This paper explores the idea that the information externalities inherent in institutions may contribute to their persistence. Institutions that form to reduce moral hazard often eliminate discretion and pool the actions of heterogeneous agents so that agents’ types cannot be determined by their actions. While in the short run such mechanisms may be optimal, in the long run inefficient institutions may persist because information about changes in the underlying environment is lost. I model and experimentally analyze a market with high quality and low quality products that are indistinguishable without a costly certification process. Sellers in the market make endogenous production decisions and are heterogeneous in their levels of moral hazard leading to two possible equilibria—non-certifying and certifying—that vary in both efficiency and information about the underlying environment. I find that the certifying equilibrium, which does not carry information about changes in the distribution of sellers, does not adjust when the underlying environment changes. This leads to the persistence of an inefficient certification institution that makes all market participants weakly worse off.

Keywords: Persistence of Institutions, Information, Externalities, Inefficient Market Structures, Experimental Economics.

JEL Classification Codes: D02, C92, D40, D62, D82
1 Introduction

Many real world markets and organizations operate in environments with both heterogeneous types and moral hazard. In the markets for public debt, for instance, firms have heterogeneity in project choice and moral hazard in taking risk and truthfully revealing their financial records. In agriculture markets, producers have heterogeneous costs for using quality inputs and moral hazard in supplying safe food.

Institutions that emerge to reduce moral hazard often do so through certification and monitoring, which takes discretion away from the individual agents in the market. The Sarbanes-Oxley Act, for instance, requires standardized auditing across firms and risk management which constrains some firms from taking undisclosed risks. The FDA sets certification standards for agriculture products and offers voluntary grading and certification programs for a variety of US agriculture goods. These programs may endogenously induce sellers to specific joint production and certification decisions.

By eliminating discretion, certification institutions agglomerate the actions of heterogeneous agents so that observing an action may reveal less information about the type of agent performing it. This pooling of types can eliminate information about the true state of the world and limit the ability for individuals to learn and make comparisons between alternative institutions. If the environment changes so that the certification institution is no longer needed, participants in the market cannot observe these changes and eliminate the institution.

The persistence of inefficient institutions may have long-term consequences on the efficiency of many markets and organizations, especially in environments where the underlying population is stochastic and the optimal institution varies over time. The persistence of monitoring may lead to a continual expansion of red tape. The persistence of certification may lead to additional verification costs and reduced competition. The persistence of regulation may lead to increased enforcement costs and a decrease in intrinsic motivation.

This paper explores the idea that persistence of agglomerating institutions may be due to inadequate information—an information externality inherent to the mechanism itself. Intuitively, if an institution or market structure suppresses information about the underlying environment, the institution is likely to persist because information necessary to evaluate relative efficiency and coordinate to a new equilibrium is missing.

I develop a theoretical market with ex ante indistinguishable high quality and low quality products where a costly certification technology can be adopted by sellers that guarantees
quality. Heterogeneity in production costs divide sellers into three categories: good sellers who produce high quality units, conditional sellers who have moral hazard and produce high quality units in the certified market and low quality units in the uncertified market, and bad sellers who always produce low quality units.

I show that for some initial distributions of seller types, two competitive equilibria may emerge—non-certifying and certifying—which vary significantly both in terms of efficiency and in the information they generate about the underlying environment.

- In the non-certifying equilibrium, no seller chooses to certify their product and the prevailing market price carries information about the proportion of good sellers. A decrease in the number of good sellers leads to an observable decrease in the non-certifying price. This decline in price leads to a potential arbitrage opportunity for good and conditional sellers by adopting certification and provides a natural channel by which a market may endogenously adopt certification.

- In the certifying equilibrium, the certification process is adopted by both good and conditional sellers so that their actions no longer reveal their types. Changes in the proportion of sellers between these two groups are not observable by the market and thus there is no information revealed when market conditions change. This information externality may lead to inefficient persistence of the certification institution since the true state of the world is not transmitted through private nor market information.

I next use laboratory experiments to explore the informational properties of the certifying and non-certifying equilibrium by comparing their response to changes in the environment. I design two environments—Safe and Hazardous—which vary in the composition of sellers in the market. In the safe environment, the proportion of good sellers in the market is large, thusfavoring the formation of the non-certifying equilibrium. In the hazardous environment, good sellers are replaced with conditional sellers, leading to significant amounts of moral hazard and the elimination of the non-certifying equilibrium. I begin half the sessions in the safe environment and switch to the hazardous environment midway through. In the other half, I reverse the order, starting in the hazardous environment and ending in the safe environment.

Consistent with the theory, I find that individuals who begin in the safe environment establish a non-certifying equilibrium and then adapt to the certifying equilibrium in response to a change in the underlying environment. Subjects who begin in the hazardous environment form the certifying equilibrium and remain in this equilibrium when the environment is
changed to safe. This persistence leads to a loss in efficiency relative to a market where the non-certifying equilibrium initially formed.

Compared to other empirical tools, experimentation has many advantages for studying informational persistence. First, supply and demand are exogenously set and do not have to be estimated in conjunction with the parameters of interest. This allows for more explicit hypothesis tests, exogenous changes to the environment, and tight control over the number of market equilibria. Second, natural data sets where certification institutions are eliminated are rare. While this may be an indication that information is important, it may also be a result of strategic actions taken by agents with a vested interest in the current institution. In an experimental setting, I can eliminate this form of agency from the analysis by ensuring that all individuals can be made weakly better off with the elimination of the certifying equilibrium. Experimentation also guarantees exogenous variation in seller costs and ensures the existence of a non-certifying equilibrium through careful assignment of seller costs. Finally, primitives that are typically unobservable such as beliefs and risk aversion can be elicited in conjunction with the main experiment and used to study what forces shape the dynamics of the market.

This paper contributes to a growing literature on the persistence of institutions. Whereas the political economy literature, developed by North (1990) and Acemoglu & Robinson (2006) has centered on the role that agency plays in persistence, I develop an informational channel of persistence where the informational properties of institutions themselves endogenously affect long-run outcomes. I show that inefficient institutions may persist even in an environment where all agents can be made better off. Given the prevalence of mechanisms that mitigate moral hazard such as red tape, regulation, certification, and monitoring, this form of informational externality may be of great importance to the function and efficiency of many markets and organizations.

The paper is organized as follows. Section 2 relates the current paper to the literature. Section 3 builds the theoretical model and characterizes its competitive equilibria in terms of efficiency and information. Section 4 develops the experimental design. Section 5 reports the main experimental results and is divided into three parts. Section 5.1 looks at initial convergence of the experimental market in the safe and hazardous environments. Section 5.2 demonstrates the difference in adaptation between the non-certifying and certifying equilibrium. Section 5.3 looks at the welfare consequences of persistence in the certifying equilibrium and section 6 concludes.
2 Related Literature

Although the theory in this paper is, to the best of my knowledge, new, it is similar to the history dependence, herding, and conventions literature developed by other authors. History dependent models such as Argenziano & Gilboa (2006) and Tirole (1996) establish links between actions today and global actions in the future. In Tirole (1996), for instance, the reputations of members within a group are observed imperfectly. A member’s current incentives are affected by his past behavior and, because of imperfect observation, by actions of other agents in his group. The destruction of reputation in one generation may lead to reduced incentives for reputation for all future generations and eliminate the ability of the market to restore good faith. By contrast, my model generates history dependence due to informational differences between institutional structures. My model is most suitable in situations in which global coordination can be mutually beneficial to all parties but where deviation from the current institution requires information that is obscured by the institution itself.

In the herding literature, pioneered by Banerjee (1992), Bikhchandani, Hirshleifer & Welch (1992), and Welch (1992), the ability to observe the actions of past actors may lead individual agents to follow past play rather than their own signal. This can lead to an information cascade where new information is eliminated from the environment and all agents continue to make the wrong, inefficient choice. Like the herding literature, my model points out situations in which information needed to make globally welfare improving changes is eliminated by actions taken in the past. Whereas the herding literature concentrates on individual-level actions, I concentrate on the endogenous formation of institutions that globally eliminates information.

In the convention literature, a game with multiple equilibria is augmented with small amounts of persistent randomness to study how random mutations in strategies might affect the persistence of equilibria. Developed by Foster & Young (1990), Kandori (1992), and Young (1993), the conventions literature has the appealing characteristic that it often selects a unique equilibrium in games with multiple equilibria. This paper differs from the conventions literature in that it studies a specific channel by which history and information

\[^{1}\text{For more general theoretical treatments of herding, see Chamley (1999) and Smith & Sorensen (2000). Information cascades have also been studied in the lab by Anderson & Holt (1997) and more recently by Goeree, Palfrey, Rogers & McKelvey (2007). In both studies, reversals of cascades are observed in the long run suggesting that individual agents may overweight their own information and mitigate inefficient herding.}\]

\[^{2}\text{For an overview of conventions see Ellison (1993). The convention literature has been experimentally explored in Van Huyck (1997) and Van Huyck (2001).}\]
together might dynamically influence final outcomes. The informative properties of signals
depend critically on the institution that have formed and thus the probability that individual
explore other strategies is based directly on the institutions that have formed from past play.

3 The Model

3.1 Overview

In this section, I build a theoretical framework that illustrates the informational properties of
institutions. I begin by developing a benchmark model of a market with unobservable quality,
costly certification, and heterogeneity in seller types where the distribution of seller types is
known. Using a simple Walrasian approach, I define and characterize the possible competitive
equilibria. I show that two equilibria may exist, non-certifying and certifying, and that these
equilibria vary in terms of both efficiency and information about the underlying environment.

The model developed here is intentionally stylized in order to concentrate on market-
level information externalities. In order to ensure that there are two stable equilibria, I use
a discrete certification technology. This eliminates many of the complications that out-of-
equilibrium beliefs pose to equilibrium selection by ensuring that payoff relevant states are
well defined. The baseline model also uses a common knowledge assumption about the
distribution of seller types to reduce complication. In the appendix, I extend the baseline
model to a game theoretic environment where there is dynamic learning about the distribu-
tion of types. Given assumptions on the learning of buyers and sellers, the prices for which
an equilibrium exists as time grows to infinity are similar to the prices in the static setting.

While the goal of this research is to understand dynamic information effects, I study the
static competitive equilibrium for three reasons. First, from an experimental perspective,
there is clear evidence that experimental markets converge toward the competitive equi-
librium when the trading mechanism is a double auction. From a theoretical perspective,
simultaneous move double auctions also converge to the competitive equilibrium as the num-
ber of players grows large. As such, the use of a competitive equilibrium as a solution concept
is meant to generate reasonable benchmark predictions for the experimental environment.

---

3For problems of existence in a competitive equilibrium with adverse selection, see Rothschild & Stiglitz
(1976). For a more general model of equilibrium selection in a market with adverse selection and a continuous
contract space see Gale (1992).

4See Walker & Williams (1988) for a discussion of convergence across varying institutions.
Second, the model environment is designed with market power on the sellers’ side, so that all rents from trade are likely to be appropriated by the sellers in any fully-specified bargaining process. In the appendix, I provide an explicit game-theoretic auction model. I show that the set of pure strategy equilibria from the game-theoretic model is identical to the competitive equilibria with the key exception being that the certifying equilibrium does not exist in an environment where a single good seller, who chooses to sell an uncertified unit over a certified unit, can generate a new equilibrium.

Finally, any particular dynamic game will be sensitive to assumptions made about matching, memory, information, updating, bargaining, utility functions, and the formation of out-of-equilibrium beliefs. On the other hand, if agents are anonymous, trade is frictionless, and the number of players grow large, a few general restrictions should hold true: A buyer should be able to buy uncertified and certified goods at the cheapest price he can negotiate, sellers should be able to sell at the highest price they can negotiate, and all buyers and sellers should be able to enter and exit as many negotiations as they would like before the final resolution of the market. These restrictions bring us naturally to outcomes that are \textit{ex post} stable, a property that directly leads to a set of prices identical to those in one of the competitive equilibria.

3.2 Benchmark Model

I consider a world with high ($H$) and low ($L$) quality units. There are $N$ buyers indexed by $i \in \{1, \ldots, N\}$ divided into a finite number of types $b \in B$. There are $M < N$ sellers indexed by $j \in \{1, \ldots, M\}$ divided into three types $s \in \{G, C, B\}$ (Good, Conditional, and Bad). The number of buyers who are of type $b$ is $N_b$. Likewise the number of sellers who are of type $s$ is $M_s$. Each buyer can consume a single high or low quality unit. Likewise, each seller can produce a single high or low quality unit.

While a more general structure is included for extensions, I initially consider the case where there is only one type of buyer denoted by $b_0$. Buyers of type $b_0$ have gross utilities for consuming the high and low quality good of $U^H$ and $U^L$ relative to a separable numéraire good, are loss and risk neutral, and receive zero utility if they do not trade.

The quality of units being traded is initially unknown to buyers. Sellers have available a costly technology that certifies quality. Certification costs $T \in (0, U^H - U^L)$ and eliminates all uncertainty over the quality of the unit to the buyer. I assume that certification costs are paid by the seller when a trade occurs and that $T$ is common knowledge. Since $U^H > U^L$,
certifying the low quality unit can not increase its value and thus a certified low quality unit will never be offered by a profit maximizing firm. I thus omit this choice from the analysis and restrict all certified units to be of high quality.

Given the choice over certification, buyers and sellers may exchange in three markets \( m \in \mathcal{M} = \{C, NC, \emptyset\} \). \( C \) is a market for high quality certified units, \( NC \) is a market of uncertified units, and \( \emptyset \) is a market without trades. In the certified market, all three types of sellers produce the high quality unit. In the uncertified market, a seller is free to exchange a unit of either quality.

If a seller exchanges a unit of quality \( L \), she pays a cost of \( C^L \) which is constant across all sellers. If a seller exchanges a unit of quality \( H \), she pays a cost \( C^H_s \) which differs by seller type. Types are defined such that

\[
(1) \quad C^H_B > C^H_C > C^H_L > C^H_G
\]

and

\[
(2) \quad C^H_B > C^H_L + U^H - U^L - T > C^H_C.
\]

Condition 1 distinguishes sellers of type-\( G \) from those of type-\( C \) and type-\( B \) by giving them incentives to produce high quality units if they trade in the uncertified market. Condition 2 distinguishes type-\( C \) sellers from type-\( B \) sellers by giving type-\( C \) sellers incentives to adopt the certification institution and produce high quality goods if the market price for low quality goods is sufficiently low.

As will be shown later, the adoption of certification by sellers of type \( C \) alters their production decision so that it coincides with the social optimum. In a proscriptive sense, the formation of a certifying equilibrium resolves the problem of hidden action for sellers of type \( C \). I thus define the “degree of moral hazard” in the environment as the proportion of sellers who are of type \( C \). Let the proportion of sellers who are type \( G \), \( C \), and \( B \) be given by \( g \), \( c \), and \( b \) respectively where \( g = \frac{MG}{M} \).

**Definition 1 Degree of moral hazard:** The proportion of type-\( C \) sellers \( c \).

I make the simplifying assumption that \( C^L < U^L \) so that trade is potentially welfare improving. In order to easily analyze the welfare implications of the model, I make the additional assumption that \( C^H_B - C^L < U^H - U^L \) so that the total surplus created from producing a high quality good is always higher than producing a low quality good.
A buyer of type $b \in B$ who matches with a seller of type $s \in \{G, C, B\}$ in market $m \in M$ at price $P^m$ receives utility $u(m, P^m, b, s)$. The market affects this utility by restricting the set of actions that a seller can take. For instance, if a buyer matches with a type-$C$ seller in market $NC$, the conditional seller is free to exchange a unit of either high or low quality and optimally supplies a low quality unit. If the buyer had matched with the same seller in market $C$, the conditional seller is constrained and would supply a high quality unit. In the baseline case:

$$u(m, P^m, b_0, s) = \begin{cases} U^H - P^C & \text{if } m \in C, s \in \{G, C, B\} \\ U^H - P^{NC} & \text{if } m \in NC, s \in \{G\} \\ U^L - P^{NC} & \text{if } m \in NC, s \in \{C, B\} \end{cases}$$

Similarly, a seller of type $s$ who matches with a buyer of type $b$ in market $m$ at price $P^m$ receives utility $v(m, P^m, b, s)$. A seller maximizes expected value and thus, given optimal action in both markets, has a utility function of:

$$v(m, P^m, b, s) = \begin{cases} P^C - C^H_s - T & \text{if } m \in C, s \in \{G, C, B\} \\ P^{NC} - C^H_s & \text{if } m \in NC, s \in \{G\} \\ P^{NC} - C^L & \text{if } m \in NC, s \in \{C, B\} \end{cases}$$

Note that in the baseline case $v(m, P^m, b, s)$ is independent of $b$.

The description of the competitive equilibrium is comprised of three parts: an attainable allocation $(D, S)$, a belief system $\mu$, and a price system $P$.

**Attainable Allocations:** The number of buyers of type $b$ who demand from market $m$ is denoted by $D(m, b)$. An allocation of buyers is a function $D : M \times B \rightarrow \mathbb{I}_+$ such that $\Sigma_{m \in M} D(m, b) = N_b$. Likewise, the number of sellers of type $s \in \{G, C, B\}$ who supply in market $m$ is denoted by $S(m, s)$. An allocation of sellers is a function $S : M \times \{G, C, B\} \rightarrow \mathbb{I}_+$ such that $\Sigma_{m \in M} S(m, s) = M_s$. An allocation $(D, S)$ is **attainable** iff $\Sigma_{s \in \{G, C, B\}} S(m, s) = \Sigma_{b \in B} D(m, B)$ for $m \in \{C, NC\}$. Note that the market clearing condition is not binding in the $\emptyset$ market.

