_2; f_O, f_N) = 1$ too.

$x_O(\rho_1; f_O, f_N) = 1$ implies that

$$\Delta V(\rho'_1; f'_O, f'_N) > \frac{(1-\pi_1)\big(1 - P(\nu)\big)\big(u(\theta_0) + z_*\big) + z/\beta}{\pi_1 + (1-\pi_1)P(\nu)}.$$

Now with a larger ρ_2, the LHS changes to $\Delta V(\rho'_2; f'_O, f'_N)$ where $\rho'_2 > \rho'_1$. Since we have proved that $\Delta V(\rho; f_O, f_N)$ is non-decreasing in $\rho \in \mathbb{R}$, we have $\Delta V(\rho'_2; f'_O, f'_N) \geq \Delta V(\rho'_1; f'_O, f'_N)$.

The RHS changes to

$$\frac{(1-\pi_2)\big(1-P(\nu)\big)\big(u(\theta_0)+z_*\big)+z/\beta}{\pi_2+(1-\pi_2)P(\nu)},$$

where $\pi_2 > \pi_1$, and obviously it is smaller.

Thus we know

$$\Delta V(\rho_2'; f_O', f_N') > \frac{(1-\pi_2)\big(1-P(\nu)\big)\big(u(\theta_0)+z_*\big)+z/\beta}{\pi_2+(1-\pi_1)P(\nu)}$$

must hold, and $x_O(\rho_2; f_O, f_N) = 1$. That concludes our proof for the existence of $\bar{\rho}_O$.

Proof for $\bar{\rho}_N$ follows the same logic, so we omit it here. What's left is to show that $\bar{\rho}_O \geq \bar{\rho}_N$. The proof is straightforward, remember that a state $i = O$ producer compares EV_O and EV_N. If we define the corresponding terms for a state $i = N$ producer, $\hat{EV}_O$ and $\hat{EV}_N$, we can easily check that

$$EV_O > \hat{EV}_O \ ; \ EV_N < \hat{EV}_N.$$

This implies that if a state $i = O$ producer finds it optimal to adopt the new technology, then a state $i = N$ producer, who is otherwise holding the same belief and in a same aggregate state, must also find it optimal to keep using the new technology. This immediately gives us $\bar{\rho}_O \geq \bar{\rho}_N$. QED

5.5 Proof of Proposition 3

By Theorem 1, we know the existence of $\bar{\rho}_O$ and $\bar{\rho}_N$. Based on the values of them and the current cutoff $\bar{\rho}$, there are three cases possible.

Case1: $\bar{\rho} > \bar{\rho}_O \geq \bar{\rho}_N$. Remember that producers with belief $\rho \in (-\infty, \bar{\rho})$ are currently in state $i = O$. Since $\bar{\rho} > \bar{\rho}_O$, those state $i = O$ producers with belief $\rho \in [\bar{\rho}_O, \bar{\rho})$ will switch to the new technology. And the next period cutoff is

$$\bar{\rho}' = \bar{\rho}_O - logP(\nu),$$

the adoption rate increases.

Case2: $\bar{\rho}_O \geq \bar{\rho} \geq \bar{\rho}_N$. In this case, neither state $i = O$ or state $i = N$ producer would want to change their current technology, and the next period cutoff is

$$\bar{\rho}' = \bar{\rho} - logP(\nu),$$

the adoption rate doesn't change.

Case3: $\bar{\rho}_O \geq \bar{\rho}_N > \bar{\rho}$. In this case, state $i = N$ producers with belief $\rho \in [\bar{\rho}, \bar{\rho}_N)$ will switch back to the old technology. And the next period cutoff is

$$\bar{\rho}' = \bar{\rho}_N - logP(\nu),$$

the adoption rate decreases.

