Essays on Institutional Persistence, Contract Structure, and Authority

by

Tom S. Wilkening

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Economics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

[June 2008]
May 2008

Author ....
Department of Economics
May 28, 2008

Certified by ....
Sergei Izmalkov
Assistant Professor
Thesis Supervisor

Certified by ....
Abhijit Banerjee
Ford Foundation International Professor of Economics
Thesis Supervisor

Accepted by ....
Peter Temin
Chairman, Department Committee on Graduate Students

Essays on Institutional Persistence, Contract Structure, and Authority

by

Tom S. Wilkening

Submitted to the Department of Economics
on May 28, 2008, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Economics

Abstract

This thesis is a collection of essays that uses theoretical and experimental methods to explore institutional persistence, contract structure, and authority.

Chapter One studies the informational and efficiency properties of institutions that form to reduce moral hazard. 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. Using experimental and theoretical methods, it analyzes 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. The certifying equilibrium, which does not carry information about changes in the distribution of sellers, does not adjust when the underlying environment changes, perpetuating a market structure that makes *all* market participants weakly worse off.

Chapter Two studies how changes in contract structure may help preserve antiquities. Most countries prohibit the export of certain antiquities. This practice often leads to illegal excavation and looting for the black market, which damages objects and destroys important aspects of the archaeological record. Chapter Two argues that many of the goals for export bans could be better accomplished through the use of long-term leases which would raise revenue for the country of origin while preserving national long-term ownership rights.

Chapter Three uses experiments to study how control rights are distributed in a setting with incentive conflicts. It shows that while effort levels are consistent with theoretical predictions, principals retain control rights even when it is strongly in their interest to delegate. Chapter Three also documents a differential response to authority by gender. As agents, women have strong fairness preferences resulting in diminished effort in asymmetric treatments but higher effort in symmetric ones. As principals, women are more likely to transfer authority when it is efficient to do so.

Thesis Supervisor: Sergei Izmalkov
Title: Assistant Professor

Thesis Supervisor: Abhijit Banerjee
Title: Ford Foundation International Professor of Economics

To my parents, Mary and Stuart Wilkening,

my brother Jon, and my sister Katie

Whose love and support accompanies me wherever I go

Acknowledgments

This thesis has benefited greatly from conversations, advice, and support by friends, colleagues, and mentors. Foremost, I would like to thank my advisors Abhijit Banerje, Ernst Fehr, Sergei Izmalkov, and Michael Kremer for their kindness and patience. I would especially like to thank Sergei Izmalkov for his infinite support in shaping my ideas and for his unwavering confidence. No matter how lonely the writing process felt, I always had at least one office with an open door.

Both of my experiments were run in Zurich at the Institute for Empirical Research at Zurich University. I am forever indebted to Ernst Fehr for granting me access to his lab and for financial assistance on the authority project. I would also like to thank Georg Hartmann for converting the experiments into German and for Holger Herz who helped design the authority experiment and oversaw the laboratory. Running experiments in a language that I did not know was a daunting task, but Holger's careful oversight put my mind at ease.

I have been blessed with an amazing set of friends and supporters. First, I wish to thank my housemates Tatiana Didier, Trang Nguyen, Chris Smith, and Eric Weese who put up with my bizarre hours and general insanity. I would also like to thank Greg Fischer, Sylvain Chassang, Raymond Guiteras, Jim Berry, Lorence Goette, Christian Zhander, and Patrick Warren for being both amazing friends and supportive colleagues. Finally I would like to thank Raphael Auer, Tal Gross, Vasia Panousi, Jose Tessada, Jeanne Lafortune, Sahar Parsa, Megan Palmer, Dave Appleyard, Johannes Spinnewijn, and Heikki Rantakari for never letting me completely disappear into my work.

Finally I would like to thank my family for their unconditional love. To my parents, who always catch me when I fall. To Jon and Anna, my oasis of understanding. To Katie, Mike, Rachael, Madaline, and Charlie, whose happiness is infectious. And to John Cox, Katie Larson, Nick Morgan, Melissa Austin, Steve Steinke, and Nick Hunter, those whose friendships have made me who I am today.

Contents

1 The Informational Properties of Institutions: An Experimental Study of Persistence in Markets with Certification 13

1.1 Introduction 14
1.2 Related Literature 17
1.3 The Model 18
1.3.1 Overview 18
1.3.2 Benchmark Model 19
1.3.3 Market Equilibria 25
1.4 The Experiment 30
1.4.1 Overview 31
1.4.2 Valuations and Trading Mechanism 32
1.4.3 Bonus Game 35
1.4.4 Treatments 35
1.4.5 Risk and Loss Aversion 38
1.4.6 Experimental Protocol 38
1.5 Experimental Results 39
1.5.1 Initial Convergence 41
1.5.2 Adaptation 44
1.5.3 Welfare 51
1.6 Conclusion 53
1.7 Appendix 54
1.7.1 Appendix 1: The Partial Certifying Equilibrium 54
1.7.2 Proofs 65
1.7.3 Instructions 68
1.7.4 Summary Statistics 72

2 Protecting Antiquities: A Role for Long-Term Leases? 75

2.1 Introduction 76
2.2 The Model 80
2.2.1 Timing 82
2.3 Aligning Incentives through Leases 83
2.3.1 Two Period Case 84
2.3.2 Infinite Horizon 85
2.3.3 The Minister of Culture's Problem 86
2.4 Transaction Costs and the Role of Leases 90

2.4.1 Asymmetric Information 91
2.4.2 Credit Constraints 96
2.5 Corruption and Intergenerational Conflict 98
2.6 Conclusion 100
2.7 Appendix 103
2.7.1 Appendix 1: Using Leases as an Information Rent 103
2.7.2 Appendix 2: Loss Aversion 105
2.7.3 Proofs 110

3 Experiments in Authority, Delegation, and Incentive Conflict 117
3.1 Introduction 118
3.2 The Model 121
3.2.1 Analysis and Theoretical Implications 123
3.2.2 Behavioral Forces 125
3.3 The Experiment 126
3.3.1 Protocol 127
3.3.2 The Principal-Agent Game 128
3.3.3 Treatments and Predictions 129
3.4 Experimental Results 132
3.4.1 Effort Choices and the Nash Equilibrium 133
3.4.2 Gender Differences in Search 138
3.4.3 Delegation Decisions 141
3.5 Conclusion 145
3.6 Appendix 147
3.6.1 Instructions (Translated from German) 147
3.6.2 Summary Statistics 152

Chapter 1

The Informational Properties of Institutions: An Experimental Study of Persistence in Markets with Certification

Abstract

This chapter 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.