**Belief System:** Buyers and sellers form beliefs about the types of agents exchanging each contract. Let $\mu_b(m, s)$ denote the subjective probability that a unit purchased in market $m$ by a buyer is in fact supplied by a seller of type $s$. Let $\mu_s(m, b)$ denote the
subjective probability that a unit sold in market \( m \) by a seller is in fact bought by a buyer of type \( b \). A belief system is a pair of beliefs \( \mu = (\mu_b, \mu_s) \) such that \( \mu_b(m, s) : M \times \{G, C, B\} \to \mathbb{R}_+ \) satisfies \( \sum_s \mu_b(m, s) = 1 \) for every \( m \) and \( \mu_s(m, b) : M \times B \to \mathbb{R}_+ \) satisfies \( \sum_b \mu_s(m, b) = 1 \) for every \( m \).

**Price System:** A price system is a function \( P : M \to \mathbb{R}_+ \). I define \( P^c, P^{NC}, P^a \) as the prices in each market.

Suppose that a buyer of type \( b \) purchases a unit in market \( m \) at price \( P^m \). If the buyer’s beliefs are given by \( \mu_b(m, s) \), his expected utility is given by

\[
\sum_s u(m, P^m, b, s)\mu_b(m, s),
\]

where \( u(m, P^m, b, s) \) is the utility received when a seller sells her market constrained optimal unit to the buyer. A buyer will choose a market that maximizes equation (5). Consequently, an equilibrium allocation must assign all buyers of type \( b \) to markets that are in the arg max of equation (5):

\[
D(m^*, b) \neq 0 \iff m^* \in \arg \max_m \sum_s u(m, P^m, b, s)\mu_b(m, s) \quad \forall b.
\]

Likewise, suppose that a seller sells a unit in market \( m \) at price \( P^m \). If the seller’s beliefs are given by \( \mu_s(m, b) \) her expected utility is given by

\[
\sum_s v(m, P^m, b, s)\mu_s(m, b),
\]

where \( v(m, P^m, b, s) \) is the value the seller receives from selling her optimal unit to a buyer of type \( b \) subject to the constrains of the market she has entered. Like the buyer, any competitive equilibrium requires:

\[
S(m^*, s) \neq 0 \iff m^* \in \arg \max_m \sum_b v(m, P^m, b, s)\mu_s(m, b) \quad \forall s.
\]

Finally, the belief that a unit in a market is supplied by a seller of type \( s \) is equal to the proportion of agents selling in a market of a specific type. Likewise, the probability a unit in a market is bought by a buyer of type \( b \) is proportion to the agents buying in a market of that type. If a market has no trades in equilibrium, then these proportions are not well-defined and beliefs may be arbitrary. In the entire analysis, I look at the case where there is at least
one type-B seller who always trades in the uncertified market. Thus buyers’ beliefs about the uncertified market are always well defined. Since the utility of other trades do not depend on beliefs, there is never a case where an equilibrium is supported by out-of-equilibrium beliefs.

**Definition 2 Competitive Equilibrium:** A Competitive Equilibrium is a triple \( (D \times S, \mu, P) \) consisting of an attainable allocation \( (D \times S) \), beliefs \( \mu \), and a price system \( P \) which satisfy:

\[
E.1 : \quad S(m^*, s) \neq 0 \iff m^* \in \arg \max_m \sum_b v(m, P^m, b, s)\mu_s(m, b) \quad \forall s,
\]

\[
E.2 : \quad D(m^*, b) \neq 0 \iff m^* \in \arg \max_m \sum_s u(m, P^m, b, s)\mu_b(m, s) \quad \forall b,
\]

\[
E.3 : \quad \mu_b(m, s) = \frac{S(m, s)}{\sum_s S(m, s)} \text{ and } \mu_s(m, b) = \frac{D(m, b)}{\sum_b D(m, b)}.
\]

Analysis of the competitive equilibrium is simplified by two characteristics of the benchmark environment. First, the sellers valuation \( v(m, P^m, b, s) \) is independent of the buyer that she is matched with and thus \( \mu_s(m, b) \) does not affect the sellers decision. It follows that condition \( E.1 \) can be reduced to

\[
E.1b : \quad S(m^*, s) \neq 0 \iff m^* \in \arg \max_m \sum_b u(m, P^m, b, s) \quad \forall s,
\]

which is the requirement that all sellers enter the market where the difference between price and the cost of their constrained optimal production choice is largest. Second, since all buyers share the same utility function given in equation (3), only beliefs about \( \mu_b(\mathcal{NC}, G) \) affect utility. Since sellers actions only depend on prices, I define a function \( \pi^H(\Delta P) \) where \( \pi^H : P \to [0, 1] \) is a buyers beliefs about the proportion of high quality units in the uncertified market for a difference in prices of \( \Delta P \). Note that \( \pi^H(\Delta P) = \mu_b(\mathcal{NC}, G) \) for \( \Delta P = P^c - P^{NC} \).

Given these two simplifications, I solve for the set of competitive equilibrium in two steps. I first determine the set of \( S(m, s) \) that satisfy \( E.1b \) for each potential price system \( P \). I then determine the set of \( D(m, b) \) for which \( E.2 \) is satisfied for each potential price system \( P \) and (correct) belief system \( \mu_b(m, s) \). I restrict attention to the case where \( M_B \geq 1 \) so that \( \mu_b(\mathcal{NC}, s) \) is well defined.

### 3.2.1 Supply Decisions by Sellers

For a price system \( P \), a seller produces in the certified market if

\[
(9) \quad v(C, P^c, b, s) > v(\mathcal{NC}, P^{NC}, b, s).
\]
For all sellers, this reduces to the condition

\[ PC - C^H_s - T \geq P^{NC} - \min(C^H_s, C_L). \]

Define \( P^c \) and as the maximum willingness to pay for a certified unit across all buyers. Similarly, define \( P^{NC} \) as the minimum willingness to pay across all buyers for an uncertified unit. In the baseline model \( P^c = U^H \) and \( P^{NC} = U^L \). In equilibrium it will be the case that \( P^{NC} \leq P^{NC} \leq P^c \leq P^c \) so that i) \( \Delta P \) is always either zero or positive and ii) both buyers and sellers have incentive to trade in either the certified or uncertified market for prices within these bounds. Given the definition of Good, Conditional, and Bad seller types:

- A seller of type \( G \) has \( C^G_H \leq C^L \) and will always produce high quality units. A type-\( G \) seller will trade in the uncertified market if \( \Delta P \leq T \).

- A seller of type \( C \) has \( C^C_H \in (C^L, C^L + P^c - P^{NC} - T) \) and will produce either low quality units to the uncertified market or high quality units to the certified market. A type-\( C \) seller will trade to the uncertified market if \( \Delta P \leq T + (C^C_H - C^L) \).

- A seller of type \( B \) has \( C^B_H \geq C^L + P^c - P^{NC} - T \). Given the bounds on possible prices, type-\( B \) sellers never sell high quality units and will always produce low quality units in the uncertified market.

**Lemma 1** For a price system \( P \) with \( P^{NC} \leq P^{NC} \leq P^c \leq P^c \):

\[
S(C, G) = \begin{cases} 
M_G & \text{if } \Delta P < T \\
[0, M_G] \in \mathbb{I}_+ & \text{if } \Delta P = T \\
0 & \text{otherwise}
\end{cases}
\]

\[
S(C, C) = \begin{cases} 
M_C & \text{if } \Delta P < T + C^H_C - C^L \\
[0, M_C] \in \mathbb{I}_+ & \text{if } \Delta P = T + C^H_C - C^L \\
0 & \text{otherwise}
\end{cases}
\]

\[
S(C, B) = \begin{cases} 
0 & \text{if } \Delta P \in (0, P^c - P^{NC})
\end{cases}
\]

\[
S(NC, s) = M_s - S(C, s) \text{ for } s \in \{G, C, B\}.
\]

**Proof.** All proofs given in the appendix. ■
3.2.2 Demand Decisions by Buyers

Suppose that a type $b_0$ buyer has a choice of buying a certified unit at price $P^C$ or a non-certified unit at price $P^{NC}$. Let $\pi^H(\Delta P)$ be a buyer’s belief about the proportion of high quality units in the uncertified market given the difference in price between certified and uncertified units. A buyer is indifferent between purchasing in the certified and uncertified market if

\begin{equation}
\pi^H(\Delta P)U^H + (1 - \pi^H(\Delta P))U^L - P^{NC} = U^H - P^C.
\end{equation}

**Lemma 2 In Equilibrium:**

- If $\Delta P > T$ the buyer believes that all type-$G$ sellers will certify their goods and thus that $\pi^H(\Delta P) = 0$. In this case, a risk neutral buyer prefers to purchase the certified unit as long as $\Delta P < U^H - U^L \equiv P^C - P^{NC}$ and is indifferent between buying a non-certified unit and not purchasing if $P^{NC} = U^L$.

- If $\Delta P \leq T$ the buyer believes that all sellers trade in the uncertified market. In this case $\pi^H(\Delta P) = g$ and a risk neutral buyer prefers to purchase the uncertified unit as long as $\Delta P \geq (1 - g)(U^H - U^L)$.

3.3 Market Equilibria

Since there are more buyers than sellers and all buyers have identical utility functions, buyers must be indifferent between purchasing and not purchasing a unit of the good. Setting payoffs in equation (11) equal to zero yields the following two equilibria$^5$

- **Certifying Equilibrium**: $P^C = U^H$, $P^{NC} = U^L$. Type-$G$ and type-$C$ sellers produce and sell certified high quality units. Type-$B$ sellers produce uncertified low quality units.

- **Non-certifying Equilibrium**: $P^{NC} = U^H - (1 - g)(U^H - U^L)$, $P^C = U^H$. Type-$G$ sellers produce uncertified high quality units. Type-$C$ and type-$B$ sellers produce uncertified $L$-quality units. All buyers buy from the uncertified market.

$^5$In general, a partial-certifying equilibrium will also exist where $\Delta P = T$ and type-$G$ sellers are indifferent to trading in the certified and uncertified market. In the baseline model, since all buyers have the same utility function and seller types are discrete, the partial-certifying equilibrium is degenerate. See section [3.3.3](#) for an extension of the model where partial-certifying equilibria are more likely to exist.
Theorem 1 **Existence:** The Certifying Equilibrium always exists. The Non-Certifying equilibrium exists if $(1 - g)(U^H - U^L) \leq T$.

Multiplicity occurs in this market due to the cost associated with certification which diminishes the incentive of type-\(G\) sellers to identify the quality of their product. The existence of the non-certifying equilibrium requires the cost of certification to be larger than the discount that buyers require to trade for uncertified goods. This will be the case if, for instance, the proportion of type-\(G\) sellers is high.

When the price difference between the certified and uncertified market is large, type-\(G\) sellers will respond by selling in the certified market. Since the probability of receiving an high quality unit in the uncertified market is zero, a buyers’ willingness to pay for uncertified units falls to \(U^L\) and the difference in price between the uncertified and certified markets becomes \(\Delta P = U^H - U^L = P^C - P^{NC}\). Type-\(C\) sellers, defined as having \(C^H_C - C^L_C < P^C - P^{NC} - T\), also choose to certify since the profit gained from switching markets is greater than the increase in production and transaction costs.

### 3.3.1 Welfare

The relative welfare of the non-certifying and certifying equilibrium depend critically on the degree of moral hazard in the market and the cost of certification. For type-\(C\) sellers, the adoption of certification increases the quality of their goods. Not factoring in the certification cost, this leads to a efficiency gain of:

\[
M_C[U^H - U^L - (C^H_C - C^L_C)].
\]

However, in the certifying equilibrium, both type-\(G\) and type-\(C\) sellers certify their product leading to a total certification cost of:

\[
(M_G + M_C)T.
\]

Combining the two terms and normalizing by \(M\) yields:

**Theorem 2** The non-certifying equilibrium is constrained Pareto efficient if

\[
(g + \epsilon)T > \epsilon[U^H - U^L - (C^H_C - C^L_C)].
\]

Otherwise, the certifying equilibrium is constrained Pareto efficient.
The certifying equilibrium is likely to be efficient when the degree of moral hazard in the environment \( c \) is high and the proportion of type-\( G \) sellers in the environment is low. It is also more likely to be efficient if the cost of certification \( T \) is low or the additional surplus for altering the production decision of a conditional type from low quality units to high quality units is large.

### 3.3.2 Market Information

Suppose that a sequence of markets generate either the certifying or non-certifying equilibrium above. If a new homogeneous buyer enters the market and can observe price and the volume of trades in each market, what can he deduce about the proportion of sellers who are good, conditional, and bad?

In the certifying equilibrium, the prices \( P^C = U^H \) and \( P^{NC} = U^L \) only provide information about the demand function of buyers. Since only bad sellers trade in the non-certified market, the share of goods traded in the uncertified market provides information on the proportion of sellers who are of type-\( B \) but provides no additional information about the relative proportion of type-\( G \) and type-\( C \) sellers.

By contrast, in the non-certifying equilibrium, the non-certifying price \( P^{NC} = U^H - (1 - g)(U^H - U^L) \) carries information about the proportion of good sellers. Given only the non-certifying price, any agent in the market can determine the proportion of the sellers who are of type \( G \). Since no sellers certify their units, sellers of type \( C \) and type \( B \) are indistinguishable.

**Theorem 3** In a non-certifying equilibrium, price is a sufficient statistic for the proportion of type-\( G \) sellers in the environment. In the certifying equilibrium, no market signal generates information that can distinguish between type-\( G \) and type-\( C \) buyers.

The difference in information that is generated in a market with or without the adoption of the certification institution is stark. In the non-certifying equilibrium, the proportion of type-\( G \) sellers in the market can be inferred directly from the market price, a primitive that is inherently observable in the market. In the certifying equilibrium, no information is generated when the proportion of type-\( G \) and type-\( C \) sellers changes. This may lead to persistence of the certification institution since the true state of the world is not transmitted through individual and group decisions.
One interesting corollary from the previous section is that if a market has converged to a certifying or non-certifying equilibrium, *ex post* revelation of uncertified trades does not generate new information about the distribution of seller types. In the case of the non-certifying equilibrium, this result arises because the pooling price is a sufficient statistic for the proportion of type-$G$ sellers in the market. In the case of the certifying equilibrium, this result occurs due to only low quality units being traded in the uncertified market.

In an experimental setting, agents typically exhibit some aversion toward accepting actuarially fair gambles. This section briefly comments on how differences in the willingness to accept gambles can lead to a partial-certifying equilibrium where *ex post* disclosure of trade quality can generate new informative. Due to its tractable nature and players’ responses to survey questions at the end of the experiment, I model the aversion toward gambles using loss aversion. All the results of this section carry over to alternative models using risk or ambiguity aversion. A full discussion of both loss aversion and the partial certifying equilibrium in relation to learning is included in the appendix.

Suppose that some buyers are loss averse and put a greater weight on aggregate losses than gains. Let $\mathcal{B} = \{\lambda_1, \lambda_2, \ldots, \lambda_N\}$ where $\lambda_i$ is the idiosyncratic loss aversion parameter for buyer $i$ with $\lambda_i \geq 1$ for $i \in \{1, 2, \ldots, N\}$. Without loss of generality, I order buyers according to their risk aversion parameter such that $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_N$ and again normalize the utility obtained from not trading to zero.

For a price system $P$ with $P^{NC} \leq P^{NC} \leq P^{C} \leq P^{C}$, a buyer $i$ buying from market $m$ at price $P^m$ from a seller of type $s$ gets utility

$$u(m, P^m, \lambda_i, s) = \begin{cases} U^H - P^C & \text{if } m \in \mathcal{C}, s \in \{G, C, B\} \\ U^H - P^{NC} & \text{if } m \in \mathcal{NC}, s \in \{G\} \\ -\lambda_i(P^{NC} - U^L) & \text{if } m \in \mathcal{NC}, s \in \{C, B\} \end{cases}.$$  

In the non-certifying equilibrium, the market price $P^{NC} > U^L$ and there is a potential for losses in the market. Since buyers are heterogeneous in loss aversion, the aggregate demand curve for uncertified units becomes downward sloping and the uncertified price is pinned down by the loss aversion of the $M^{th}$ buyer. If the $M^{th}$ buyer is sufficiently loss averse, he may be unwilling to trade for uncertified units at a price where $\Delta P \geq T$. In this case, partial certifying equilibria may form. Let $S^C$ be the number of certified units in an
equilibrium. Then for each $S^C < M$, a partial certifying equilibrium may exist with the following properties:

- **Partial-Certifying Equilibrium:** $P^{NC} = U^H - T$, $P^C = U^H$. Type-$C$ and type-$B$ sellers produce uncertified low quality units. $S^C$ type-$G$ sellers produce certified uncertified high quality goods. $M - S^C$ type-$G$ sellers produce high quality goods. Buyers $i \in \{1, \ldots, M - S^C\}$ buy uncertified units. $S^C$ other buyers buy certified units.