We are confident that the Case3 won't happen on equilibrium path. But nevertheless, Proposition 2 holds, on equilibrium or off equilibrium path. QED

5.6 Algorithm for Numerical Solution

The model is solved numerically using a hybrid of value function iteration and policy function iteration technics. The algorithm contains two loops, the inner loop conducts value function iteration. While fixing the current guess of the aggregate policy function $\tilde{\rho}'(\bar{\rho}, m_\rho)$, the inner loop iterates on the two value functions $V_O(\rho; \bar{\rho}, m_\rho)$, $V_N(\rho; \bar{\rho}, m_\rho)$, until they both converge.

The outer loop conducts policy function iteration. It uses the converged value functions to generate a new aggregate policy function. Remember that those value functions are derived from the inner loop while fixing the current guess of the aggregate policy function. If the new aggregate policy function and the current guess of aggregate policy function are close enough, we are done.

Here are some more details. First, the state space is discretized. Originally, these three state variables ρ, $\bar{\rho}$, $m_\rho \in \mathbb{R}$, while their counterparts π, $\bar{\pi}$, $m_\pi \in [0, 1]$. We take an even grid in the π-space, that is, N grid points between some $[\pi_{min}, \pi_{max}]$. Then we use the logit transformation to get a corresponding uneven grid in the ρ-space. Thus in the loop, $V_O(\rho; \bar{\rho}, m_\rho)$ and $V_N(\rho; \bar{\rho}, m_\rho)$ are stored as $N \times N \times N$ matrices, while $\tilde{\rho}'(\bar{\rho}, m_\rho)$ is stored as a $N \times N$ matrix.

Second, to avoid coarseness from discretization, we allow the $\tilde{\rho}'(\bar{\rho}, m_\rho)$ matrix to take values in $\mathbb{R}$ (or in $[0, 1]$, depending on its format). And when we need to map the value back on the grid, we use linear interpolation. For example, suppose $\tilde{\rho}' \in \left[\rho_{\text{grid}}(n), \rho_{\text{grid}}(n+1)\right]$, we have

$$V_i(n_\rho; \tilde{\rho}', n_{m_\rho}) = \frac{\rho_{\text{grid}}(n+1) - \tilde{\rho}'}{\rho_{\text{grid}}(n+1) - \rho_{\text{grid}}(n)} V_i(n_\rho; n, n_{m_\rho}) + \frac{\tilde{\rho}' - \rho_{\text{grid}}(n)}{\rho_{\text{grid}}(n+1) - \rho_{\text{grid}}(n)} V_i(n_\rho; n+1, n_{m_\rho}).$$

or use π_{grid} instead, when its format is a $\tilde{\pi}'$.

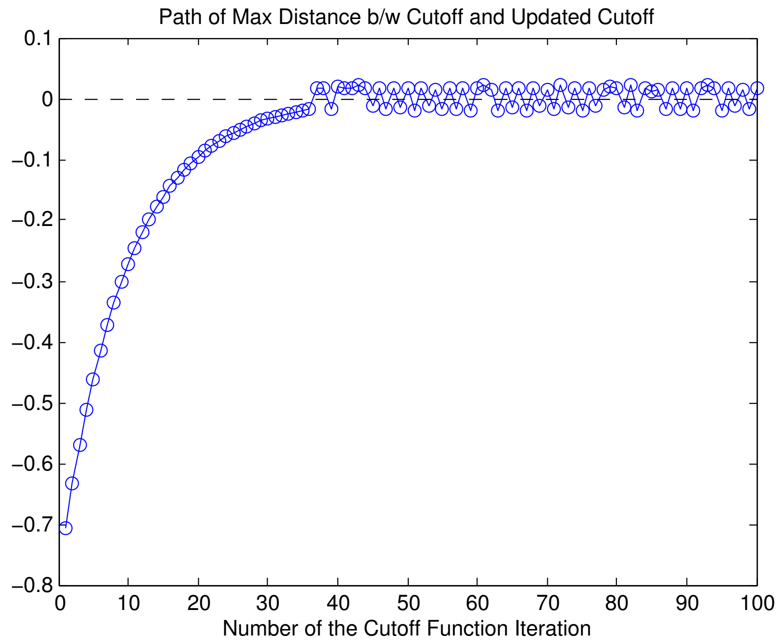
As in any typical iteration algorithm, we calculate the differences between the current guess of function and the newly derived function. When the distance, defined as the maximum absolute value of the differences, is small enough, we have ourselves a fixed point and that gives us the numerical solution. Since they are all real-valued functions, the convergence criteria for the value functions and the aggregate policy function are certain thresholds. The threshold for the value function iteration is set to be 10^{-7}. And the convergence in the value function iteration is universal, cause it is a contraction mapping.