1.1 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 chapter 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, thus favoring 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 chapter 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 chapter is organized as follows. Section 1.2 relates the current chapter to the literature. Section 1.3 builds the theoretical model and characterizes its competitive equilibria in terms of efficiency and information. Section 1.4 develops the experimental design. Section 1.5 reports the main experimental results and is divided into three parts. Section 1.5.1 looks at initial convergence of the experimental market in the safe and hazardous environments. Section 1.5.2 demonstrates the difference in adaptation between the non-certifying and certifying equilibrium. Section 1.5.3 looks at the welfare consequences of persistence in the certifying equilibrium and section

1.6 concludes.

1.2 Related Literature

Although the theory in this chapter 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.[1] 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 equilib-

[1] 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.

ria.[2] This chapter differs from the conventions literature in that it studies a specific channel by which history and information 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.

1.3 The Model

1.3.1 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.[3] 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 distribution 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 equilibrium when the trading mechanism is a double auction.[4] From

[2]For an overview of conventions see Ellison (1993). The convention literature has been experimentally explored in Van Huyck (1997) and Van Huyck (2001).

[3]For 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).

[4]See Walker & Williams (1988) for a discussion of convergence across varying institutions.

a theoretical perspective, simultaneous move double auctions also converge to the competitive equilibrium as the number 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.

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 *ex post* stable, a property that directly leads to a set of prices identical to those in one of the competitive equilibria.

1.3.2 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 $b \in \mathcal{B}$. 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 $m \in \mathcal{M} = \{\mathcal{C}, \mathcal{NC}, \varnothing\}$. $\mathcal{C}$ is a market for high quality certified units, $\mathcal{NC}$ is a market of uncertified units, and $\varnothing$ 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

$$C^H_B > C^H_C > C^L > C^H_G \tag{1.1}$$

and

$$C^H_B > C_L + U^H - U^L - T > C^H_C. \tag{1.2}$$

Condition 1.1 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 1.2 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 $\mathfrak{g}$, $\mathfrak{c}$, and $\mathfrak{b}$ respectively where $\mathfrak{g} = \frac{M_G}{M}$.

Definition 1 ***Degree of moral hazard:*** *The proportion of type-C sellers $\mathfrak{c}$.*

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 $C_B^H - C^L < U^H - U^L$ 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 $b \in \mathcal{B}$ who matches with a seller of type $s \in \{G, C, B\}$ in market $m \in \mathcal{M}$ 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 $\mathcal{NC}$, 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 $\mathcal{C}$, the conditional seller is constrained and would supply a high quality unit. In the baseline case:

$$u(m, P^m, b_0, s) = \begin{cases} U^H - P^{\mathcal{C}} & \text{if } m \in \mathcal{C}, s \in \{G, C, B\} \\ U^H - P^{\mathcal{NC}} & \text{if } m \in \mathcal{NC}, s \in \{G\} \\ U^L - P^{\mathcal{NC}} & \text{if } m \in \mathcal{NC}, s \in \{C, B\}. \end{cases} \tag{1.3}$$

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:

$$v(m, P^m, b, s) = \begin{cases} P^{\mathcal{C}} - C_s^H - T & \text{if } m \in \mathcal{C}, s \in \{G, C, B\}, \\ P^{\mathcal{NC}} - C_s^H & \text{if } m \in \mathcal{NC}, s \in \{G\}, \\ P^{\mathcal{NC}} - C^L & \text{if } m \in \mathcal{NC}, s \in \{C, B\}. \end{cases} \tag{1.4}$$

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 μ, 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 $\Sigma_{m \in \mathcal{M}} D(m, b) = N_b$. 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 $\Sigma_{m \in \mathcal{M}} S(m, s) = M_s$. An allocation (D, S) is *attainable* iff $\Sigma_{s \in \{G,C,B\}} S(m, s) = \Sigma_{b \in B} D(m, b)$ for $m \in \{\mathcal{C}, \mathcal{NC}\}$. Note that the market clearing condition is not binding in the $\varnothing$ market.

Belief System: Buyers and sellers form beliefs about the types of agents ex-changing 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 $\Sigma_s \mu_b(m,s) = 1$ for every m and $\mu_s(m,b) : M \times B \to \mathbb{R}_+$ satisfies $\Sigma_b \mu_s(m,b) = 1$ for every m.

Price System: A price system is a function $P : M \to \mathbb{R}_+$. I define $P^{\mathcal{C}}, P^{\mathcal{NC}}, P^{\varnothing}$ 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

$$\Sigma_s u(m, P^m, b, s)\mu_b(m,s), \tag{1.5}$$

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 (1.5). Consequently, an equilibrium allocation must assign all buyers of type b to markets that are in the arg max of equation (1.5):

$$D(m^*, b) \neq 0 \Leftrightarrow m^* \in \arg\max_m \Sigma_s u(m, P^m, b, s)\mu_b(m,s) \quad \forall b. \tag{1.6}$$

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

$$\Sigma_s v(m, P^m, b, s)\mu_s(m,b), \tag{1.7}$$

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:

$$S(m^*, s) \neq 0 \Leftrightarrow m^* \in \arg\max_m \Sigma_b v(m, P^m, b, s)\mu_s(m,b) \quad \forall s. \tag{1.8}$$

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 $\langle (D \times S), \mu, P \rangle$ consisting of an attainable allocation $(D \times S)$, beliefs μ, and a price system P which satisfy:*

$$E.1: \quad S(m^*, s) \neq 0 \Leftrightarrow m^* \in \arg\max_m \Sigma_b v(m, P^m, b, s)\mu_s(m, b) \qquad \forall s,$$

$$E.2: \quad D(m^*, b) \neq 0 \Leftrightarrow m^* \in \arg\max_m \Sigma_s u(m, P^m, b, s)\mu_b(m, s) \qquad \forall b,$$

$$E.3: \quad \mu_b(m, s) = \frac{S(m,s)}{\Sigma_s S(m,s)} \text{ and } \mu_s(m, b) = \frac{D(m,b)}{\Sigma_b D(m,b)}.$$

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

$$E.1b: \quad S(m^*, s) \neq 0 \Leftrightarrow m^* \in \arg\max_m \Sigma_b u(m, P^m, b, s) \qquad \forall s,$$

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 (1.3), 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 ΔP. Note that $\pi^H(\Delta P) = \mu_b(\mathcal{NC}, G)$ for $\Delta P = P^{\mathcal{C}} - P^{\mathcal{NC}}$.