In the benchmark model, the partial-certifying equilibrium was degenerate because both type-$G$ sellers and all buyers needed to be indifferent between trading in the certified and uncertified market. With heterogeneity in buyer preferences, however, partial-certifying equilibrium may be stable since the willingness to pay for uncertified units is decreasing in loss aversion leading to a downward sloping aggregate demand function.

In the partial certifying equilibrium, since $P^{NC} = U^H - T$ and $P^C = U^H$, price alone does not convey information about the proportion of type-$G$ sellers. However, there are two pieces of information that can be used in estimation. First, the certified market is composed entirely of type-$G$ sellers. Second, the buyer with the highest level of risk aversion must be willing to trade in the uncertified market. Combining these two elements, a lower bound for $g$ can be constructed:

\[
g \geq g = \frac{M(U^H - U^L - T)\lambda_{M-S^C} + S^CT}{M(U^H - U^L - T)\lambda_{M-S^C} + MT}.
\]

$g$ can be thought of as the smallest proportion of type-$G$ sellers that could support the partial certifying equilibrium at a given time. For instance, if $S^C = 0$ and $\lambda_M = 0$, then equation (16) reduces to

\[
g \geq \frac{(U^H - U^L - T)}{(U^H - U^L)}
\]

which is a rewriting of the necessary condition for existence of the non-certifying equilibrium given in Theorem 1. As is clear from this example, $g$ may be far away from the true proportion of type-$G$ sellers leading to limited inference from market information alone.

Since there are both high and low quality units being traded in the uncertified market and inference is imperfect, public information about the proportion of high quality units in the uncertified market can generate new information unavailable from market signals. If, for instance, the number of uncertified units was publicly revealed in a partial-certifying
equilibrium, an observer could determine the proportion of type-$G$ sellers in the environment by adding the number of high units in the uncertified market to the size of the certified market and dividing by $M$.

4 The Experiment

In the theory section, I developed and characterized the set of possible market equilibria that could arise in a world with moral hazard, costly certification, and seller heterogeneity. I showed that if an equilibrium forms where a certification institution is adopted, information about the degree of moral hazard in the underlying environment was eliminated. This could lead to persistence of inefficient institutions since information about changes in the underlying environment was lost.

In this section I empirically explore the informational properties of institutions via laboratory experiments. The experimental treatments are designed to study the initial formation of non-certifying and certifying equilibria, adaptation of these market equilibria to changes in the underlying environment, and the information generated in the market. I use a pre/post design in which there were two environments — Safe and Hazardous — which vary in the composition of sellers in the market. In the safe environment, the proportion of good sellers in the market is large, thus favoring the formation of the non-certifying equilibrium. In the hazardous environment, the proportion of good sellers is small, so only the certifying equilibrium should exist. I begin half the sessions in the safe environment and switch to the hazardous environment midway through. In the other half, I reverse the order, starting in the hazardous environment and ending in the safe environment.

Compared to other empirical tools, experimentation has many advantages for studying informational persistence. First, supply and demand are exogenously set and do not have to be estimated in conjunction with the parameters of interest. This generates more explicit hypothesis tests, allows for exogenous changes to the environment, and guarantees that alternative market equilibria do or do not exist. Second, experimentation allows for the isolation of the information channel and eliminates other potential channels of persistence. Finally, primitives that are typically unobservable such as beliefs and risk aversion can be elicited in conjunction with the main experiment and used to study what forces shape the dynamics of the market.

Where possible, the experimental design matches the model developed in the previous
section. I choose to allow buyers and sellers to trade multiple units and implement a downward sloping demand function in an effort to increase stability in the market. Multiple units increase the thickness of the market in each period and allow for an excess demand without the use of passive buyers who might cause noise in the experiment by trying to play. A downward sloping demand function also generates some surplus for the buyers, which has been shown by Holt, Langan & Villamil (1986) to improve the speed of convergence in markets. The supply and demand curves are constructed so that no seller or buyer could change the equilibrium price by more than 10 points by withholding their entire supply or demand from the market. Since no buyer or seller has market power, the competitive equilibrium for the experimental environment varies only slightly from the simplified model of section 3.2.

4.1 Overview

I constructed an experimental market where each subject was assigned to the role of buyer or a seller and allowed to exchange units of two possible qualities: a high quality “red” unit and a low quality “blue” unit. Buyers and sellers had access to two possible markets — an uncertified market and certified market — each conducted simultaneously as a double auction.

Each experimental session consisted of 6 sellers and 5 buyers. Sellers could produce a maximum of two units creating an aggregate supply of 12 units. Buyers could each purchase up to a total of three units creating an aggregate demand of 15 units. Excess demand in the market was implemented to allow sellers to capture any residual surplus that existed in either of the two markets and to capture rents generated through certification.

The main experiment lasted 24 periods and was composed of two phases: the market game and a bonus game. The market game consisted of two simultaneous double auctions — one with certification and one without — in which buyers and sellers were anonymous and could not create reputations. In the first three periods of the experiment, the market game lasted four minutes to allow for subjects to become accustomed to the interface. In the remaining periods, the market game lasted two minutes.

In the bonus game, subjects were asked to guess how many of the sellers had lower cost for producing the high valued unit than the low valued unit. If a subject correctly guessed the proportion of type-$G$ sellers in the market, they earned a bonus for the round. I tracked responses for the bonus game across periods and use them as a proxy for beliefs about the underlying degree of moral hazard in the game.
Experimental treatments were designed to study the initial formation of non-certifying and certifying equilibria, how these market equilibria adapt to changes in the underlying environment, and the information generated in the market. Experimental sessions were divided into four treatments which varied in the degree of moral hazard (the number of type-C sellers) and in the amount of public information available about past trades. In half the sessions, subjects began in a Safe (S) environment where the degree of moral hazard was low and switched to a Hazardous (H) environment with a high degree of moral hazard at period 13. The order of environments was reversed in the remaining sessions.

Ex post information about the quality of uncertified trades may improve inference about the underlying environment in market where a partial-certifying equilibrium forms. To test this idea, I introduced an information treatment where subjects were informed about the number of high and low quality units traded in the uncertified market in the previous period. This data was used in conjunction with beliefs data from the bonus game to study the informational properties of the non-certifying equilibrium.

Finally, after the main experiment, risk and loss aversion measures were constructed based on a risk aversion game and an exit survey. In the risk aversion game, subjects made a decision between a safe gamble and a series of risky gambles that varied in their expected value. The number of safe choices was recorded as a measure of risk aversion. In the buyers version of the survey, buyers were asked how they decided upon a price they were willing to pay for an uncertified good. 53% of buyers responded to this question by saying that they were unwilling to take losses on an uncertified trade. As a measure of loss aversion, I coded a binary variable which was one if a subject wrote that they were “unwilling to take losses” and zero otherwise.

4.2 Valuations and Trading Mechanism

In an experimental period, each of the six sellers could sell a total of two units across both markets in any combination of high and low quality. Trades were conducted in points and converted to Swiss Francs at the end of the experiment at a conversion rate of 30 points to 1 Swiss franc (37 points to 1 Dollar US).

As shown in Table 1, sellers could be assigned one of three possible cost functions for producing high and low quality units which I designate as G, C, and B. Following the theory model, type-G sellers had a lower cost for producing a high quality unit, type-C.

---

6The loss aversion parameter will be covered in greater detail in sections 4.5 and 5.1.2
sellers were conditional, and type-$B$ sellers sold uncertified low quality units only.

<table>
<thead>
<tr>
<th>Table 1: Seller Costs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Uncertified Low</strong></td>
</tr>
<tr>
<td>$G$</td>
</tr>
<tr>
<td>$C$</td>
</tr>
<tr>
<td>$B$</td>
</tr>
</tbody>
</table>

The certification cost, known to both buyers and sellers, was 60 points. If the difference in price between the certified and uncertified market grew larger than the certification cost, type-$G$ sellers had an incentive to sell a high quality unit in the certified market rather than a high quality unit in the uncertified market. Likewise if the difference in price between the certified and uncertified market grew larger than 90, type-$C$ sellers had an incentive to sell a high unit in the certified market rather than a low unit in the uncertified market.

In an experimental period, each of the five buyers could purchase a total of three units across both markets. As shown in Table 2, each buyer’s demand schedule was downward sloping. Conditional on buying a unit, the valuation of both the high and low quality units declined for each unit purchased. Thus, if buyer 1 had purchased a low quality unit and then purchased a high quality unit, he would have received 140 points for the first purchase and 220 for the second minus the price he paid for each unit. The demand functions of buyers four and five were staggered slightly to smooth the aggregate demand function.

<table>
<thead>
<tr>
<th>Table 2: Buyer Valuations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Buyers 1-3</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Unit 1</strong></td>
</tr>
<tr>
<td>High</td>
</tr>
<tr>
<td>Low</td>
</tr>
</tbody>
</table>

Earnings from one period did not carry over into the following periods. At the start of each period, buyers were given 100 points as an initial cash endowment. After each trade within a period, the type of unit purchased was revealed and his earnings or losses from the transaction were added to or subtracted from his current cash. If at any point in a period a buyer’s total cash was negative, he was not allowed to make further trades until the start of the next period. This form of bankruptcy was only observed in two instances across all experimental sessions.

Information about seller costs and buyer valuations was private information. Sellers were
shown the three possible cost functions that they might be assigned in the instructions and told that their cost schedule might change across periods. Sellers were not given information on the assignment of other sellers or on the demand schedule of the buyers. Buyers were given only their own demand schedule and were informed via the discussion of the bonus game that some of the sellers might have a lower cost for producing high valued units than low valued units. Buyers and sellers were both informed about the actual cost of certification.

Trading in the market was conducted through a computerized exchange where both buyers and sellers were anonymous and the only distinguishable feature between the various seller offers and buyer bids were the public price and quality characteristics available in each of the two exchanges.

A seller who posted an offer to the uncertified market publicly submitted an asking price and secretly selected the quality of the offered unit. During the period, quality was revealed only to the buyer who purchased the unit and was not publicly disclosed to the market. A buyer who bid in the uncertified market publicly submitted a bid price and a quality request. Quality requests in the uncertified market were not binding and a seller who filled a request had the option of supplying either quality good.

In the certified market, the quality of the seller’s offer was observable and quality requests by buyers were binding. If a seller transacted in the certified market, either by having an offer accepted or fulfilling a buyer’s trade request, she was charged the certification fee of 60 points.

In each period, a history of trades from the current period was available in graph form for all subjects in the market. Certified trades showed up in this graph in the color of the actual unit traded while uncertified trades showed up as black lines. Each seller could have one certified offer and one uncertified offer open at one time. Likewise, each buyer could have one certified bid and one uncertified bid open at any given time. If a seller sold her last unit or a buyer exhausted his demand, all remaining open contracts were automatically withdrawn from the market. Bids and offers could be changed or withdrawn at any time with no restriction on pricing. Agents could select any offer from the opposite side of the market and were not bound to accept the lowest possible price.

4.3 Bonus Game

In the bonus game, individuals were asked to guess how many of the sellers had lower cost for producing the high quality unit than for producing the low quality unit. If a subject
correctly guessed the proportion of type-G sellers in the market, they earned a 20 point bonus for the round. The monetary payment for the bonus game was intentionally low to ensure that subjects did not change actions in the market game for the sole purpose of gaining information to increase their bonus game earnings. I tracked responses for the bonus game across periods and use them as a proxy for beliefs about the underlying degree of moral hazard in the market.

4.4 Treatments

Experimental treatments were designed to study 1) the initial formation of equilibria, 2) how these market equilibria adapt to changes in the underlying environment, and 3) the information generated by the market. Experimental sessions were divided into four treatments which varied in the degree of moral hazard (the number of type-C sellers) and in the amount of public information available about past trades. In half the sessions, subjects began in a safe (S) environment where the degree of moral hazard was low and switched to a hazardous (H) environment with a high degree of moral hazard at period 13. This order was reversed in the remaining sessions. To distinguish between periods before and after the switch, I use Pre and Post superscript appended to the environment identifier.

<table>
<thead>
<tr>
<th>Table 3: Treatments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
</tbody>
</table>

The hazardous and safe environments varied in the number of sellers who were assigned to the three seller types. In the safe environment, five of the sellers were of type G and one seller was of type B. In the hazardous environment, one seller was of type G, four sellers were of type C, and one seller was of type B. The single type-B seller was included in both treatments in order to have both certified and uncertified prices available when the certifying equilibrium formed.

The hazardous environment was designed so that only the certifying equilibrium existed. This single equilibrium design was adopted for two reasons. First the existence of the non-certifying equilibrium environment for intermediate distributions of good and conditional sellers depended critically on the distribution of loss aversion in the buyer population. Given
that I could not directly control the loss preferences of individuals and maintain a random sample, I elected to study the polar cases where the existence of equilibria were clear. Second, the experiments run in this paper were meant as a baseline for future research into institution adoption. As such, I started with the treatments that I expected to be the most stable so that future experimental work had a baseline on which to compare results.

The safe environment was designed so that under full information about the distribution of types, the certifying equilibrium was extremely unlikely to form or persist. Under full information, if a single type-\(G\) sellers switched to the uncertified market, a loss neutral buyer who knew the proportion of agents in each market would be willing to pay \(0.5U^H + 0.5U^L\) for an uncertified good and \(U^H\) for an uncertified good. Since \(U^H - U^L\) was 100 points across all units, the difference in willingness to pay for a certified and an uncertified unit was \(0.5(U^H - U^L) = 50\). This difference was less than the certification cost of 60 points. Thus under full information, a paired deviation from the certifying equilibrium by a seller and risk neutral buyer could eliminate the certifying equilibrium.

If all buyers were loss neutral, the non-certifying and separating equilibrium under the safe treatment were as follows:

- **Non-Certifying Equilibria for Safe Environment**: \(P^{NC} = 183\). Type-\(G\) sellers produce uncertified \(H\) quality units for a surplus of 153 points per unit. Type \(B\) sellers produce uncertified \(L\) quality units for a surplus of 133 per unit.

- **Certifying Equilibrium for Safe Environment**: \(P^C = 200, P^{NC} = 100\). Type \(G\) sellers sell certified \(H\) quality units for a surplus of 133 per unit. Type \(B\) sellers produce uncertified \(L\) quality units for a surplus of 50 per unit.

Prices in these equilibria were determined by the valuation for the twelfth unit traded. Under loss and risk neutral preferences, this corresponded to the marginal valuation of Unit 3 for a buyer with the higher set of valuations.

Comparing the two equilibria, type-\(G\) sellers received a surplus of 153 points in the non-certifying equilibrium versus 133 points in the separating equilibrium. The type-\(B\) seller

---

<table>
<thead>
<tr>
<th>Table 4: Moral Hazard Environments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good</td>
</tr>
<tr>
<td>Safe ((S))</td>
</tr>
<tr>
<td>Hazardous ((H))</td>
</tr>
</tbody>
</table>

---
received a surplus of 133 points in the non-certifying equilibrium versus 50 points in the certifying equilibrium. All sellers were thus better off in the non-certifying equilibrium and had group incentives to coordinate to this equilibrium.

In the safe environment, if a single type-\(G\) seller traded both of her units in the uncertified market, the proportion of high quality units in the uncertified market would have been 1/2. If the proportion of high and low quality units in the uncertified market were known, a loss neutral buyer who knew this proportion would be willing to pay 150 for an uncertified unit versus 200 for a certified unit. Since the difference in price between the two markets was less than the certification cost of 60, switching to the uncertified market would lead to greater profit for all sellers.

As the statements above indicate, the non-certifying equilibrium was more efficient than the certifying equilibrium under the safe environment. If the quality of both goods was known without certification, the globally efficient outcome was for all seller types to produce the high quality good. In this perfect world, the total efficiency of the safe environment would be \(\Sigma U^H_i - \Sigma C^H_j = 2100\). Under the non-certifying equilibrium, the type-\(B\) seller was expected to produce two low quality units rather than high quality units leading to an expected efficiency of 2060. Similarly, under the certifying equilibrium, the total expected efficiency in the market was 1460.

![Table 5: Efficiency](image)

<table>
<thead>
<tr>
<th></th>
<th>Perfect Information</th>
<th>non-certifying</th>
<th>Certifying</th>
</tr>
</thead>
<tbody>
<tr>
<td>Safe</td>
<td>2100*</td>
<td>2060</td>
<td>1460</td>
</tr>
<tr>
<td>Hazardous</td>
<td>1700*</td>
<td>1100*</td>
<td>1060</td>
</tr>
</tbody>
</table>

*not supportable as an equilibrium

Ex post information about the quality of uncertified trades may provide new information about the underlying environment to buyers and sellers in a market where a partial-certifying equilibrium forms. To test this idea, I introduced an information treatment where subjects were informed about the number of high and low quality units traded in the uncertified market in the previous period. This data was used in conjunction with beliefs data from the bonus game to study the informational properties of the non-certifying equilibrium. Information was given ex post rather than during the trading period to eliminate intra period strategic waiting that would make the trading patterns of the private and public information treatments different.
4.5 Risk and Loss Aversion

After all 24 periods of the main experiment, subjects participated in a risk aversion game and an exit survey. In the risk aversion game, subjects made a series of decisions between a guaranteed return of 90 points and a 50-50 gamble between earning 0 and \( x \), where \( x \) varied between 90 and 300 in increments of 30. I recorded the number of decisions in which the agent chose the safe gamble and used this as a proxy for risk preferences. A high risk score corresponded to more risk aversion.