The convergence in the policy function iteration is complicated though, we don't get universal convergence. At most of the grid points, convergence is achieved within a few rounds of iteration. But at some of the grid points, we find the aggregate policy function oscillate. For example, with $N = 101$ and 100 rounds of iteration, we end up with large differences[3] at 36 out of 101^2 grid points. And along the way, the set of grid points at which the differences are large becomes stable. Here is a graph of the maximum difference[4] during the 100 rounds of iteration.

[3]A difference is regarded as large if its absolute value is no less than 10^{-3}.

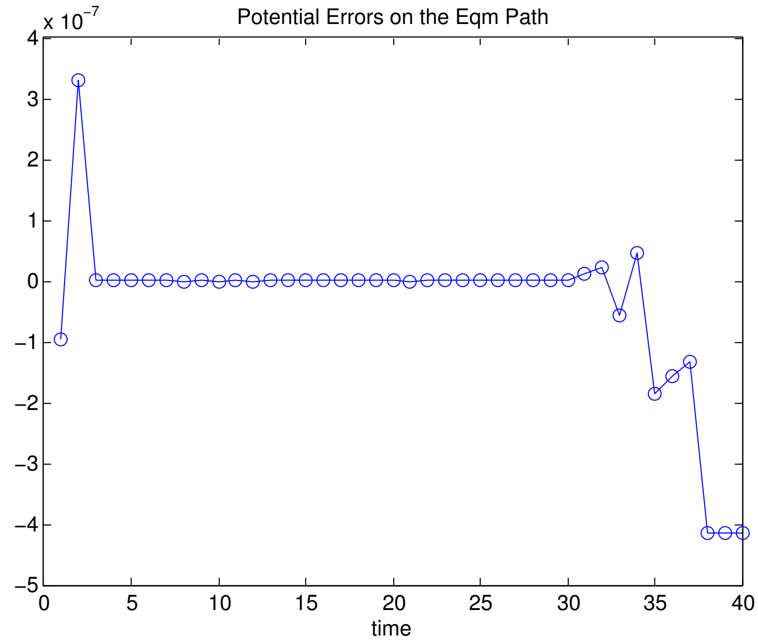
[4]In terms of the absolute value of course. In this numerical exercise, they are all achieved at the same grid point, except for the first few rounds of iteration.

Graph A1: Maximum Difference during Iterations



Those grid points at which the aggregate policy function oscillates, though being a very small portion of the total grid points, could jeopardize our simulation if they were on the equilibrium path. So the next task is to show that they are never visited, unless we start the economy there. We run several simulations, by varying the mean of the initial distribution from $\rho = -2.2$ ($\pi = 0.1$) to $\rho = 0$ ($\pi = 0.5$). In each simulation, we record the differences along the equilibrium path of state. It turns out that those differences never exceed a scale of 10^{-6}. Here is a graph of them when the initial mean is $\rho = -2.2$ ($\pi = 0.1$).

Graph A2: Errors along the Equilibrium Path



We conclude that the simulation results are valid, though we don't get universal convergence with the aggregate policy function. We could also reach such conclusion by switching to a more lenient measure of distance, for example, a weighted average $\frac{1}{N}\sqrt{\sum_{i,j} \text{difference}_{i,j}^2}$.