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.

Supply Decisions by Sellers

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

$$v(\mathcal{C}, P^{\mathcal{C}}, b, s) > v(\mathcal{NC}, P^{\mathcal{NC}}, b, s). \tag{1.9}$$

For all sellers, this reduces to the condition

$$P^{\mathcal{C}} - C_s^H - T \geq P^{\mathcal{NC}} - min(C_s^H, C_L). \tag{1.10}$$

Define $\overline{P}^{\mathcal{C}}$ and as the maximum willingness to pay for a certified unit across all buyers. Similarly, define $\underline{P}^{\mathcal{NC}}$ as the minimum willingness to pay across all buyers for an uncertified unit. In the baseline model $\overline{P}^{\mathcal{C}} = U^H$ and $\underline{P}^{\mathcal{NC}} = U^L$. In equilibrium it will be the case that $\underline{P}^{\mathcal{NC}} \leq P^{\mathcal{NC}} \leq P^{\mathcal{C}} \leq \overline{P}^{\mathcal{C}}$ so that i) ΔP 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 $C_C^H \in (C^L, C^L + \overline{P}^{\mathcal{C}} - \underline{P}^{\mathcal{NC}} - T)$ 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 $C_B^H \geq C_L + \overline{P}^{\mathcal{C}} - \underline{P}^{\mathcal{NC}} - T$. 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 $\underline{P}^{\mathcal{NC}} \leq P^{\mathcal{NC}} \leq P^{\mathcal{C}} \leq \overline{P}^{\mathcal{C}}$:*

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

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

$$S(\mathcal{C}, B) = \begin{cases} 0 & \text{if } \Delta P \in (0, \overline{P}^{\mathcal{C}} - \underline{P}^{\mathcal{NC}}). \end{cases}$$

$S(\mathcal{NC}, s) = M_s - S(\mathcal{C}, s)$ for $s \in \{G, C, B\}$.

Proof. All proofs given in the appendix. ■

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

$$\pi^H(\Delta P)U^H + (1-\pi^H(\Delta P))U^L - P^{NC} = U^H - P^C. \tag{1.11}$$

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 $\Delta P < U^H - U^L \equiv \overline{P}^C - \underline{P}^{NC}$ 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 $\pi^H(\Delta P) = \mathfrak{g}$ and a risk neutral buyer prefers to purchase the uncertified unit as long as $\Delta P \geq (1-\mathfrak{g})(U^H - U^L)$.*

1.3.3 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 (1.11) 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:** $P^{NC} = U^H - (1-\mathfrak{g})(U^H - U^L), P^C = U^H$. 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 1.3.3 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* $(1-\mathfrak{g})(U^H - U^L) \leq T$.

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 $\Delta P = U^H - U^L = \overline{P}^C - \underline{P}^{NC}$. Type-$C$ sellers, defined as having $C_C^H - C_L < \overline{P}^C - \underline{P}^{NC} - T$, also choose to certify since the profit gained from switching markets is greater than the increase in production and transaction costs.

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:

$$M_C[U^H - U^L - (C_C^H - C^L)]. \tag{1.12}$$

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

$$(M_G + M_C)T. \tag{1.13}$$

Combining the two terms and normalizing by M yields:

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

$$(\mathfrak{g} + \mathfrak{c})T > \mathfrak{c}[U^H - U^L - (C_C^H - C^L)]. \tag{1.14}$$

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 $\mathfrak{c}$ 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.

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 $P^{NC} = U^H - (1-\mathfrak{g})(U^H - U^L)$ 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 sellers.*

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.

Loss Aversion, and Partial Certification and Public Information

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 $\mathcal{B} = \{\lambda_1, \lambda_2, \dots, \lambda_N\}$ where λ_i is the idiosyncratic loss aversion parameter for buyer i with $\lambda_i \geq 1$ for $i \in \{1, 2, ..., N\}$. Without loss of generality, I order buyers according to their risk aversion parameter such that $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_N$ and again normalize the utility obtained from not trading to zero.

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

$$u(m, P^m, \lambda_i, s) = \begin{cases} U^H - P^{\mathcal{C}} & \text{if } m \in \mathcal{C}, s \in \{G, C, B\} \\ U^H - P^{\mathcal{NC}} & \text{if } m \in \mathcal{NC}, s \in \{G\} \\ -\lambda_i(P^{\mathcal{NC}} - U^L) & \text{if } m \in \mathcal{NC}, s \in \{C, B\} \end{cases} . \tag{1.15}$$

In the non-certifying equilibrium, the market price $P^{\mathcal{NC}} > U^L$ 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 $S^{\mathcal{C}}$ be the number of certified units in an equilibrium. Then for each $S^{\mathcal{C}} < M_G$, a partial certifying equilibrium may exist with the following properties:

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

$$\mathfrak{g} \geq \underline{\mathfrak{g}} = \frac{M(U^H - U^L - T)\lambda_{M-S^{\mathcal{C}}} + S^{\mathcal{C}}T}{M(U^H - U^L - T)\lambda_{M-S^{\mathcal{C}}} + MT}. \tag{1.16}$$

$\underline{\mathfrak{g}}$ 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 $S^{\mathcal{C}} = 0$ and $\lambda_M = 0$, then equation (1.16) reduces to

$$\mathfrak{g} \geq \frac{(U^H - U^L - T)}{(U^H - U^L)} \tag{1.17}$$

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, $\underline{\mathfrak{g}}$ 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.

1.4 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 1.3.2.