After the risk experiment, subjects were asked to fill out a survey prior to payment. In the buyers version of the survey, buyers were asked:

- How did you decide on the price you were willing to pay for an uncertified good?

53% of buyers responded to this question by saying that they were unwilling to take losses on an uncertified trade. As a measure of loss aversion, I coded a binary variable which was one if a subject wrote that they were “unwilling to take losses” and zero otherwise. I included in the loss averse group two subjects who reported they were unwilling to make an uncertified trade that was guaranteed a profit of at least 5.

While it is somewhat unsatisfying to use survey data for a measure of loss aversion, there are a number of reasons to believe that the measure is capturing heterogeneity of loss preferences. First, the survey question itself did not prompt subjects to talk about losses. Given that all the responses coded for loss aversion explicitly talk of an unwillingness to take a losses, it is likely that this is an issue of loss aversion rather than risk aversion. Second, as will be shown in section 5.1.2, there is a clear distinction in the riskiness of trades carried out by those who self report loss aversion than to those who do not.

4.6 Experimental Protocol

Subjects for this experiment were volunteers recruited through a database maintained at the Institute for Empirical Research in Economics. The subject pool was primarily comprised of undergraduate students from the University of Zurich and UTH-Zurich. Given the complexity of the experiment, I actively recruited subjects who were experienced in participating in laboratory economic experiments. I excluded economic and psychology students from participating and only permitted each subject to participate in one experimental session.

Subjects were recruited in groups of 33 and randomly assigned to the role of buyer or seller in one of three simultaneous experimental sessions. Subjects were given extensive written
and oral instructions which explained the trading interface, valuations, market procedures, bonus game, and payments. All subjects took a small quiz to ensure they understood how trading worked and how their profits for trading in the certified and uncertified market were calculated.

Once all subjects had finished reading the written instructions, a verbal summary of the experiment was read aloud and subjects began a short computer program that allowed them to become accustomed to all the components of the computer interface. Each seller and buyer was given the chance to post a bid or offer, accept a bid or offer, and play the bonus game. No prices were presented in the computer program to prevent spillover from the practice market to the real market.

Subjects next played all 24 periods of the main experiment. A period of the main experiment consisted of the market game, the bonus game, and an information screens that varied between the information treatments. In the private information treatment, individuals were given a summary of their trades at the close of each trading period prior to playing the bonus game. In the public information session, this summary screen also included the total number of uncertified low, uncertified high, certified low, and certified high trades from the previous period.

At the conclusion of the market game, subjects participated in the risk aversion game and the final survey. I randomly selected 6 of the periods for payment at the end of the main experiment. Including the final questionnaire an experimental session lasted between 2 and 2.5 hours and paid an average of 45 Swiss Francs ($38). All programs for this experiment were written in Z-Tree.

7 Sessions were run in Zurich at the Institute for Empirical Research in Economics and conducted in Swiss German.

5 Experimental Results

Analysis of the experiment is divided into three parts: Initial convergence, adaptation, and welfare. Individual session level data as well as some preliminary findings on belief formation and information is located in the appendix.

1. Initial Convergence: In section 5.1 I compare the equilibrium that forms under the safe and hazardous environments absent a preexisting market organization. I find that in the hazardous environment the certifying equilibrium forms and under the

7 See Fischbacher (2007) for a description of Z-Tree.
safe environment a non-certifying or partial-non-certifying equilibrium forms. I then study how aggregate levels of loss aversion influence the non-certifying price and the formation of the partial certifying equilibrium.

2. Adaptation: In section 5.2, I study how the non-certifying and certifying equilibrium adapt to changes in the environment. I find that the non-certifying equilibrium adapts to the certifying equilibrium when the environment changes from safe to hazardous while the certifying equilibrium persists when the environment changes from hazardous to safe. Looking at the dynamics of the market, I find that when a market is non-certifying initially and the environment changes from safe to hazardous, the composition of uncertified units changes leading to a gradual change in price and ultimately a shift from uncertified to certified trades. In markets that reach the certifying equilibrium, there is no observable difference in the composition of trades associated with a change in the environment.

3. Welfare In section 5.3, I compare welfare outcomes from subjects who were in treatments 1 & 2 with those in treatments 3 & 4. I find a strong welfare loss in treatments 3 & 4 due to the persistence of the certifying equilibrium in the safe environment.

As noted in section 4.4, experimental sessions were divided into four groups based on the ordering of the $S$ and $H$ environments and the amount of information about past trades. I designate trades in the first 12 periods with a $Pre$ superscript and trades in the last 12 periods with a $Post$ superscript. For reference, a copy of Table 3 is reprinted here.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Periods 1-12</th>
<th>Periods 13-24</th>
<th>Information</th>
<th>Identifiers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Safe</td>
<td>Hazardous</td>
<td>Private</td>
<td>$S^{Pre}, H^{Post}$</td>
</tr>
<tr>
<td>2</td>
<td>Safe</td>
<td>Hazardous</td>
<td>Public</td>
<td>$S^{Pre}, H^{Post}$</td>
</tr>
<tr>
<td>3</td>
<td>Hazardous</td>
<td>Safe</td>
<td>Private</td>
<td>$H^{Pre}, S^{Post}$</td>
</tr>
<tr>
<td>4</td>
<td>Hazardous</td>
<td>Safe</td>
<td>Public</td>
<td>$H^{Pre}, S^{Post}$</td>
</tr>
</tbody>
</table>

5.1 Initial Convergence

5.1.1 Hypothesis and Empirical Strategy

In this section, I compare the equilibria that forms under the safe and hazardous environments absent a preexisting market organization. I establish that in the hazardous envi-
ronment the certifying equilibrium forms while under the safe environment a non-certifying or partial-certifying equilibrium forms. I then study how aggregate levels of loss aversion influence the non-certifying price and the formation of the partial-certifying equilibrium.

Based on the theoretical model, the relative welfare and probability of existence of the non-certifying equilibrium is increasing in the proportion of type-\(G\) sellers in the environment. It follows:

**Hypothesis 1** In markets with the possibility of certification, the likelihood of the non-certifying equilibrium forming in a new market is increasing in the proportion of agents who always produce high quality units.

I test this hypothesis by comparing the prices of uncertified trades in the \(S^\text{Pre}\) environment where the degree of moral hazard is low with those in the \(H^\text{Pre}\) environment where the degree of moral hazard is high. To allow time for the market to converge, I restrict attention to periods 7-12.\(^8\) Using session fixed effects, I estimate:\(^9\)

\[
P_{i,s} = \alpha_0 + \Sigma \alpha_s + \beta \text{Cert} I_{\text{Cert}} + \beta S^\text{Pre} I_{S^\text{Pre}} + \epsilon_{i,s}
\]

where \(P_{i,s}\) is the price of an individual trade \(i\) in session \(s\), \(\alpha_s\) are individual session fixed effects, \(I_{\text{Cert}}\) is an indicator for a certified trade, and \(I_{S^\text{Pre}}\) is an indicator variable for uncertified trades in the safe environment.

In markets where the certifying equilibrium forms, the equilibrium prices for certified and uncertified units is 200 and 100. In the \(S^\text{Pre}\) environment, if the non-certifying equilibrium forms, the expected non-certifying price with no loss aversion is 183. Expecting the non-certifying equilibrium to form in the \(S^\text{Pre}\) environment and the certifying equilibrium to form in the \(H^\text{Pre}\) environment, I predict \(\alpha_0 = 100, \alpha_0 + \beta \text{Cert} = 200\) and \(\alpha_0 + \beta S^\text{Pre} = 183\).

One possibly reason for a deviation from our non-certifying prediction price of 183 is loss aversion. As noted in the theory section, the loss aversion of the last buyer who is willing to

\(^8\)The number of omitted periods was decided prior to running the experiment and based on the initial pilots. As can be seen in the individual experiments included in the appendix, the price of the uncertified market converges to the non-certifying or partial-certifying equilibrium from below. Thus, increasing the number of periods in the analysis decreases the estimated uncertified price for treatments that converge to the non-certifying equilibrium. All results are still statistically significant in the full sample with attenuated magnitudes on \(\beta S^\text{Pre}\).

\(^9\)Note that since loss aversion only affects uncertified trades and the estimation includes both certified and uncertified trades, session level fixed effects do not eliminate the variation in uncertified trades across treatments.
trade in the uncertified market impacts the market price of a non-certifying equilibrium. In order to get a simple aggregate measure of loss aversion, I take the total number of buyers in each session who self report loss aversion. Interacting this number with the safe treatment (the only treatment where the non-certifying equilibrium is likely) I estimate:

\[ P_{i,s} = \alpha_0 + \sum \alpha_s + \beta_{LA}(LA \times IS_{Pre}) + \beta_{Cert}I_{Cert} + \beta_{SP_{Pre}}IS_{Pre} + \epsilon_{i,s} \]

Where \( LA \) is the total number of buyers in a session who report that they were unwilling to make a trade if they had the potential for taking a loss.

5.1.2 Results

Table 6 presents the price regressions from equation 18 and 19 with varying degrees of control for loss aversion. As can be seen in column (1), when loss aversion is not taken into account, the predicted non-certifying price \((\alpha_0 + \beta_{SP_{Pre}} = 147)\) is lower than our predicted value of 183 but above the minimum price that could sustain a partial-certifying equilibrium. When an aggregate measure of loss aversion is used, as in column (3), the predicted non-certifying price for a market with no loss aversion is 184, remarkably close to our theoretical prediction. The estimated price for uncertified and certified units in the \( \mathcal{H}_{Pre} \) environment are 105 and 196, both close to their predicted values of 100 and 200.

In order to confirm that my measure of loss aversion is indeed capturing heterogeneity of preference across agents, I look at the purchase history of buyers conditional of their measured loss aversion. The left hand side of Table 7 compares the number of certified and uncertified trades made by loss averse buyers in the certified and uncertified markets. As expected, loss averse individuals are more likely to make safe transactions.

Some trades in the uncertified market carry no risk for the buyer. A unit purchased at price 100, for instance, has no possibility of loss for a buyer who has a value of 140 for a low quality unit. As a better measure of risk, I partition trades into those that there are “safe” and have no possibility for a loss and those that are “risky” and may result in a loss if a low quality unit is transacted. As shown on the right hand side of Table 7, loss averse individuals are significantly more likely to make safe transactions.
Table 6: Hypothesis 1: Convergence of Pre Treatments to the Non-Certifying or Certifying Equilibrium

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Certification</td>
<td>91.414***</td>
<td>91.414***</td>
<td>91.414***</td>
</tr>
<tr>
<td>Treatment $I_{Pre}$</td>
<td>39.100***</td>
<td>60.867***</td>
<td>79.229***</td>
</tr>
<tr>
<td>Risk Aversion of Buyer in $I_{Pre}$</td>
<td>-2.270*</td>
<td>-0.090</td>
<td>(1.096)</td>
</tr>
<tr>
<td>Individual Loss Aversion in $I_{Pre}$</td>
<td>-20.827***</td>
<td>(6.097)</td>
<td></td>
</tr>
<tr>
<td>Number of Loss Averse Buyers in $I_{Pre}$</td>
<td>-16.216***</td>
<td>(2.113)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>108.496***</td>
<td>106.630***</td>
<td>105.349***</td>
</tr>
<tr>
<td>Fixed Effects$^b$</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.838</td>
<td>0.849</td>
<td>0.856</td>
</tr>
<tr>
<td>Observations (Trades in Period 7-12)</td>
<td>834</td>
<td>834</td>
<td>834</td>
</tr>
</tbody>
</table>

$^a$Since loss aversion is an aggregate measure in specification (3), the standard error from the trade level regression is improperly calculated. As a better measure, I use randomization inference to construct a confidence interval. I begin by estimating the session level regression $AvgP_s = \alpha_0 + \beta_{LA}(LA_s)$. I then take every permutation of possible loss aversion assignments to construct placebo estimates of the loss aversion parameter. This generates a distribution of possible loss aversion parameters centered at zero. The true estimated value of $\beta_{LA}$ lies outside the 95% confidence of this placebo distribution.

$^b$Fixed Effects are at the session level. Standard errors in parenthesis. Errors clustered by individual buyer.

Table 7: Transaction History Conditional on Loss Aversion

<table>
<thead>
<tr>
<th></th>
<th>Uncertified</th>
<th>Certified</th>
<th>Risky</th>
<th>Safe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss Averse</td>
<td>84</td>
<td>120</td>
<td>51</td>
<td>153</td>
</tr>
<tr>
<td>Normal</td>
<td>134</td>
<td>85</td>
<td>128</td>
<td>92</td>
</tr>
<tr>
<td>Fisher Exact Test $p &lt; 0.01$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whitney-Mann-Wilcoxon $p = 0.2211$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$One problem with using a Fisher Exact Distribution test is that each individual accounts for multiple observations. In order to more accurately account for this multiplicity, I create two new continuous variables which measure the proportion of uncertified trades and the proportion of safe trades respectively. I then use a two-sample Whitney-Mann-Wilcoxon rank-sum test to test for differences in the distributions of the proportions.
5.2 Adaptation

5.2.1 Hypothesis and Empirical Strategy

I next look at how the equilibrium that formed in the initial 12 periods adapts to changes in the underlying environment. In the theoretical model, I showed that when the certifying equilibrium is reached, there is no aggregate information observable when type-C sellers are replaced with type-G sellers and thus the equilibrium does not change. By contrast, I showed that when the non-certifying equilibrium is reached, a replacement of type-G with type-C sellers leads to a reduction in the uncertified price and an eventual change to the certifying equilibrium. This leads to:

Hypothesis 2 Any market equilibrium that reaches the certifying equilibrium will remain certifying for any changes in the number of type-C and type-G sellers.

I test this hypothesis by comparing the price of uncertified trades that occur in the last six trading periods of each treatment. If there is no aggregate information observable when the environment changes from hazardous to safe, equilibrium prices in periods under the $S_{Post}$ treatment should be the same as those from $H_{Pre}$ and significantly differ from those in $S_{Pre}$. I thus estimate:

\[
P_{i,s} = \alpha_0 + \sum \alpha_s + \beta_{LA}(LA * I_{SPre}) + \beta_{Cert}I_{Cert} \\
+ \beta_{SPre}I_{SPre} + \beta_{SPost}I_{SPost} + \beta_{HPost}I_{HPost} + \epsilon_{i,s},
\]

where $P_{i,s}$ is the price of an individual trade $i$ in session $s$, $\alpha_s$ are individual session fixed effects, $I_{Cert}$ is an indicator for a certified trade, and $I_{SPre}$, $I_{SPost}$, and $I_{HPost}$ are indicator variables for uncertified trades in their respective environment. I hypothesize that $\alpha_0 + \beta_{SPre} = 183$, and $\beta_{SPost} = \beta_{HPost} = 0$.

An informational theory of persistence predicts that markets in a certifying equilibrium should show no observable change when a type-C seller is replaced by a type-G seller. I next look at the composition of trades over time in the uncertified market and between the certified and uncertified market.

In treatments where the certifying market forms initially, there should be no observable difference in the composition of trades before and after the change in environment. I hypothesize:
Hypothesis 3  *In a certifying equilibrium, the composition of goods within a market and between the certified and uncertified markets is unaffected by a replacement of type-C sellers with type-G sellers.*

To study this, I compare the composition of trades in the six periods before and after the change in environment. In the treatments that start in the $H^{Pre}$ environment (where the certifying equilibrium was consistently observed), I expect to see no change in the composition of trades when the environment changes to $S^{Post}$.

As a comparison, I also look at the composition of trades within the uncertified market and between the certified and uncertified markets when the non-certifying equilibrium forms initially and type-G sellers are replaced by type-C sellers. When sellers change from type $G$ to type $C$ in the non-certifying equilibrium, their initial incentive is to sell low quality units in the uncertified market as long as the difference in prices is greater than 90 (the certification cost + the difference in production costs for a type-C seller). In transition from a non-certifying to a certifying equilibrium, the expected dynamics of the market are thus an initial shift from uncertified high quality units to uncertified low quality units followed by an eventual shift from the uncertified market to the certified market.

Hypothesis 4  *In a non-certifying equilibrium, the replacement of type-G sellers with type-C sellers leads to an immediate shift from uncertified high quality units to uncertified low quality units followed by a transition to the certifying equilibrium.*

In the treatments that start in the $S^{Pre}$ environment (where the non-certifying equilibrium is consistently observed), I expect to see a shift to uncertified low quality units followed by a gradual shift to certified units when the environment changes to $H^{Post}$.

5.2.2 Results

The persistence of the certifying equilibrium is most easily seen by comparing an individual session in treatments 1 & 2 with an individual session in treatments 3 & 4. Figure 1 makes this comparison, showing the complete trade history of session 6 and session 12. The horizontal dashed lines show the predicted price of the certified and uncertified market in the case of the non-certifying equilibrium for the $S^{Pre}$ environment and the certifying equilibrium in the case of the other three environments. The vertical dashed lines split trades into 6 period increments with the aggregate number of certified and uncertified trades reported at the bottom of each block.
As can be seen in the top half of figure 1, a session in treatment 1 & 2 where the environment is initially set to safe converges to the partial-certifying equilibrium in the first 12 periods and then adapts to the certifying equilibrium when the environment changes. Note that in the safe environment, there is always a single type-$B$ seller and thus the predicted composition of units without loss aversion is 60 uncertified high quality units and 12 uncertified low quality units. Typical of all sessions in treatments 1 & 2, convergence of the uncertified price to a partial certifying equilibrium is from below with a subset of certified trades conducted in each period at a premium 60 points above the prevailing uncertified market price. When the environment changes, sellers who switched from type $G$ to type $C$ sell low quality units leading to a fall in price and the eventual establishment of a certifying equilibrium.

In the session that began in the hazardous environment, the certifying equilibrium is established in the first 12 periods. When the environment switches to safe at period 13, there is no noticeable change in the uncertified price nor in the composition of certified and uncertified trades. This is the case in the bottom half of figure 1 where convergence to the certifying equilibrium is rapid and the convergence of the uncertified price is from above.

While the two sessions used in figure 1 were chosen for clarity, the patterns of adaption and persistence are typical across all sessions. Figure 2 shows average uncertified prices for the last six periods of each environment. Notice that the uncertified price in the $S^{Post}$ environment is nearly identical to both the $H^{Pre}$ and $H^{Post}$ treatments and markedly different to the $S^{Pre}$ treatment.

Again using a price regression, I extend the initial regression to include the last six periods of both environments in all 12 experimental sessions. In support of hypothesis 2, I find that there is no significant difference between the uncertified prices in the $S^{Post}$ and $H^{Post}$ environments relative to the baseline environment of $H^{Pre}$.

Turning attention to market dynamics, I begin by looking at the composition of trades over time. Based on the theoretical model, in sessions where the market is initially non-certifying and the degree of moral hazard is increased, I expect to see an initial shift of units from uncertified high quality units to uncertified low quality units followed by a gradual transition to certified trades as the uncertified market price falls. In sessions where the certifying equilibrium has formed, an information based story of adaptation would predict no change in the composition of goods when moral hazard is decreased.

---

10Summary statistics of the other treatments are available in the appendix. Additional data is available upon request.
**Session 6:** Formation of the Pooling Equilibrium and Adaptation to the Certifying Equilibrium

**Session 12:** Formation and Persistence of the Certifying Equilibrium

Figure 1: Hypothesis 2 — Persistence of the Certifying Equilibrium
Figure 2: Average Uncertified Prices by Environment

Table 8: Hypothesis 2: Persistence of the Certifying Equilibrium

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Certification</td>
<td>89.003***</td>
<td>89.003***</td>
</tr>
<tr>
<td></td>
<td>(2.878)</td>
<td>(2.879)</td>
</tr>
<tr>
<td>Treatment $S^{Pre}$</td>
<td>36.610***</td>
<td>75.948***</td>
</tr>
<tr>
<td></td>
<td>(4.698)</td>
<td>(7.196)</td>
</tr>
<tr>
<td>Treatment $S^{Post}$</td>
<td>2.776</td>
<td>2.776</td>
</tr>
<tr>
<td></td>
<td>(3.721)</td>
<td>(3.722)</td>
</tr>
<tr>
<td>Treatment $H^{Post}$</td>
<td>3.620</td>
<td>3.867</td>
</tr>
<tr>
<td></td>
<td>(3.592)</td>
<td>(3.503)</td>
</tr>
<tr>
<td># of Loss Averse Buyers in $S^{Pre}$</td>
<td>-16.207***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.528)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>112.190***</td>
<td>110.616***</td>
</tr>
<tr>
<td></td>
<td>(3.369)</td>
<td>(3.209)</td>
</tr>
<tr>
<td>Fixed Effects$^a$</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.861</td>
<td>0.873</td>
</tr>
<tr>
<td>Observations</td>
<td>1675</td>
<td>1675</td>
</tr>
</tbody>
</table>

$^a$Fixed Effects are at the session level. Standard Errors in parenthesis. Errors are clustered at the individual level.
Table 9 shows the aggregate number of trades in Treatments 1 & 2 for the last 18 periods of the experiment split into 6 period increments. As seen in this aggregate data, there is a strong shift in the composition of units in the uncertified market in response to the change in environment followed by a gradual shift from uncertified to certified markets between periods 13-18 and 19-24. In comparison, table 10 shows the same aggregate data for Treatments 3 & 4 where the safe and hazardous environments are reversed. In these sessions, there is very little change in the composition of trades across time.

Table 9: Aggregate Trades in Treatments 1 & 2 (Sessions 1-6)

<table>
<thead>
<tr>
<th></th>
<th>Uncertified Red</th>
<th>Uncertified Blue</th>
<th>Certified Red</th>
</tr>
</thead>
<tbody>
<tr>
<td>Last 6 periods of $S^{Pre}$</td>
<td>135</td>
<td>95</td>
<td>208</td>
</tr>
<tr>
<td>First 6 periods of $H^{Post}$</td>
<td>15</td>
<td>206</td>
<td>177</td>
</tr>
<tr>
<td>Last 6 periods of $H^{Post}$</td>
<td>2</td>
<td>156</td>
<td>254</td>
</tr>
</tbody>
</table>

Table 10: Aggregate Trades in Treatments 3 & 4 (Sessions 7-12)

<table>
<thead>
<tr>
<th></th>
<th>Uncertified Red</th>
<th>Uncertified Blue</th>
<th>Certified Red</th>
</tr>
</thead>
<tbody>
<tr>
<td>Last 6 periods of $H^{Pre}$</td>
<td>6</td>
<td>150</td>
<td>255</td>
</tr>
<tr>
<td>First 6 periods of $S^{Post}$</td>
<td>23</td>
<td>95</td>
<td>311</td>
</tr>
<tr>
<td>Last 6 periods of $S^{Post}$</td>
<td>21</td>
<td>89</td>
<td>319</td>
</tr>
</tbody>
</table>

As more precise support for hypotheses 3 and 4, I break the composition of trades down to the period level. As with the aggregated data, the period level data shown in figures 3 and 4 show significant compositional change in the non-certifying equilibrium and no change in the certifying equilibrium.

Apparent in Figure 3, the change in environment from safe to hazardous results in an immediate shift from uncertified high quality units to uncertified low quality units. There is also a small but consistent shift of transactions from certified high quality units to uncertified low quality units in the two periods following the change in treatment. Recall that in the non-certifying equilibrium with loss aversion, it may be the case that the type-$G$ sellers are indifferent between trading certified and uncertified units while type-$C$ sellers strictly prefer to sell uncertified units. Given a replacement of type-$G$ sellers with type-$C$ sellers, there is an increase in incentives of sellers to sell in the uncertified market. This effect may increase the speed of adaptation from non-certifying to certifying by increasing the number of uncertified low quality observations that occur after the change in moral hazard.

Whereas a replacement of type-$G$ sellers with type-$C$ sellers increases the number of
uncertified trades, replacing type-\(C\) sellers with type-\(G\) sellers, as done in treatments 3 & 4, may reduce incentives and diminish uncertified trade. Notice that in figure 4, there is a slight shift away from the uncertified market when the environment improves. This again is most likely a result of weaker incentives for type-\(G\) sellers to trade uncertified goods than those of type \(C\). Unlike treatments 1 & 2 where this effect improved adaptation, here the change in environment reduces experimentation and increases the likelihood that the separating equilibrium persists.

5.3 Welfare

5.3.1 Hypothesis and Empirical Strategy

In the theory section of the paper, I showed that the relative earnings of the non-certifying and certifying equilibria are conditional on the change in quality of the conditional sellers.
and the proportion of sellers who were of type $G$ and type $C$. It follows

**Hypothesis 5** *In environments where the proportion of conditional sellers is small, the certifying equilibrium will be less efficient than the non-certifying equilibrium.*

To test this statement, I look at the overall efficiency of the last six periods of the $S^{Pre}$ environment and compare it to the efficiency of the last 6 periods in the $S^{Post}$ environment. I estimate

$$\text{Efficiency}_{t,s} = \alpha_0 + \beta_{S^{Pre}} I_{S^{Pre}} + \beta_{LA}(LA \times I_{S^{Pre}}) + \beta_{LA}(LA \times I_{S^{Post}}) + \epsilon_{t,s}$$

I predict that periods in the $S^{Pre}$ period will have a higher overall efficiency than those in the $S^{Post}$ treatment with the efficiency gains decreasing in the number of buyers who are loss averse and separate from the uncertified market.

### 5.3.2 Results

In figure 5, I compare period by period efficiency of the $S^{Pre}$ periods with those of the $S^{Post}$ treatment. The dashed horizontal line in the graph shows the predicted efficiency of a pure non-certifying equilibrium for the $S^{Pre}$ treatment and the certifying equilibrium in the $S^{Post}$ treatment. On the left hand side of the figure, it can be seen that the overall efficiency of the partial-certifying equilibrium is significantly below what is predicted in the pure non-certifying equilibrium case. This decrease in efficiency is a result of three factors: the exit of loss averse buyers from the uncertified market, missed trade opportunities that often occurred in the non-certifying equilibrium as buyers and sellers negotiated trades, and the adoption of certification in one of the treatments that began with $S^{Pre}$. On the right hand side of the figure, it can be seen that all 6 treatments have consistent efficiency levels in line with the predictions of the certifying equilibrium.

Unsurprising given the visible difference in efficiency, the price regression in table 11 shows a significant increase in efficiency in the $S^{Pre}$ environment relative to $S^{Post}$. The overall efficiency of the non-certifying equilibrium is lower than the theoretical prediction, however, suggesting additional inefficiencies that are created from the dynamic process that are not accounted for by loss aversion alone.
Figure 5: Hypothesis 5 — Efficiency Loss due to the Persistence of Certification

### Table 11: Hypothesis 5: Efficiency Loss due to the Persistence of Certification

<table>
<thead>
<tr>
<th>Treatment $S^{Pre}$</th>
<th>330.010**</th>
</tr>
</thead>
<tbody>
<tr>
<td># Loss Averse Buyers in $S^{Pre}$</td>
<td>$-64.242$</td>
</tr>
<tr>
<td>(122.232)(a)</td>
<td>(50.930)</td>
</tr>
<tr>
<td># of Loss Averse Buyers in $S^{Post}$</td>
<td>$-2.073$</td>
</tr>
<tr>
<td>(3.568)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>1456.707***</td>
</tr>
<tr>
<td>(13.634)</td>
<td></td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.439</td>
</tr>
<tr>
<td>Observations</td>
<td>72</td>
</tr>
</tbody>
</table>

\(a\)Errors are clustered at the session level
6 Conclusion

This paper represents a first step in understanding how the informational properties of institutions may lead to their inefficient persistence. I showed formally that, in a market with endogenously formed certification institutions, observable information about changes in the underlying environment could be lost. This lost information could lead to the persistence of a certifying equilibrium where all participants in the environment are weakly worse off relative to a world without the certification institution. The experimental evidence of inefficient persistence of the certifying equilibrium was striking. No session that initially adopted the certification institution showed observable changes in price or the distribution of trades in response to a change in the underlying distribution of seller types. This led to a loss of efficiency relative to a market with the same underlying environment but where the certifying equilibrium had not initially formed.

The experiments described in this paper constitute a stable baseline on which to guide future theoretical and experimental work. I showed that in a double auction environment where trades were centralized and buyers and sellers were anonymous, the benchmark model performed extremely well in predicting both initial convergence and adaptation. I further demonstrated that for some initial distribution of seller types, both the non-certifying and certifying equilibrium were stable. Building on the consistency of these initial experiments, future research will focus on the types of information necessary to adapt away from the certifying equilibrium and on the dynamic learning processes that generate persistence.

The information externality highlighted in this paper represents a general phenomenon that extends beyond the simple certification market considered here. Common mechanisms designed to mitigate moral hazard such as red tape, regulation, certification, monitoring, process management, and credit scoring all share the common characteristic that they group heterogenous agents into the same action. Given the importance of these institutions in everyday markets and organizations, developing an understanding of how information externalities dynamically alter the institutional landscape is of great importance.
7 Appendix

7.1 Appendix 1: The Partial Certifying Equilibrium

In experimental sessions beginning in the safe environment, a partial-certifying equilibrium often formed in which the difference in prices between the certified and uncertified market was equal to the certification cost. This section looks at two possible forces that might contribute to this partial certifying equilibrium: loss aversion and incomplete learning by buyers.

Appendix 1 is divided into three parts. In section 7.1.1, I explicitly characterize the competitive equilibrium with loss aversion and a common prior about the distribution of seller types. I show that loss aversion can lead to a downward sloping demand curve that can potentially support a partial-certifying equilibrium. In section 7.1.2, I develop an alternative game theory model where I relax the common prior assumption in the benchmark model to study learning. I show that in the partial-certifying and certifying equilibrium, the beliefs of buyers need not converge to the true distribution of types leading to persistence of the partial-certifying equilibrium due to incomplete learning. Finally, in section 7.1.3, I empirically distinguish between these two effects using the information treatment as exogenous variation. I find that loss neutral buyers respond to the information treatment with more trade in the uncertified market while loss averse buyers do not. These results are consistent with a model where loss aversion generates a partial-certifying equilibrium and other buyers, due to limited market signals, do not have beliefs that converge to the true distribution of types.

7.1.1 Loss Aversion

Suppose that instead of all buyers being the same, some buyers are loss averse and put a greater weight on aggregate losses than gains. Let \( B = \{\lambda_1, \lambda_2, \ldots, \lambda_N\} \) where \( \lambda_i \) is the idiosyncratic loss aversion parameter for buyer \( i \) with \( \lambda_i \geq 1 \) for \( i \in \{1, 2, \ldots, N\} \). Without loss of generality, I order buyers according to their risk aversion parameter such that \( \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_N \) and again normalize the utility obtained from not trading to zero.

For a price system \( P \) with \( P_{NC} \leq P_{NC} \leq P_C \leq P_C \), a buyer \( i \) buying from market \( m \) at price \( P^m \) from a seller of type \( s \) gets utility

\[
\begin{align*}
    u(m, P^m, \lambda, s) &= \begin{cases} 
        U^H - P^C & \text{if } m \in C, s \in \{G, C, B\} \\
        U^H - P_{NC} & \text{if } m \in NC, s \in \{G\} \\
        -\lambda_i (P_{NC} - U^L) & \text{if } m \in NC, s \in \{C, B\} 
    \end{cases}
\end{align*}
\]

(22)

A loss-averse buyer prefers to purchase in the uncertified market any time that:

\[
U^H - P^C \leq \pi^H(\Delta P)(U^H - P_{NC}) + (1 - \pi^H(\Delta P))\lambda_i(U^L - P_{NC}).
\]

Rewriting this condition in terms of the uncertified market price, a buyer prefers to purchase
in the uncertified market as long as:

\[ P^{NC} \leq \frac{P^C - U^H}{\pi^H(\Delta P) + (1 - \pi^H(\Delta P))\lambda_i} + \frac{\pi^H(\Delta P)U^H + (1 - \pi^H(\Delta P))\lambda_iU^L}{\pi^H(\Delta P) + (1 - \pi^H(\Delta P))\lambda_i}. \]

Note that with a surplus of buyers, there is always indifference between buying and not buying. Since the certified market has no risk, this means that \( P^C = U^H \). Thus the first term on the right side of the inequality will be zero and (23) reduces to

\[ P^{NC} \leq \frac{\pi^H(\Delta P)U^H + (1 - \pi^H(\Delta P))\lambda_iU^L}{\pi^H(\Delta P) + (1 - \pi^H(\Delta P))\lambda_i}. \]

In equilibrium:

- If \( \Delta P > T \) the buyer believes that all type-\( G \) sellers will certify their goods and thus that \( \pi^H(\Delta P) = 0 \). In this case, a buyer prefers to purchase the certified unit as long as \( \Delta P < U^H - U^L \equiv \bar{T}^C - P^{NC} \) and is indifferent between buying a non-certified unit and not purchasing if \( P^{NC} = U^L \).

- If \( \Delta P < T \) the buyer believes that all agents trade in the uncertified market. In this case \( \pi^H(\Delta P) = g \). A loss-averse buyer prefers to purchase the uncertified unit as long as inequality (24) is satisfied.

- If \( \Delta P = T \), the buyer believes that all sellers that trade in the certified market are type \( G \). Given that \( \Sigma_sS(C, s) \) sellers trade in the certified market, \( \pi^H(\Delta P) = \frac{M_G - \Sigma_sS(C, s)}{M - \Sigma_sS(C, s)} \) and a loss-averse buyer prefers to purchase the uncertified unit as long as inequality (24) is satisfied.

The demand decision of an individual buyer \( i \) is directly affected by his own loss aversion parameter \( \lambda_i \) and indirectly affected by the loss aversion of other agents. In the non-certifying equilibrium, the market price \( P^{NC} \) is greater than \( U^L \) and there is a potential for losses in the uncertified market. Since buyers are heterogeneous in loss aversion, the aggregate demand curve for uncertified units becomes downward sloping and the market price is pinned down by the loss aversion of the \( M^{th} \) buyer. If the \( M^{th} \) buyer (the last to trade in equilibrium) is sufficiently loss averse, he may be unwilling to trade for uncertified units at a price where \( \Delta P \geq T \). In this case, the \( M^{th} \) buyer may exit the uncertified market and a partial certifying equilibrium may form.