1.4.1 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 ($\mathcal{S}$) environment where the degree of moral hazard was low and switched to a Hazardous ($\mathcal{H}$) 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.[6]

1.4.2 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).

[6]The loss aversion parameter will be covered in greater detail in sections 1.4.5 and 1.5.1

As shown in Table 1.1, 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 sellers were conditional, and type-B sellers sold uncertified low quality units only.

Table 1.1: Seller Costs

	Uncertified Low	Uncertified High	Certified High
G	50	30	90
C	50	80	140
B	50	130	190

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 1.2, 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 1.2: Buyer Valuations

Buyers 1-3				Buyers 4-5			
	Unit 1	Unit 2	Unit 3		Unit 1	Unit 2	Unit 3
High	240	220	200	High	230	210	190
Low	140	120	100	Low	130	110	90

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.

1.4.3 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.

1.4.4 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 ($\mathcal{S}$) environment where the degree of moral hazard was low and switched to a hazardous ($\mathcal{H}$) 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 1.3: Treatments

Treatment	Periods 1-12	Periods 13-24	Information	Identifiers
1	Safe	Hazardous	Private	$\mathcal{S}^{Pre}$,$\mathcal{H}^{Post}$
2	Safe	Hazardous	Public	$\mathcal{S}^{Pre}$,$\mathcal{H}^{Post}$
3	Hazardous	Safe	Private	$\mathcal{H}^{Pre}$,$\mathcal{S}^{Post}$
4	Hazardous	Safe	Public	$\mathcal{H}^{Pre}$,$\mathcal{S}^{Post}$

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.

Table 1.4: Moral Hazard Environments

	Good	Conditional	Bad
Safe ($\mathcal{S}$)	5	0	1
Hazardous ($\mathcal{H}$)	1	4	1

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 chapter 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 $.5U^H + .5U^L$ 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 $.5(U^H - U^L) = 50$. 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 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 $\Sigma U_i^H - \Sigma C_j^H = 2100$. 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 1.5: Efficiency

	Perfect Information	non-certifying	Certifying
Safe	2100*	2060	1460
Hazardous	1700*	1100*	1060

*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.

1.4.5 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 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 1.5.1, 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.

1.4.6 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.

1.5 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

[7]See Fischbacher (2007) for a description of Z-Tree.

formation and information is located in the appendix.

1. **Initial Convergence**: In section 1.5.1, 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 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 1.5.2, 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 1.5.3, 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 1.4.4, experimental sessions were divided into four groups based on the ordering of the $\mathcal{S}$ and $\mathcal{H}$ 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 1.3 is reprinted here.

Table 3: Treatments

Treatment	Periods 1-12	Periods 13-24	Information	Identifiers
1	Safe	Hazardous	Private	$\mathcal{S}^{Pre}$,$\mathcal{H}^{Post}$
2	Safe	Hazardous	Public	$\mathcal{S}^{Pre}$,$\mathcal{H}^{Post}$
3	Hazardous	Safe	Private	$\mathcal{H}^{Pre}$,$\mathcal{S}^{Post}$
4	Hazardous	Safe	Public	$\mathcal{H}^{Pre}$,$\mathcal{S}^{Post}$

1.5.1 Initial Convergence

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 environment 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 $\mathcal{S}^{Pre}$ environment where the degree of moral hazard is low with those in the $\mathcal{H}^{Pre}$ 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]

$$P_{i,s} = \alpha_0 + \Sigma\alpha_s + \beta_{Cert}I_{Cert} + \beta_{\mathcal{S}^{Pre}}I_{\mathcal{S}^{Pre}} + \epsilon_{i,s} \tag{1.18}$$

where $P_{i,s}$ is the price of an individual trade i in session s, α_s are individual session fixed effects, I_{Cert} is an indicator for a certified trade, and $I_{\mathcal{S}^{Pre}}$ 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 $\mathcal{S}^{Pre}$ environment, if the non-certifying equilibrium forms, the expected non-certifying price with no loss aversion

[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 $\beta_{\mathcal{S}^{Pre}}$.

[9]Note that since loss aversion only affects uncertified trades and the estimation includes both certified and uncertified trades, session level fixed affects do not eliminate the variation in uncertified trades across treatments.

is 183. Expecting the non-certifying equilibrium to form in the $\mathcal{S}^{Pre}$ environment and the certifying equilibrium to form in the $\mathcal{H}^{Pre}$ environment, I predict $\alpha_0 = 100$, $\alpha_0 + \beta_{Cert} = 200$ and $\alpha_0 + \beta_{\mathcal{S}^{Pre}} = 183$.

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 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:

$$P_{i,s} = \alpha_0 + \Sigma\alpha_s + \beta_{LA}(LA * I_{\mathcal{S}^{Pre}}) + \beta_{Cert}I_{Cert} + \beta_{\mathcal{S}^{Pre}}I_{\mathcal{S}^{Pre}} + \epsilon_{i,s} \tag{1.19}$$

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.

Results

Table 1.6 presents the price regressions from equation 1.18 and 1.19 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 ($\alpha_0 + \beta_{\mathcal{S}^{Pre}} = 147$) 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 1.7 compares the number of certified and uncertified trades made by loss averse buyers in the certified and uncertified markets. As expected, loss averse agents are more likely to buy from the certified market than their counterparts.

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

Table 1.6: Hypothesis 1: Convergence of *Pre* Treatments to the Non-Certifying or Certifying Equilibrium

	(1)	(2)	(3)
Certification	91.414***	91.414***	91.414***
	(2.816)	(2.816)	(2.820)
Treatment $I_{S^{Pre}}$	39.100***	60.867***	79.229***
	(4.685)	(8.122)	(7.905)
Risk Aversion of Buyer in $I_{S^{Pre}}$		−2.270*	−0.090
		(1.096)	(0.946)
Individual Loss Aversion in $I_{S^{Pre}}$		−20.827***	
		(6.097)	
Number of Loss Averse Buyers in $I_{S^{Pre}}$			−16.216***
			(2.113)[a]
Constant	108.496***	106.630***	105.349***
	(3.500)	(3.597)	(3.331)
Fixed Effects[b]	Yes	Yes	Yes
Adj. R^2	0.838	0.849	0.856
Observations (Trades in Period 7-12)	834	834	834

[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 β_{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 1.7: Transaction History Conditional on Loss Aversion

	Uncertified	Certified
Loss Averse	84	120
Normal	134	85
Fisher Exact Test $p < 0.01$		
Whitney-Mann-Wilcoxon$p = 0.2211$		

	Risky	Safe
Loss Averse	51	153
Normal	128	92
Fisher Exact Test $p < 0.01$		
Whitney-Mann-Wilcoxon $p = 0.0264$[a]		

[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.