As before, I characterize a competitive equilibrium where the supply and demand decision of the buyers and sellers are optimal and where both the certified and uncertified markets clear. Let \( \lambda_M \) be the loss aversion parameter for the \( M^{th} \) buyer and define the prices at which he is indifferent between his three options of buying certified units, buying uncertified
units, and not buying as:

\[
\tilde{P}^{\text{NC}}_M = \frac{gU^H + (1 - g)\lambda_M U^L}{g + (1 - g)\lambda_M},
\]

\[
\tilde{P}^C = U^H.
\]

If \(\tilde{P}^C - \tilde{P}^{\text{NC}}_M \geq T\), the \(M\)th highest buyer is unwilling to buy in the uncertified market at the minimal price supporting a pure non-certifying equilibrium. Due to differences in costs across the three seller types, for any price system where a type-\(C\) seller prefers to sell in the uncertified market, all type-\(G\) sellers must also wish to certify. Thus any partial certifying equilibrium has only type-\(G\) sellers in the certified market.

Iteratively, if the \(M\)th buyer exits the uncertified market, the \(M - 1\) remaining buyers face a pool of uncertified goods where one type-\(G\) seller has been removed. Looking at buyer \(M-1\), I again define prices at which this buyer is indifferent across markets, taking into account the higher risk in the uncertified market:

\[
\tilde{P}^{\text{NC}}_{M-1} = \frac{\tilde{g}_{-1}U^H + (1 - \tilde{g}_{-1})\lambda_{M-1} U^L}{\tilde{g}_{-1} + (1 - \tilde{g}_{-1})\lambda_{M-1}},
\]

\[
\tilde{P}^C_{M-1} = U^H,
\]

\[
\tilde{g}_{-1} = \frac{M_G - 1}{M - 1}
\]

If buyer \(M-1\) is unwilling to trade in the uncertified market, I iterate the process again. Define prices at which the buyer with the \(M - k\)th lowest loss aversion is indifferent between all three markets as

\[
\tilde{P}^{\text{NC}}_{M-k} = \frac{\tilde{g}_{-k}U^H + (1 - \tilde{g}_{-k})\lambda_{M-k} U^L}{\tilde{g}_{-k} + (1 - \tilde{g}_{-k})\lambda_{M-k}},
\]

\[
\tilde{P}^C_{M-k} = U^H,
\]

\[
\tilde{g}_{-k} = \frac{M_G - k}{M - k}
\]

A partial certifying equilibrium occurs any time their exists a \(k\) such that \(\tilde{P}^C_{M-k} - \tilde{P}^{\text{NC}}_{M-k} \leq T\) and \(\tilde{P}^C_{M-k+1} - \tilde{P}^{\text{NC}}_{M-k+1} > T\):

**Theorem 4** Let \(k^*\) be the smallest \(k\) such that \(\tilde{P}^C_{M-k} - \tilde{P}^{\text{NC}}_{M-k} \leq T\). Then, if \(k^* \in [1, M_G]\), a partial certifying equilibrium exists where \(P^C = U^H, P^{NC} = U^H - T\).

In a partial certifying equilibrium, buyers \(i = \{1, \ldots, (M - k^*)\}\) buy from the uncertified market. These buyers have strictly positive expected utility since loss aversion by other buyers pin down the uncertified price. \(k^*\) other buyers trade in the certified market. \(k^*\)
type-$G$ buyers sell certified high quality units. The remaining type-$G$ sellers mix with the type-$C$ and type-$B$ sellers and exchange in the uncertified market.

7.1.2 Alternative Game-Theoretic Model

In this section, I relax the common prior assumption of the benchmark model and develop a game-theoretic model where strategic play over time may lead to price and allocation rules similar to the Competitive Equilibria of Sections 3.2 and 7.1.1. I first build a single period game which guarantees \textit{ex post} stability. I then study under what conditions buyers who update their beliefs solely from private signals can learn about the true distribution of types. I show that even under the best conditions where the non-certifying equilibrium is selected any time that it exists, incorrect beliefs by buyers can lead to the partial-certifying or certifying equilibrium where buyers no longer learn.

Game theoretic models with simultaneous action can often support equilibrium based on the fear of being left unmatched when all players reveal their actions. In order to eliminate this type of equilibria from analysis, I take a nonstandard modeling approach in which buyers are allowed to make contingent offers to the market, sellers sequentially and repeatedly decide whether to enter or stay out of the market, and resolution is via a Vickrey auction.

By having the buyers act first and giving them weakly dominant strategies, I eliminate the need to condition strategies on higher order beliefs. This eliminates some of the complications that arise out of a pure signalling model where the sellers are first to act. By allowing bids of buyers to be contingent on the overall size of the market and allowing sequential offers by the sellers, I ensure \textit{ex post} stability in the market. This stability property coupled with assumptions about seller rationality allows buyers to update their beliefs about the distribution of seller types directly from the quality of units that they receive in the uncertified market.

The modeling approach taken in this paper is not meant to be exhaustive of the interesting theoretical questions that arise in this environment. Instead, I take the most direct route to myopic belief updating to underscore how myopic learning in the non-certifying equilibrium differs from the other two equilibria that might form. Interesting questions related to higher order belief formation in signalling games without common knowledge and dynamic learning on the seller side is left for future study.\footnote{I sidestep some of the complications that arise on the seller side of the problem by restricting analysis to the case of sequential action with known valuations for the buyer. More explicit dynamic learning of sellers via fictitious play or reinforced learning opens up the possibility that type-$C$ sellers may certify in some periods before type-$G$ sellers. This eliminates the ability to make inferences about the distribution of types from signals in the partial-certifying equilibrium but does not affect the results from the non-certifying equilibrium.}

**Primitives**

Consider an extension to the benchmark model in which all buyers and sellers are infinitely lived over periods indexed by $t \in \{0, \ldots, \infty\}$. There are $M$ sellers who can each sell one unit
and are divided into three types \( s \in \{ G, C, B \} \). In each period, sellers \( j \in \{1, \ldots, M-1\} \) are randomly assigned to either type-\( G \) or type-\( C \). Seller \( j = M \) is assigned to be of type-\( B \) so that there is always at least one seller in the non-certified market. The true proportion of sellers of type \( G \) in a period is \( \hat{g} \).

There are \( N \) buyers who each demand a single high or low quality unit in each period. Initially, the distribution of seller types is unknown to buyers. Buyers form beliefs about \( \{ M_G, M_C, M_B \} \) based on an initial prior distribution and update this prior between periods based on the quality of unit that they received in the last period, the market they traded in, and the volume of trades in the high and low market. For simplicity, I assume that \( M_B = 1 \) is known to all parties so that the estimation problem can be reduced to estimating a single parameter \( \hat{g} = \frac{M_G}{M} \) where \( \hat{g} \in \{ \hat{g}_0, \hat{g}_1, \ldots, \hat{g}_{M-1} \} \) and \( \hat{g}_k = \frac{k}{M} \).

Each buyer \( i \) has an initial prior \( p_0^i(\hat{g}) = \{ p_0^i(\hat{g}_0), p_0^i(\hat{g}_1), \ldots, p_0^i(\hat{g}_{M-1}) \} \) where \( p_0^i(\hat{g}_k) > 0 \) and \( \sum_k p_0^i(\hat{g}_k) = 1 \). Buyers are strongly myopic in the sense that their beliefs change only with their own history of trades and not on information that is directly observable in the market. Let \( h^i(t) \) be the buyers history at time \( t \). This history is comprised of the buyers prior \( p_t^i(\hat{g}) \), the type of unit that he received in the current period \( x_t^i \in \{ H, L \} \), the market that he traded in \( m_t^i \), and the volume of certified trades in the market \( S^C \). If a buyer \( i \) at time \( t \) is given the trade history of all other buyers \( h^1(t), h^2(t), \ldots, h^N(t) \) and the prices and allocations of the market \( P^{NC}, P^C, S^{NC}, S^C \), a buyer’s posterior is

\[
(25) \quad p_{t+1}^i(\hat{g}|h^1(t), h^2(t), \ldots, h^N(t), P^{NC}, P^C, S^{NC}, S^C) = p_{t+1}^i(\hat{g}|h^i(t)).
\]

As such, even though the market is a common value auction, I treat each individual as if he has a private value for uncertified goods based only on his individual beliefs\(^{12}\).

Let \( \mu_b(m, s|\hat{g}_i, S^{NC}) \) be the belief of a buyer that a seller in market \( m \) is of type \( s \) conditional on the size of the uncertified market being \( S^{NC} \) and the true proportion of type-\( G \) sellers being \( \hat{g}_i \). A buyers of type \( b_{p_t^i(\hat{g})} \) (where type is defined by his prior in the period) trading in period \( t \) in market \( m \) at price \( P^m \) given \( S^{NC} \) has (perceived) expected utility:

\[
(26) \quad EU_t^i(m|S^{NC}, p_t^i(\hat{g})) = \sum_k \sum_s \mu_b(m, P^m, b_{p_t^i(\hat{g})}, s) \mu_b(m, s|\hat{g}_k, S^{NC}) p_t^i(\hat{g}_k|S^{NC}).
\]

The Stage Game

At each point in time \( t \), buyers and sellers take part in the following two part auction:

1. Part 1: The Uncertified Auction

   a. **Stage 1**: Each buyer \( i \) submits bid schedule \( \beta_{t}^{NC}(S^{NC}) \) which is his bid for an uncertified unit conditional on the number of sellers in the uncertified market

\(^{12}\)On the other extreme, a model in which all buyers understand the common value nature of the problem could be resolved with a direct mechanism in which all buyers submit their signals and the market clears using a Vickrey auction. In this case, each period, buyers perfectly learn the information of other buyers and a single common knowledge prior emerges each period. The results from this common prior model is similar to our common knowledge benchmark considered in section\(^{42}\).
being $S^{NC}$.

(b) **Stage 2:** All $M$ sellers enter into stage 2 and are sequentially ordered with type-$B$ sellers first, type-$C$ sellers next, and type-$G$ sellers last. Sellers play a sequence of rounds in which they decide whether to stay in the market or exit. At the beginning of each round, sellers who stayed in the market from the previous round sequentially and publicly decide whether to stay in the uncertified market or exit. Rounds are repeated until all sellers who started the round remain at the end of the round.

(c) **Stage 3:** Given that $K$ sellers remain in the uncertified market, the following Vickrey mechanism is implemented: Ordering buyers $i$ by their values of $\beta_i^{NC}(K)$, $K$ units are traded at a price equal to $\beta_{K+1}^{NC}(K)$. If $\beta_{K+1}^{NC} < U_H - T$ and $\beta_K^{NC} \geq U_H - T$, an exogenous reservation price $U_H - T$ is used.

2. **Part 2: The Certified Auction**

   (a) **Stage 4:** Each buyer $i'$ who did not purchase a unit in stage 3 submits a bid $\beta_i^{C}$ which is his bid for a certified unit.

   (b) **Stage 5:** All remaining sellers $j'$ who did not trade in stage 2 enter into stage 5 with the same ordering as in stage 2. Sellers sequentially and publicly decide whether to stay in the certified market or withdraw their bids.

   (c) **Stage 6:** Given that $S^C$ sellers enter the certified market, a Vickrey auction takes place where $S^C$ units are traded at the reservation price of the $S^C + 1$ highest bid.

In order to study myopic dynamics, I assume that buyers and sellers both have sufficiently high discount rates that they maximize their (perceived) expected utility in the current period. Buyers in this game also adopt the weakly dominant strategy of bidding their valuation for an object allowing the stage game to be solved via backward induction. Analysis of the game can be deduced via the following steps:

1. Starting in Stage 4, assume that $K$ buyers and sellers exchanged goods in the certified market. Buyers valuation for a certified unit is unaffected by beliefs and thus each buyer bids $U_H$ for a certified unit. Since there are more buyers than sellers and $U_H > T + C_{s}^{H}$ for $s \in \{G, C\}$, all type-$G$ and type-$C$ sellers that remain will choose to enter and trade in the certified market. Buyers in the Stage 4-6 auction all receive zero utility and all remaining sellers in the market receive a utility of $U_H - T - C_{s}^{H}$.

---

13. This ordering can be though of as coming out of a pre auction where sellers bid on the right to chose their order in the sequence.

14. Note that all buyers and sellers weakly prefer this reservation price rule in the market rather than letting the price fall to $U_L$.

15. If $U_H > T + C_{B}^{H}$, the type-$B$ seller will also enter the uncertified market if she has not sold a unit in Stage 3. This will never occur in equilibrium.
2. In Stage 2, given a set of contingent bids, sellers must decide whether to stay in the market or exit. Suppose that in a round there are \( k \) sellers remaining and that the first \( k - 1 \) sellers have elected to stay in the market. The last seller will stay in the market if 
\[
\beta_{K+1}^{NC}(K) - \min(C_L, C_H^k) > U^H - T - C_H^k
\]
and exit otherwise. Since the sellers are ordered in terms of their reservation values, a buyer earlier in the sequence can always wait to exit unit they are the last seller in the sequence. Thus, the exit of sellers from the market occurs in reverse order and any partial-certifying equilibrium will have type-\( G \) sellers being the first to exit the uncertified market.

3. In stage 1, a buyer knows that his utility from trading in the certified market is zero. Thus, his reservation value for waiting is zero and a truthful bid in the uncertified market must make his expected utility (based on his potentially wrong beliefs) equal to zero. Buyers know that sellers of type-\( G \) are the first to exit from the uncertified market. Thus, any equilibrium in which \( S^{NC} < M \) will have type-\( G \) sellers being the first to exit from the uncertified market. It follows, that conditional on \( \hat{g}_k \) and \( S^{NC} \), a buyers belief about the proportion of good types in the environment is

\[
\mu_b(NC, G|\hat{g}_k, S^{NC}) = \max\left(\frac{M\hat{g}_k - M + S^{NC}}{S^{NC}}, 0\right).
\]

Since in equilibrium, the type-\( B \) seller will always be in the uncertified market, \( S^{NC} > 0 \) and this belief is well defined.

4. A buyers perceived expected value for bidding in the uncertified market conditional on \( S^{NC} \) is \( EU_i(m|S^{NC}, p_i(\hat{g})) \). If this value is greater than \( U^H - T \) the buyer bids his actual value. Otherwise, the buyer correctly assumes that no high type buyer would accept his bid and he bids \( U^L \).

Equilibria

Observe that since sellers begin in the uncertified market, if the non-certifying equilibrium exists it will be selected. Further, if the non-certifying equilibrium does not exist, but partial-certifying equilibria do exist, the stage game will select the partial certifying equilibrium with the largest number of uncertified units. A slight modification to the game here where there is a positive cost to entering Stage 2 and entry is simultaneous will generate multiple equilibrium consistent with the model considered in the main text. The current model is considered to simplify the dynamic analysis.

Dynamics and Convergence

Given the structure of the entrance/exit round of the game, getting high and low quality units in the uncertified market can generate new information about the distribution of types for sellers who trade uncertified units in a non-certifying or partial-certifying equilibrium. I assume that in each period, a buyer updates his beliefs based on the type of unit that he receives conditioning on the fact that sellers do not select dominated outcomes. Basing
updating only on a buyer's history at the end of the period, a myopic Bayesian buyer updates his beliefs in the following way:

1. If a buyer trades in the uncertified market when $\Delta P_t > T$, the buyer will update his beliefs based on Bayes rule. Given a new unit of quality $x \in \{H, L\}$, a buyer recognizes that an uncertified unit must be from a type-$G$ seller and an uncertified unit must be from a type-$C$ or type-$B$ seller. He thus updates his prior as follows:

$$p_{t+1}(\hat{g}_i|x) = \frac{p_t(\hat{g}_i)q(x|\hat{g}_i)}{\sum_k p_t(g_k)q(x|\hat{g}_k)}$$

where

$$q(x|\hat{g}_k) = \begin{cases} \hat{g}_k & \text{for } x = H \\ 1 - \hat{g}_k & \text{for } x = L \end{cases}$$

is the probability of getting a unit of quality $x$ in the uncertified market given the proportion of high types is $\hat{g}_k$ and all sellers trade in the uncertified market.\(^{16}\)

2. If a buyer trades in the uncertified market when $\Delta P_t = T$, the buyer recognizes that the portion of sellers in the certified market must all be type-$G$ types. Taking this into account, the seller weights his update in relation to the number of units being traded in the certified market. Let $\frac{SC}{M}$ be the share of sellers who trade in the certified market. With probability $\frac{SC}{M}$, a buyer ignores his own signal and updates as if he received a high value units. With probability $1 - \frac{SC}{M}$, the buyer updates his beliefs with the actual quality of unit he purchases.\(^{17}\)

3. If a buyer trades for a certified unit or an uncertified unit with $\Delta P_t < T$, then

$$p_{h(t+1)}(\hat{g}) = p_{h(t)}(\hat{g}).$$

Repeated play of the stage game using these updating rules yields the following theorem:

\(^{16}\)The use of $q(x|\hat{g}_k)$ in this equation is to highlight that there is actually two steps taking place in updating the posterior over types. The first is an empirical update on the likelihood of getting a high quality unit in the uncertified market. The second is mapping this empirical data back into implications about the proportion of type-$G$ sellers in the environment under the assumption that sellers do not play dominated strategies.