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 1.7, loss averse individuals are significantly more likely to make safe transactions.

1.5.2 Adaptation

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 $\mathcal{S}^{Post}$ treatment should be the same as those from $\mathcal{H}^{Pre}$ and significantly differ from those in $\mathcal{S}^{Pre}$. I thus estimate:

$$\begin{aligned} P_{i,s} &= \alpha_0 + \Sigma\alpha_s + \beta_{LA}(LA * I_{\mathcal{S}^{Pre}}) + \beta_{Cert}I_{Cert} \\ &\quad + \beta_{\mathcal{S}^{Pre}}I_{\mathcal{S}^{Pre}} + \beta_{\mathcal{S}^{Post}}I_{\mathcal{S}^{Post}} + \beta_{\mathcal{H}^{Post}}I_{\mathcal{H}^{Post}} + \epsilon_{i,s}, \end{aligned} \tag{1.20}$$

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

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 $\mathcal{H}^{Pre}$ environment (where the certifying equilibrium was consistently observed), I expect to see no change in the composition of trades when the environment changes to $\mathcal{S}^{Post}$.

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 $\mathcal{S}^{Pre}$ 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 $\mathcal{H}^{Post}$.

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 $\mathcal{S}^{Pre}$ 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 1-1, 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 1-1 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 1-1 were chosen for clarity, the patterns of adaption and persistence are typical across all sessions.[10] Figure 1-2 shows average uncertified prices for the last six periods of each environment. Notice that the uncertified price in the $\mathcal{S}^{Post}$ environments is nearly identical to both the $\mathcal{H}^{Pre}$ and $\mathcal{H}^{Post}$ treatments and markedly different to the $\mathcal{S}^{Pre}$ 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 $\mathcal{S}^{Post}$ and $\mathcal{H}^{Post}$ environments relative to the baseline environment of $\mathcal{H}^{Pre}$.

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

[10]Summary statistics of the other treatments are available in the appendix. Additional data is available upon request.

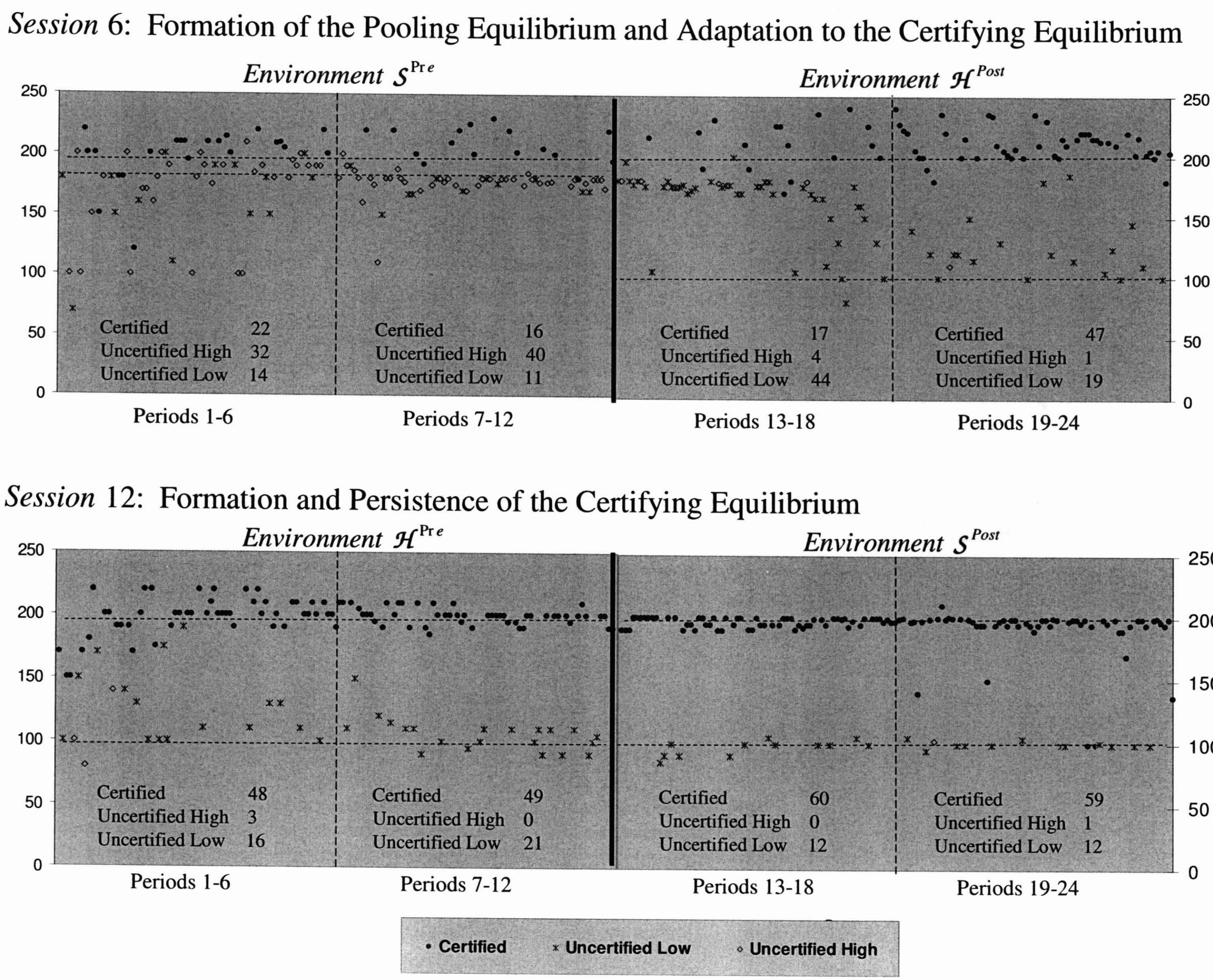


Figure 1-1: Hypothesis 2 — Persistence of the Certifying Equilibrium

Figure 1-2: Average Uncertified Prices by Environment

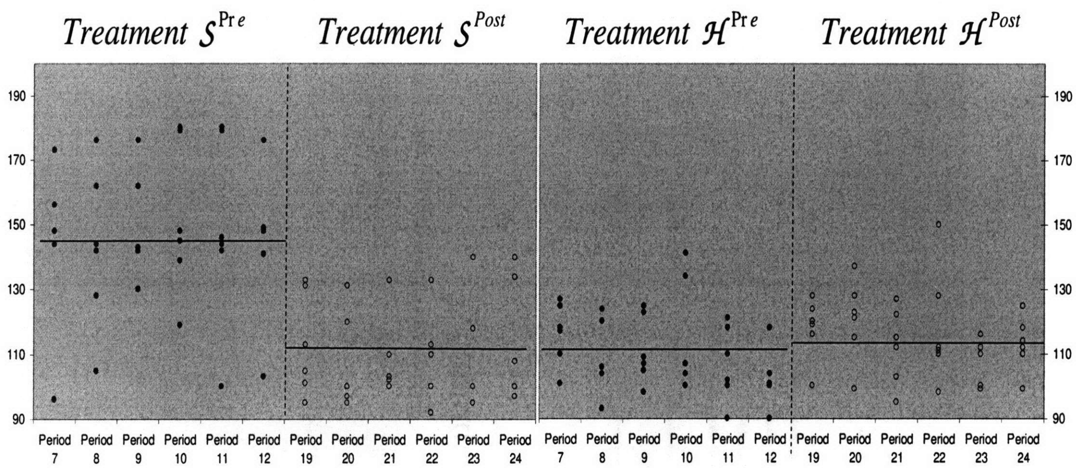


Table 1.8: Hypothesis 2: Persistence of the Certifying Equilibrium

	(1)	(2)
Certification	89.003***	89.003***
	(2.878)	(2.879)
Treatment $\mathcal{S}^{Pre}$	36.610***	75.948***
	(4.698)	(7.196)
Treatment $\mathcal{S}^{Post}$	2.776	2.776
	(3.721)	(3.722)
Treatment $\mathcal{H}^{Post}$	3.620	3.867
	(3.592)	(3.503)
# of Loss Averse Buyers in $\mathcal{S}^{Pre}$		–16.207***
		(2.528)
Constant	112.190***	110.616***
	(3.369)	(3.209)
Fixed Effects[a]	Yes	Yes
Adj. R^2	0.861	0.873
Observations	1675	1675

[a]Fixed Effects are at the session level. Standard Errors in parenthesis. Errors are clustered at the individual level.

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.

Table 1.9 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 1.10 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 1.9: Aggregate Trades in Treatments 1 & 2 (Sessions 1-6)

	Uncertified Red	Uncertified Blue	Certified Red
Last 6 periods of $\mathcal{S}^{Pre}$	135	95	208
First 6 periods of $\mathcal{H}^{Post}$	15	206	177
Last 6 periods of $\mathcal{H}^{Post}$	2	156	254

Table 1.10: Aggregate Trades in Treatments 3 & 4 (Sessions 7-12)

	Uncertified Red	Uncertified Blue	Certified Red
Last 6 periods of $\mathcal{H}^{Pre}$	6	150	255
First 6 periods of $\mathcal{S}^{Post}$	23	95	311
Last 6 periods of $\mathcal{S}^{Post}$	21	89	319

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 1-3 and 1-4 show significant compositional change in the non-certifying equilibrium and no change in the certifying equilibrium.

Apparent in Figure 1-3, 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 1-4, 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.

Figure 1-3: Hypothesis 4 — Composition of Trades: Treatments 1 & 2

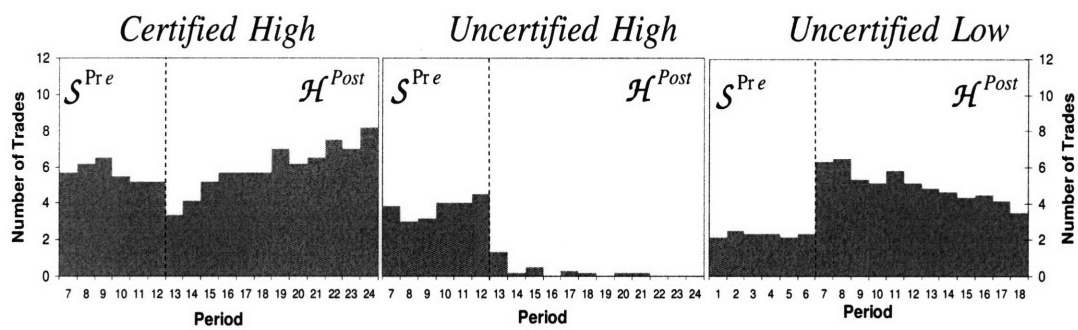
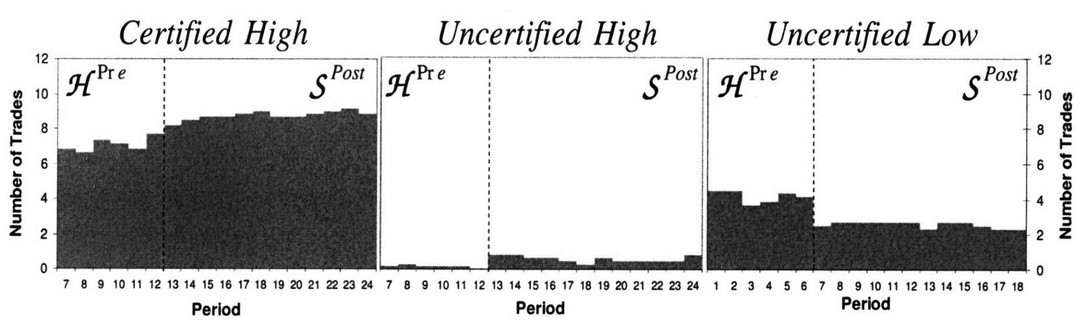