\(^{17}\)Beliefs using this update method converge to $\frac{SC}{M} MG + \frac{M-SC}{M} \frac{SC}{M} MG = \frac{MG}{M} = g$. Note that updating a prior in the partial-certifying equilibrium requires stronger assumptions about the strategies of sellers. A buyer observing empirical data in a partial-certifying equilibrium requires specific sorting and strategies in Stage 2 of the game in order to map empirical data back into types. Strict ordering of seller types may not hold in more general versions of the game where sellers learn about their best strategies through experimentation (such as a model with fictitious play or reinforced learning). This suggests that beliefs are significantly more likely to be incorrect in a partial-certifying equilibrium relative to the non-certifying equilibrium.
Theorem 5 As \( t \to \infty \):

1. If the non-certifying equilibrium occurs, then there exists at least \( M \) buyers such that

\[
(28) \quad p_i^t(\hat{g}) \overset{a.s.}{\longrightarrow} g.
\]

2. If a partial certifying equilibrium occurs with \( M - k^* \) trades in the uncertified market, then there exists at least \( M - k^* \) buyers such that:

\[
(29) \quad p_i^t(\hat{g}) \overset{a.s.}{\longrightarrow} g
\]

3. If the certifying equilibrium occurs at time \( t^* \), then

\[
(30) \quad p_i^t(\hat{g}) = p^t_*(\hat{g}) \quad \forall t > t^*
\]

Theorem 5 stems from the fact that sellers choices in Stage 2 of the game always correspond to the equilibrium with the largest number of buyers in the uncertified market. In each period that the non-certifying equilibrium occurs, the \( M \) buyers with the highest (perceived) expected utility for the round get units and update their beliefs. As long as the non-certifying equilibrium continues to exist, these buyers will continue to learn from the market leading to eventual convergence of beliefs.

If there exists a period where there are no longer \( M \) buyers with an expected utility for uncertified units larger than \( U^H - T \), the non-certifying equilibrium no longer exists and a partial-certifying equilibrium forms. Given the structure of Stage 2, good type sellers are the first to exit, and thus belief updates for buyers based on the size of the certified market and the proportion of high quality goods still provide a consistent estimate for the proportion of type-\( G \) sellers. Over time, the sellers who stay in the uncertified market learn, but the other sellers do not.

Finally, if there is ever a period in which all non-certifying and partial-certifying equilibrium do not exist, only type-\( B \) sellers will stay in the uncertified market. Quality updates no longer reveal information about the proportion of type-\( G \) sellers in the environment and thus buyers no longer receive private signals on which to update beliefs.

Given the convergence of beliefs, prices in the dynamic game converge as follows:

**Corollary 1** Suppose that a series of auctions are played in which 1) buyers play the weakly dominant strategy of truth telling and 2) buyers update their beliefs myopically. Then, as \( t \to \infty \), the prices generated in the uncertified auctions must converge to one of the following three prices:

- **Certifying Equilibrium**: \( P^{NC} = U^L \)
- **Partial-Certifying Equilibrium**: \( P^{NC} = U^H - T \).
- **Non-Certifying Equilibrium**: \( P^{NC} \leq U^H - (1 - \hat{g})(U^H - U^L) \).
The prices in the Non-Certifying Equilibrium are not strictly equal to the price found in the competitive equilibrium model since it is the value for player $M + 1$ that pins down the uncertified price.

7.1.3 Information

The theoretical models in the last two sections point to two reasons why a partial-certifying equilibrium might form: loss aversion and incorrect beliefs by buyers. In this section, I use the public information treatment as exogenous variation in order to differentiate these two effects. I find that both information and heterogeneous preferences contribute to the formation of the partial-certifying equilibrium and may contribute toward the adoption of inefficient certification technologies.

Hypothesis and Empirical Strategy

If the partial-certifying equilibrium is due purely to loss aversion, the creation of public information about the proportion of uncertified goods should have no effect on the partial-certifying equilibrium. In contrast, if the partial-certifying equilibrium is due to limited information, public information should increase the number of uncertified trades since it decreases the likelihood that buyers will have incorrect beliefs about the proportion of sellers in the marketplace. Using this second prediction as the null hypothesis I predict:

**Hypothesis 6** In the $S^{Pre}$ environment, the information treatment should increase the likelihood that buyers purchase risky units from the uncertified market.

To test this hypothesis, I run a probit regression where I study the relationship between buying units that have the potential for losses and the information environment. I interact the information treatment with the $S^{Pre}$ treatment since this is the only treatment where the information treatment should matter and estimate:

$$\text{Risky}_{i,s} = \alpha_0 + \sum \alpha_s + \beta_{\text{LossAverse}}I_{\text{LossAverse}} + \beta_{\text{Beliefs}}I_{\text{Beliefs}} + \beta_{\text{Public}}(I_{\text{Public}} \ast I_{S^{Pre}}) + \beta_{\text{Public} \ast \text{LN}}(I_{\text{LossNeutral}} \ast I_{\text{Public}} \ast I_{S^{Pre}}) + \epsilon$$

where $\text{Risky}_{i,s}$ is 1 if a buyer purchased an uncertified unit that had the potential for a loss, $\alpha_s$ are session level fixed effects, $I_{\text{LossAverse}}$ is an indicator variable for individual loss aversion, $I_{\text{Public}}$ is an indicator for the public information treatment, and $I_{S^{Pre}}$ is an indicator for environment $S^{Pre}$. I predict that $\beta_{\text{Public} \ast \text{LN}} > 0$ and that $\beta_{\text{Public}} > 0$.

Results

Evidence from the experimental sessions shows a clear effect of information on the proportion of uncertified trades in the environment. Figure 6 shows the proportion of risky trades undertaken by loss neutral and loss averse individuals divided between the $S^{Pre}$ treatment where the non-certifying equilibrium formed and all other treatments where the certifying
equilibrium formed. As can be seen on the left hand side, public information greatly increases the probability that a loss neutral individual takes risk in the uncertified market and has a positive much smaller effect on buyers who are loss averse. In other treatments, there is minimal differences between risk and loss neutral individuals in their response to information. These same forces can be seen in the probit regression shown in table 12.

**Figure 6: Effect of Information on the Propensity to Purchase Risky Units**

<table>
<thead>
<tr>
<th>Environment $S^{Pre}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public Information</td>
</tr>
<tr>
<td>Private Information</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Other Environments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public Information</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Loss Aversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.568***</td>
</tr>
<tr>
<td>(0.069)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Public * $S^{Pre}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.349***</td>
</tr>
<tr>
<td>(0.130)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Loss Neutral * Public * $S^{Pre}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.480***</td>
</tr>
<tr>
<td>(0.151)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Beliefs</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.005</td>
</tr>
<tr>
<td>(0.017)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Constant</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.835***</td>
</tr>
<tr>
<td>(0.119)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Psuedo Adj. $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>.216</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>3353</td>
</tr>
</tbody>
</table>

18 Probit regression with fixed effects is shown to be biased by Fernandez-Val (2007). The current specification does not take this bias into account.
7.2 Proofs

**Proof. Lemma 1:** By the definition of $s \in \{G, C, B\}$, $C^H_B \geq C_L + U^H - U^L - T \geq C^H_C \geq C^H_G$. Thus, in the uncertified market, only type-$G$ sellers will produce high-quality goods. Writing out the utility of the seller:

$$ v(m, P^m, b, s) = \begin{cases} 
P^C - C^H_s - T & \text{if } m \in C, s \in \{G, C, B\}, \\
P^{NC} - C^H_s & \text{if } m \in NC, s \in \{G\}, \\
P^{NC} - C^L & \text{if } m \in NC, s \in \{C, B\}.
\end{cases} $$

By Definition 2

$$ S(m^*, s) \neq 0 \Leftrightarrow m^* \in \underset{m}{\arg \max} \sum_b v(m, P^m, b, s) \quad \forall s. $$

Finding the points where each seller type is indifferent between the certified and uncertified markets lead directly to Lemma 1. ■

**Proof. Lemma 2:** In the baseline model, there is only one type of buyer which I denote as $b_0$ whose utility is as follows:

$$ u(m, P^m, b_0, s) = \begin{cases} 
U^H - P^C & \text{if } m \in C, s \in \{G, C, B\} \\
U^H - P^{NC} & \text{if } m \in NC, s \in \{G\} \\
U^L - P^{NC} & \text{if } m \in NC, s \in \{C, B\}.
\end{cases} $$

It follows:

1. When $\Delta P > T$, $v(C, P^C, b, G) > v(NC, P^{NC}, b, G)$ and thus $S(NC, G) = 0$. By the definition of a price equilibrium, $\mu_b(NC, G) = \frac{S(NC, G)}{\sum_s S(NC, s)} = 0$ and thus

$$ \sum_s u(NC, P^{NC}, b_0, s) \mu_b(NC, s) = U^L - P^{NC}. $$

Since $\forall s, u(C, P^C, b_0, s) = U^H - P^C$ and $u(\emptyset, P^{NC}, b_0, s) = 0$, it follows that an agent is indifferent between all three markets when $P^{NC} = U^H, P^C = U^L$

2. When $\Delta P \leq T$, $\forall s, v(C, P^C, b, s) < v(NC, P^{NC}, b, s)$ and thus $S(NC, G) = M_G$. By the definition of a price equilibrium, $\mu_b(NC, G) = \frac{S(NC, G)}{\sum_s S(NC, s)} = \frac{M_G}{M} = g$. It follows that

$$ \sum_s u(NC, P^{NC}, b_0, s) \mu_b(NC, G) = g U^H + (1-g) U^L - P^{NC}. $$

A buyer is indifferent across all three markets if $P^{NC} = U^H - (1-g)(U^H - U^L)$ and $P^C = U^H$. ■
Proof. Theorem 1:

1. When $\Delta P = U^H - U^L$:
   
   (a) By Lemma 1, $S(NC, B) = M_B$, $S(C, G) = M_G$, and $S(C, C) = M_C$.
   
   (b) By Lemma 2, if $P^{NC} = U^H$, $P^C = U^L$, $D(C, b_0) = [0, N_{b_0}] \in \mathbb{I}_+$, $D(NC, b_0) = [0, N_{b_0}] \in \mathbb{I}_+$, with $\Sigma_m D(m, b_0) = N_{b_0}$.

   Thus the attainable allocation where $P^{NC} = U^H$, $P^C = U^L$, $D(C, b_0) = M_G + M_C$, $D(NC, b_0) = M_B$, and $D(\emptyset, b_0) = N_{b_0} - M$ always exists.

2. When $\Delta P > T$:
   
   (a) By Lemma 1, $S(NC, B) = M_B$, $S(NC, G) = M_G$, and $S(NC, C) = M_B$.
   
   (b) By Lemma 2, a buyer is indifferent between all three markets if $P^{NC} = U^H - (1 - g)(U^H - U^L)$ and $P^C = U^H$.

   If $P^C - P^{NC} = (1 - g)(U^H - U^L) > T$, then $D(NC, b_0) = M$, $D(\emptyset, b_0) = N_{b_0} - M$ is an equilibrium. Otherwise, there does not exist a set of prices such that $\Delta P > T$ and a buyer is indifferent between the certified and uncertified market.

Proof. Theorem 2: By Lemma 1 $S(C, C) > 0 \rightarrow S(C, G) = M_G$. It follows that for any competitive equilibrium where type-C sellers certify their good, all type-G sellers certify their good. Define

$$W = \Sigma_m u(m, P^m, b_0, s)\mu_m(m, s)D(m, b_0) + \Sigma_{s,m} v(m, P^m, b_0, s)S(m, s)$$

Since $u(C, P^C, b_0, G) + v(C, P^C, b_0, G) - T \leq u(NC, P^{NC}, b_0, G) + v(NC, P^{NC}, b_0, G)$, $W$ is decreasing in $S(C, G)$. Likewise, since $u(C, P^C, b_0, C) + v(C, P^C, b_0, C) - T \geq u(NC, P^{NC}, b_0, C) + v(NC, P^{NC}, b_0, C)$, $W$ is increasing in $S(C, C)$. Thus, the constrained Pareto Efficient equilibrium must either be the certifying equilibrium where all the type-C sellers trade in the certified market or the non-certifying equilibrium where no type-G sellers certify their goods. In cases where the non-certifying equilibrium does not exist but where partial-certifying equilibria do exist, it is either the certifying equilibrium or the partial-certifying equilibrium with the least amount of certification that is constrained Pareto efficient.

Proof. Theorem 3: Let $\hat{g}$, $\hat{c}$, $\hat{b}$ be the prior beliefs about the proportion of good, conditional and bad agents in the market. When a non-certifying equilibrium exists, $P^{NC} = U^H - (1 - \hat{g})(U^H - U^L)$. When a certifying equilibrium exists, only type-B sellers are in the uncertified market.

Define $S^{NC}$ as the number of sellers trading in the uncertified market. Then, if the number of buyers in each market are known, prices are observable, and the marginal valuations for the pivotal buyer are known, the posteriors $\hat{g}$, $\hat{c}$, $\hat{b}$ under the non-certifying and certifying equilibrium are as follows:

54
When $U^H$ and $U^L$ are known and the market is in a non-certifying equilibrium, $\hat{g} = g$ and thus price is a sufficient statistic for the proportion of type-$G$ sellers in the environment. 

**Proof. Theorem 4:** For a partial certifying equilibrium to exist, it must be the case that at $\Delta P = T$, $\sum_b D(\mathcal{C}, b) = \sum_s S(\mathcal{C}, s)$ and $\sum_b D(\mathcal{NC}, b) = \sum_s S(\mathcal{NC}, s)$. By Lemma 1, at $\Delta P = T$, $S(\mathcal{C}, G) \in \mathbb{I}_+$, $S(\mathcal{NC}, G) = M_G - S(\mathcal{C}, G)$, $S(\mathcal{NC}, C) = M_C$ and $S(\mathcal{NC}, B) = M_G$. Buyers are ordered with increasing loss aversion, thus if there exists at least one buyer with loss aversion $\lambda_i$ and $\lambda_{i+1}$, if $D(m, \lambda_i) = 0 \rightarrow D(m, \lambda_{i+1}) = 0$. Further, if $D(m, \lambda_{i+1}) > 0 \rightarrow D(m, \lambda_i) > N_{\lambda_i}$.

Given this monotonicity, it is sufficient to look for solutions that have the first $k$ buyer types demanding from the uncertified market, and the remaining buyers demanding from the certified and null markets. Define the prices at which the $k^{th}$ buyer is willing to buy given $D(\mathcal{NC}, \lambda_{k+1}) = 0$ and $D(\mathcal{NC}, \lambda_{k-1}) = N_{k-1}$ as:

\[
\tilde{P}_{\mathcal{NC}}^{M-k} = \frac{\tilde{g}_{-k} U^H + (1 - \tilde{g}_{-k}) \lambda_{M-k} U^L}{\tilde{g}_{-k} + (1 - \tilde{g}_{-k}) \lambda_{M-k}},
\]

\[
\tilde{P}_{\mathcal{C}}^{M-k} = U^H,
\]

\[
\tilde{g}_{-k} = \frac{M_G - k}{M - k}.
\]

If $\tilde{P}_{\mathcal{C}}^{M-k} - \tilde{P}_{\mathcal{NC}}^{M-k} > T$, $D(\mathcal{NC}, \lambda_k) = 0$. Otherwise $D(\mathcal{NC}, \lambda_{k}) > 0$ and thus buyers $i \in \{1, \ldots, k\}$ all have strictly positive demand. For each $k \in \{1, \ldots, M - 1\}$ such that $\tilde{P}_{\mathcal{C}}^{M-k} - \tilde{P}_{\mathcal{NC}}^{M-k} > T$ and $\tilde{P}_{\mathcal{C}}^{M-k+1} - \tilde{P}_{\mathcal{NC}}^{M-k+1} \leq T$, a partial certifying equilibrium exists where buyers $i \in \{1, \ldots, k\}$ demand 1 unit from the uncertified unit, and $M - k$ other buyers are randomly matched with type-$G$ sellers in the certified market. 

**Proof. Theorem 5:**

The proof of theorem 5 is built in two steps. I first show that for the non-certifying equilibrium to exist as $t \rightarrow \infty$, the non-certifying must occur an infinite number of times. I then show that if a buyer samples from the uncertified market an infinite number of times, his beliefs will converge to true proportion.

**Step 1** By construction, in any period where the non-certifying equilibrium exists, this equilibrium is selected. In any period where the partial-certifying equilibrium there exists a $k$ such that buyers $k + 1, \ldots, N$ trade in either the certified market or don’t trade. Since buyers in these markets do not update their beliefs, a market where the partial-certifying equilibrium has been adopted never returns to the non-certifying equilibrium.
By similar logic, the non-certifying equilibrium does not return to the partial-certifying equilibrium or non-certifying equilibrium. Thus, for the non-certifying equilibrium to exist as \( t \to \infty \), all periods leading up to the current period must all have had the non-certifying equilibrium form.