Figure 1-4: Hypothesis 3 — Composition of Trades: Treatments 3 & 4



1.5.3 Welfare

Hypothesis and Empirical Strategy

In the theory section of the chapter, 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 $\mathcal{S}^{Pre}$ environment and compare it to the efficiency of the last 6 periods in the $\mathcal{S}^{Post}$ environment. I estimate

$$\text{Efficiency}_{t,s} = \alpha_0 + \beta_{\mathcal{S}^{Pre}} I_{\mathcal{S}^{Pre}} + \beta_{LA}(LA * I_{\mathcal{S}^{Pre}}) + \beta_{LA}(LA * I_{\mathcal{S}^{Post}}) + \epsilon_{t,s} \quad (1.21)$$

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

Results

In figure 1-5, I compare period by period efficiency of the $\mathcal{S}^{Pre}$ periods with those of the $\mathcal{S}^{Post}$ treatment. The dashed horizontal line in the graph shows the predicted efficiency of a pure non-certifying equilibrium for the $\mathcal{S}^{Pre}$ treatment and the certifying equilibrium in the $\mathcal{S}^{Post}$ 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 $\mathcal{S}^{Pre}$. 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 1.11 shows a significant increase in efficiency in the $\mathcal{S}^{Pre}$ environment relative to $\mathcal{S}^{Post}$. 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 1-5: Hypothesis 5 — Efficiency Loss due to the Persistence of Certification

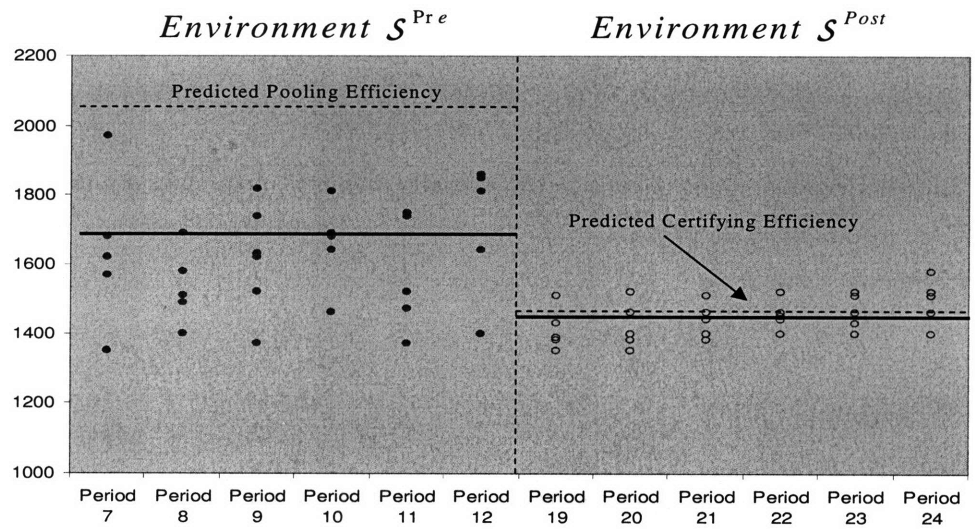


Table 1.11: Hypothesis 5: Efficiency Loss due to the Persistence of Certification

Treatment $\mathcal{S}^{Pre}$	330.010**
	(122.232)[a]
# Loss Averse Buyers in $\mathcal{S}^{Pre}$	-64.242
	(50.930)
# of Loss Averse Buyers in $\mathcal{S}^{Post}$	-2.073
	(3.568)
Constant	1456.707***
	(13.634)
Adj. R^2	0.439
Observations	72

[a]Errors are clustered at the session level

1.6 Conclusion

This chapter 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 chapter 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 chapter 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.

1.7 Appendix

1.7.1 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 1.7.1, 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 1.7.1, 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 1.7.1, 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.

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 $\mathcal{B} = \{\lambda_1, \lambda_2, \dots, \lambda_N\}$ where λ_i is the idiosyncratic loss aversion parameter for buyer i with $\lambda_i \geq 1$ for $i \in \{1, 2, ..., N\}$. Without loss of generality, I order buyers according to their risk aversion parameter such that $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_N$ and again normalize the utility obtained from not trading to zero.

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

$$u(m, P^m, \lambda_i, s) = \begin{cases} U^H - P^{\mathcal{C}} & \text{if } m \in \mathcal{C}, s \in \{G, C, B\} \\ U^H - P^{\mathcal{NC}} & \text{if } m \in \mathcal{NC}, s \in \{G\} \\ -\lambda_i(P^{\mathcal{NC}} - U^L) & \text{if } m \in \mathcal{NC}, s \in \{C, B\} \end{cases} . \tag{1.22}$$

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

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

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

$$P^{NC} \leq \frac{P^C - U^H}{\pi^H(\Delta P) + (1-\pi^H(\Delta P))\lambda_i} + \frac{\pi^H(\Delta P)U^H + (1-\pi^H(\Delta P))\lambda_i U^L}{\pi^H(\Delta P) + (1-\pi^H(\Delta P))\lambda_i}. \quad (1.23)$$

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 (1.23) reduces to

$$P^{NC} \leq \frac{\pi^H(\Delta P)U^H + (1-\pi^H(\Delta P))\lambda_i U^L}{\pi^H(\Delta P) + (1-\pi^H(\Delta P))\lambda_i}. \quad (1.24)$$

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 $\Delta P < U^H - U^L \equiv \overline{P}^C - \underline{P}^{NC}$ 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 $\pi^H(\Delta P) \doteq \mathfrak{g}$. A loss-averse buyer prefers to purchase the uncertified unit as long as inequality (1.24) is satisfied.
- If $\Delta P = T$, the buyer believes that all sellers that trade in the certified market are type G. Given that $\Sigma_s S(C,s)$ sellers trade in the certified market, $\pi^H(\Delta P) = \frac{M_G - \Sigma_s S(C,s)}{M - \Sigma_s S(C,s)}$ and a loss-averse buyer prefers to purchase the uncertified unit as long as inequality (1.24) is satisfied.