Step 2 Let \( x = (x_1, \ldots, x_T) \) be observations of a single buyer where \( x_i = \{H, L\} \). As before, let \( \hat{g} \in \{0, \frac{1}{M}, \ldots, \frac{M-1}{M}\} \) be the possible number of type-\( G \) sellers in the market. Given an initial prior \( p_i(\hat{g}) = \{p_i(\hat{g}_0), p_i(\hat{g}_1), \ldots, p_i(\hat{g}_{M-1})\} \) where \( p_i(\hat{g}_k) > 0 \) and \( \Sigma_k p_i(\hat{g}_k) = 1 \), the posterior \( p_t(\hat{g} | x) \) converges almost surely to the true proportion as \( t \to \infty \) as long as \( g \in \hat{g} \) and

\[
\Sigma x q(x | \hat{g}_i) \log \left[ \frac{q(x | \hat{g}_i)}{q(x | \hat{g}_j)} \right] > 0.
\]

Expanding condition (32) yields:

\[
\hat{g}_i \log \left( \frac{\hat{g}_i}{\hat{g}_j} \right) + (1 - \hat{g}_i) \log \left( \frac{1 - \hat{g}_i}{1 - \hat{g}_j} \right)
\]

Rewriting \( \hat{g}_j = \hat{g}_i + z \) and taking the derivative with respect to \( z \), the first derivative is zero at \( z = 0 \) and the second derivative is strictly positive for all \( z \). Thus condition (32) holds. Since \( g \in \{0, \frac{1}{M}, \ldots, \frac{M-1}{M}\} \), convergence is guaranteed as \( t \to \infty \).

The proof of convergence of beliefs in the partial-certifying equilibrium follows a similar logic. By construction, if the non-certifying equilibrium does not exist, the partial certifying equilibrium with the largest number of uncertified buyers is selected. Over time, buyers in the uncertified market are always the same and thus they sample an infinite amount of time. Since their updating rule is consistent, beliefs converge to the true values.

**Proof. Corollary 1:**

In each period of the non-certifying equilibrium, the \( M \) buyers with the highest valuations for an uncertified unit win units from the auction. Since \( N \) is finite, as \( t \to \infty \), there exists at least \( M \) buyers who get infinite draws. For these buyers, the willingness to pay for an uncertified unit converges to \( gU^H + (1-g)U^L \). If a buyer has a higher value than \( gU^H + (1-g)U^L \), he will be included in the \( M \) buyers who win a unit and as such, his valuation will decrease over time toward the true valuation. The price paid at auction, based on the valuation for the \( M - 1 \)th seller must therefore be at or below \( P^{NC} = gU^H + (1-g)U^L \).

7.3 Instructions
Sellers Instructions

Before the experiment, subjects were randomly split into two groups: buyers and sellers. These are a translated version of the instructions given to the sellers. Instructions for the buyers as well as the computerized instructions are available upon request.

Today you will take part in a market experiment. Please read through the following instructions carefully. All the information you need to successfully participate in this experiment is written here. If you have questions regarding the experiment or the instructions, please raise your hand. An instructor will come to your desk and will answer your question.

By participating in this experiment, you automatically receive a show-up fee of 10 Francs. In the course of the experiment you can earn additional money by earning points through trading. The amount of points you will earn depends on your decisions and the decisions of other participants during the experiment.

The experiment is split up into 24 separate periods. In each period you will interact with other participants in the experiment using the computer in front of you. The points that you earn during this experiments are converted into francs at the end of the experiment. The conversion rate is:

30 Points = 1 Swiss Franc

At the end of the experiment, six periods are randomly chosen and you will receive the amount of money you earned in these periods plus the 10 francs show-up fee in cash.

Please be aware that communication is strictly forbidden during the time you are in the laboratory. Also note that the use of the computer is restricted to the experimental program only. Communication or manipulating of the computer will result in exclusion from the experiment. If you have any questions please raise your hand and an instructor will answer them.

Overview of the course of the experiment

In this study you are a seller in a market with RED and BLUE products. The market consists of 5 buyers and 6 sellers. As a seller, you may sell up to two products. You will earn a number of points on a transaction equal to the price that you sell a unit minus the cost for producing the unit and any certification costs that you incur.

Your Earnings = Price – Production Cost – Certification Fee

In the market, you may sell two types of products: RED and BLUE. These products are of different quality and may have different valuations to the buyers in the market. A buyer earns money if he pays less than his valuation for a product. A buyer’s valuation for a product depends on the quality of the product that he receives and the total number of units that he has already bought in the period.

Initially, the buyers and other sellers can not observe the quality of the unit that you are selling. You may choose to offer certified units instead of normal units which guarantee a specific color to the buyer. If you sell a certified unit, you will be charged 60 points in certification fees at the time of transaction.

In total the experiment consists of 24 Periods. The course of each period is as follows:

1. The Trading Phase: In the trading phase, you will trade with buyers in the market. The trading phase in the first 3 periods will be 4 minutes. The trading phase for the remaining periods will be 2 minutes. During the trading phase, you may complete trades either by posting offers that a buyer accepts or by accepting bids from the buyers.

Your offer to sell:
- Your offer to sell consists of the following specifications:
  1) the price that buyers have to pay for a unit of the product
  2) the quality of the product
  3) whether there is a certificate for the product
- The other participants can only see the actual quality of a product if the product is certified. If the product is not certified, the product quality will be labeled “UNKNOWN”.

The offers from buyers:
- A buyer’s bid to buy consists of the following specifications:
  1) the price he is willing to pay for a unit of the product
  2) the desired quality of the product
  3) whether the buyer requires a certificate or not
- If a buyer requests a certificate you must sell the buyer his desired quality. If the buyer doesn’t request a certificate you can sell either quality.

2. The Bonus Phase: The next phase is the bonus phase. In this phase you have to guess how many of the sellers had lower cost producing the RED quality than producing the BLUE quality during the respective period. If your guess is correct you will earn 20 points.

3. The Earnings Screen: At the end of each period you will see the earnings screen. Each participant is informed how much he has earned during the last trading period.

6 out of the 24 Periods are randomly chosen and the earnings of these periods and the show-up fee will be paid out in cash at the end of the experiment.

Detailed course of the experiment

During the experiment you will enter your decisions using the computer. In the following instructions, all the functions will be explained in detail.

1. The Trading Phase

At the beginning of the trading phase, you will be informed of the production costs for the following period. When all players have reviewed their cost and value information, the trading phase will begin.

During the first three periods the trading phase will last for 4 minutes. In the remaining periods, the trading phase will last for 2 minutes. The clock in the upper right hand corner of the screen will show the remaining time in a period in seconds. When this clock reaches zero the game will immediately end and you will not be able to make any more trades.

During each trading phase you will see the following screen:
Product Quality

There are two possible product qualities: RED and BLUE. Your production costs as well as the valuations of the buyers differ with the quality. In each period either the RED or the BLUE quality can be cheaper for you to produce.

Sellers Production Costs

The production costs of a product depend on two things. First the quality (RED or BLUE) of the product influence the costs and second certification increases the production costs. In every period you will see your costs on the lower left side of the trading screen.

Your costs can change from period to period, so please pay close attention to your production costs.

The following cost structures can occur during the experiment. In each period one of the following cost structures will be applicable. Please note that different sellers may have different costs during each period.

Case 1, RED Quality is cheaper to produce:

<table>
<thead>
<tr>
<th>Quality</th>
<th>Costs without certification</th>
<th>Costs with certification</th>
</tr>
</thead>
<tbody>
<tr>
<td>RED</td>
<td>30</td>
<td>30 + 60 = 90</td>
</tr>
<tr>
<td>BLUE</td>
<td>50</td>
<td>50 + 60 = 110</td>
</tr>
</tbody>
</table>

Case 2a, BLUE Quality is cheaper to produce:

<table>
<thead>
<tr>
<th>Quality</th>
<th>Costs without certification</th>
<th>Costs with certification</th>
</tr>
</thead>
<tbody>
<tr>
<td>RED</td>
<td>80</td>
<td>80 + 60 = 140</td>
</tr>
<tr>
<td>BLUE</td>
<td>50</td>
<td>50 + 60 = 110</td>
</tr>
</tbody>
</table>

Case 2b, BLUE Quality is cheaper to produce:

<table>
<thead>
<tr>
<th>Quality</th>
<th>Costs without certification</th>
<th>Costs with certification</th>
</tr>
</thead>
<tbody>
<tr>
<td>RED</td>
<td>130</td>
<td>130 + 60 = 190</td>
</tr>
<tr>
<td>BLUE</td>
<td>50</td>
<td>50 + 60 = 110</td>
</tr>
</tbody>
</table>

Certification

The other participants, buyers and sellers, can only see the quality of a product if the product is certified. A buyer can see the quality of products without a certificate only after the purchase. In this case the quality of the product will be labeled “UNKNOWN”.

To reveal the quality of a product to the buyers, you can elect to certify your product. As you can see in the table above, certification increases the production cost by 60 Points. The certification costs only occur when a product is sold. So you don’t have to pay certification costs for an unsold unit.

Your offers to buyers

You and all the other sellers can post offers to buyers during the whole period. If you want to post an offer you have to specify the following:

- You have to specify a price, which the buyer has to pay for the product. The price has to lie between 0 and 400:

  \[ 0 \leq \text{Price} \leq 400 \]

- You have to specify the quality:

  \[ \text{Quality} = \text{RED or BLUE} \]

- You have to decide whether you will issue a certificate:

  \[ \text{Certificate} = \text{Yes or No} \]

  \[ \text{Costs of certification} = 60 \]

As soon as you have made all the required specifications you can validate your offer by clicking on the “post offer”-button.

This information will appear on the screen in the field offers to sell and all the other participants, buyers and sellers can see it. Your own offers will appear in blue, the offers of all the other sellers appear in black. The offers to sell appear in descending order of the price on the screen.

As soon as a buyer accepts an offer, the respective offer disappears from the screen. If you want to post the same offer again, you have to reenter all the specifications.

As long as you can sell at least one unit you can have two standing offers, one that is certified and one that is not certified. After your second sale all of your standing offers will be deleted. If you have a standing offer, and you enter a new offer, the new offer replaces the old one, if both offers have the same certification status.
You have the following standing offers:

<table>
<thead>
<tr>
<th>Quality</th>
<th>Price</th>
<th>Certified</th>
</tr>
</thead>
<tbody>
<tr>
<td>RED</td>
<td>400</td>
<td>Yes</td>
</tr>
<tr>
<td>BLUE</td>
<td>50</td>
<td>No</td>
</tr>
</tbody>
</table>

Now you enter an offer for a RED quality product at the price of 350 and you offer a certificate. Your standing offers will change to:

<table>
<thead>
<tr>
<th>Quality</th>
<th>Price</th>
<th>Certified</th>
</tr>
</thead>
<tbody>
<tr>
<td>RED</td>
<td>350</td>
<td>Yes</td>
</tr>
<tr>
<td>BLUE</td>
<td>50</td>
<td>No</td>
</tr>
</tbody>
</table>

Now you enter an offer for a RED quality product at the price of 250 and you do not offer a certificate. Your standing offers will change to:

<table>
<thead>
<tr>
<th>Quality</th>
<th>Price</th>
<th>Certified</th>
</tr>
</thead>
<tbody>
<tr>
<td>RED</td>
<td>350</td>
<td>Yes</td>
</tr>
<tr>
<td>RED</td>
<td>250</td>
<td>No</td>
</tr>
</tbody>
</table>

To withdraw offers you can click the “withdraw offers”-button and all your offers are withdrawn.

Accepting offers from buyers

The offers to buy are sorted in descending order of the price.

To accept an offer from a buyer, you select the line of the respective offer and click the “sell RED”-button, if you want to sell the RED Quality or click the “sell BLUE”-button if you want to sell the blue quality.

- If the buyer doesn’t request certification, you can sell either quality.
- If the buyer request certification, you have to sell the desired quality AND you have to pay the certification cost.

<table>
<thead>
<tr>
<th>Quality</th>
<th>Price</th>
<th>Certified</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROT</td>
<td>80</td>
<td>140</td>
</tr>
<tr>
<td>BLAU</td>
<td>60</td>
<td>120</td>
</tr>
</tbody>
</table>

Exercise 1: A buyer bids 180 for a product and doesn’t request a certificate, how much do you earn with this sale?

Earnings if you sell a BLUE quality product =
Earnings if you sell a RED quality product =

Exercise 2: You sell a RED Quality good for which a buyer paid 150. How high are your earnings if the buyer requests a certificate and what do you earn if he doesn’t request a certificate?

Earnings with certificate =
Earnings without certificate =

Exercise 3: There are the following two standing offers of buyers:

<table>
<thead>
<tr>
<th>Offer number</th>
<th>Price</th>
<th>Quality</th>
<th>Certificate requested</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>220</td>
<td>BLUE</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>180</td>
<td>RED</td>
<td>No</td>
</tr>
</tbody>
</table>

Through which sale can you make the higher earnings?

Possible earnings through offer number 1 =
Possible earnings through offer number 2 =

2. The bonus phase

Following the trading phase is the bonus phase. In this phase you have to guess how many of the sellers had lower cost producing the RED Quality than producing the BLUE quality during the respective period. If your guess is correct you will get 20 points.

3. The earning screen

At the end of each period you will see the earnings screen. There you will find your market earnings of the period.

Six out of the 24 Periods are randomly chosen and the earnings of these periods and the showup fee will be paid out in cash at the end of the experiment.

Omitted: Examples of How Earnings Is Calculated, Example of Randomized Payment

Exercises
7.4 Summary Statistics

The summary statistics reported here are for the last 6 periods of each environment.

1. **Session:**
   
   (a) Sessions 1-6 are Safe/Hazardous treatments.
   
   (b) Sessions 7-12 are Hazardous/Safe treatments.
   
   (c) Sessions 4-6 and 10-12 are public information treatments.

2. **Uncertified Price:** Average Price across both uncertified low-quality units and uncertified high-quality units.

3. **Certified Price:** Average Price of certified high-quality units.

4. **Uncertified High Quality:** Total number of uncertified high-quality units.

5. **Uncertified Low Quality:** Total number of uncertified low-quality units.

6. **Certified High Quality:** Total number of certified high-quality units.

7. **Number Loss Aversion:** Number of buyers reporting that they were unwilling to take losses in a period

<table>
<thead>
<tr>
<th>Session</th>
<th>Uncertified Price</th>
<th>Certified Price</th>
<th>Uncertified High Quality</th>
<th>Uncertified Low Quality</th>
<th>Certified High Quality</th>
<th>Number Loss Averse</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>141</td>
<td>194</td>
<td>22</td>
<td>12</td>
<td>28</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>170</td>
<td>203</td>
<td>19</td>
<td>13</td>
<td>40</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>103</td>
<td>193</td>
<td>1</td>
<td>17</td>
<td>53</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>145</td>
<td>202</td>
<td>19</td>
<td>12</td>
<td>39</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>143</td>
<td>197</td>
<td>34</td>
<td>18</td>
<td>19</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>176</td>
<td>209</td>
<td>40</td>
<td>11</td>
<td>16</td>
<td>2</td>
</tr>
</tbody>
</table>
### Table 14: Summary Statistics for $H^{Post}$ Environment

<table>
<thead>
<tr>
<th>Session</th>
<th>Uncertified Price</th>
<th>Certified Price</th>
<th>Uncertified High Quality</th>
<th>Uncertified Low Quality</th>
<th>Certified High Quality</th>
<th>Certified Low Quality</th>
<th>Number</th>
<th>Loss Averse</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>114</td>
<td>200</td>
<td>0</td>
<td>44</td>
<td>25</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>126</td>
<td>203</td>
<td>1</td>
<td>31</td>
<td>34</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>98</td>
<td>199</td>
<td>0</td>
<td>19</td>
<td>51</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>119</td>
<td>204</td>
<td>0</td>
<td>23</td>
<td>47</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>114</td>
<td>211</td>
<td>0</td>
<td>20</td>
<td>50</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>124</td>
<td>201</td>
<td>1</td>
<td>19</td>
<td>47</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 15: Summary Statistics for $H^{Pre}$ Environment

<table>
<thead>
<tr>
<th>Session</th>
<th>Uncertified Price</th>
<th>Certified Price</th>
<th>Uncertified High Quality</th>
<th>Uncertified Low Quality</th>
<th>Certified High Quality</th>
<th>Certified Low Quality</th>
<th>Number</th>
<th>Loss Averse</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>125</td>
<td>202</td>
<td>0</td>
<td>39</td>
<td>28</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>97</td>
<td>193</td>
<td>1</td>
<td>14</td>
<td>51</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>106</td>
<td>201</td>
<td>0</td>
<td>22</td>
<td>46</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>109</td>
<td>202</td>
<td>0</td>
<td>25</td>
<td>45</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>118</td>
<td>212</td>
<td>2</td>
<td>29</td>
<td>36</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>106</td>
<td>200</td>
<td>0</td>
<td>21</td>
<td>49</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 16: Summary Statistics for $S^{Post}$ Environment

<table>
<thead>
<tr>
<th>Session</th>
<th>Uncertified Price</th>
<th>Certified Price</th>
<th>Uncertified High Quality</th>
<th>Uncertified Low Quality</th>
<th>Certified High Quality</th>
<th>Certified Low Quality</th>
<th>Number</th>
<th>Loss Averse</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>134</td>
<td>201</td>
<td>12</td>
<td>24</td>
<td>36</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>102</td>
<td>192</td>
<td>0</td>
<td>11</td>
<td>60</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>99</td>
<td>198</td>
<td>0</td>
<td>18</td>
<td>54</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>111</td>
<td>198</td>
<td>2</td>
<td>12</td>
<td>56</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>121</td>
<td>205</td>
<td>6</td>
<td>12</td>
<td>54</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>101</td>
<td>191</td>
<td>1</td>
<td>12</td>
<td>59</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### References