The demand decision of an individual buyer i is directly affected by his own loss aversion parameter λ_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 λ_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:

$$\begin{aligned}\widetilde{P}_M^{NC} &= \frac{\mathfrak{g}U^H + (1-\mathfrak{g})\lambda_M U^L}{\mathfrak{g} + (1-\mathfrak{g})\lambda_M},\\ \widetilde{P}_M^{C} &= U^H.\end{aligned}$$

If $\widetilde{P}_M^C - \widetilde{P}_M^{NC} \geq T$, 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:

$$\begin{aligned}\widetilde{P}_{M-1}^{NC} &= \frac{\widetilde{\mathfrak{g}}_{-1}U^H + (1-\widetilde{\mathfrak{g}}_{-1})\lambda_{M-1} U^L}{\widetilde{\mathfrak{g}}_{-1} + (1-\widetilde{\mathfrak{g}}_{-1})\lambda_{M-1}},\\ \widetilde{P}_{M-1}^{C} &= U^H,\\ \widetilde{\mathfrak{g}}_{-1} &= \frac{M_G - 1}{M-1}\end{aligned}$$

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

$$\begin{aligned}\widetilde{P}_{M-k}^{NC} &= \frac{\widetilde{\mathfrak{g}}_{-k}U^H + (1-\widetilde{\mathfrak{g}}_{-k})\lambda_{M-k} U^L}{\widetilde{\mathfrak{g}}_{-k} + (1-\widetilde{\mathfrak{g}}_{-k})\lambda_{M-k}},\\ \widetilde{P}_{M-k}^{C} &= U^H,\\ \widetilde{\mathfrak{g}}_{-k} &= \frac{M_G - k}{M-k}\end{aligned}$$

A partial certifying equilibrium occurs any time their exists a k such that $\widetilde{P}_{M-k}^C - \widetilde{P}_{M-k}^{NC} \leq T$ and $\widetilde{P}_{M-k+1}^C - \widetilde{P}_{M-k+1}^{NC} > T$:

Theorem 4 *Let k^* be the smallest k such that $\widetilde{P}_{M-k}^C - \widetilde{P}_{M-k}^{NC} \leq T$. 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 $i = \{1, ..., (M - k^*)\}$ 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.

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 1.3.2 and 1.7.1. I first build a single period game which guarantees *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 *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 chapter 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.[11]

Primitives

Consider an extension to the benchmark model in which all buyers and sellers are

[11] 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.

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 $\mathfrak{g}$.

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 buyers 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 $\Sigma_k p_0^i(\hat{g}_k) = 1$. 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 $P^{\mathcal{NC}}, P^{\mathcal{C}}, S^{\mathcal{NC}}, S^{\mathcal{C}}$, a buyer's posterior is

$$p_{t+1}^i(\hat{g}|h^1(t), h^2(t), \ldots, h^N(t), P^{\mathcal{NC}}, P^{\mathcal{C}}, S^{\mathcal{NC}}, S^{\mathcal{C}}) = p_{t+1}^i(\hat{g}|h^i(t)). \tag{1.25}$$

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 $\mu_b(m, s|\hat{g}_i, S^{\mathcal{NC}})$ 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 $S^{\mathcal{NC}}$ 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 $S^{\mathcal{NC}}$ has (perceived) expected utility:

$$EU_t^i(m|S^{NC}, p_t^i(\hat{g})) = \Sigma_k \Sigma_s u(m, P^m, b_{p_t^i(\hat{g})}, s)\mu_b(m, s|\hat{g}_k, S^{\mathcal{NC}})p_t^i(\hat{g}_k|S^{\mathcal{NC}}). \tag{1.26}$$

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**

[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 1.3.2

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

(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.[13] 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 $\beta_i^{\mathcal{NC}}(K)$, K units are traded at a price equal to $\beta_{K+1}^{\mathcal{NC}}(K)$. If $\beta_{K+1}^{\mathcal{NC}} < U^H - T$ and $\beta_K^{\mathcal{NC}} \geq U^H - T$, an exogenous reservation price $U^H - T$ is used.[14]

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

(a) **Stage 4:** Each buyer i' who did not purchase a unit in stage 3 submits a bid $\beta_{i'}^{\mathcal{C}}$ 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 $S^{\mathcal{C}}$ sellers enter the certified market, a Vickrey auction takes place where $S^{\mathcal{C}}$ units are traded at the reservation price of the $S^{\mathcal{C}}+1$ 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 biding 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.[15] Buyers

[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.

in the Stage 4-6 auction all receive zero utility and all remaining sellers in the market receive a utility of $U^H - T - C_s^H$.

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 $\beta_{K+1}^{\mathcal{NC}}(K) - min(C^L, C_s^H) > U^H - T - C_s^H$ 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 $S^{\mathcal{NC}} < M$ will have type-G sellers being the first to exit from the uncertified market. It follows, that conditional on $\hat{g}_k$ and $S^{\mathcal{NC}}$, a buyers belief about the proportion of good types in the environment is

$$\mu_b(\mathcal{NC}, G|\hat{g}_k, S^{\mathcal{NC}}) = max\Big(\frac{M\hat{g}_k - M + S^{\mathcal{NC}}}{S^{\mathcal{NC}}}, 0\Big). \tag{1.27}$$

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

4. A buyers perceived expected value for bidding in the uncertified market conditional on $S^{\mathcal{NC}}$ is $EU_t^i(m|S^{NC}, p_t^i(\hat{g}))$. 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 buyers 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:
$$p_{t+1}(\hat{g}_i|x) = \frac{p_t(\hat{g}_i)q(x|\hat{g}_i)}{\Sigma_k p_t(g_k)q(x|\hat{g}_k)}$$
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 $\frac{S^C}{M}$ be the share of sellers who trade in the certified market. With probability $\frac{S^C}{M}$, a buyer ignores his own signal and updates as if he received a high value units. With probability $1 - \frac{S^C}{M}$, 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 $\frac{S^C}{M}M_G + \frac{M-S^C}{M}\frac{M_G-S^C}{M-S^C} = \frac{M_G}{M} = \mathfrak{g}$. 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 must assume 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*
$$p_t^i(\hat{g}) \xrightarrow{a.s.} \mathfrak{g}. \tag{1.28}$$

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:*
$$p_t^i(\hat{g}) \xrightarrow{a.s.} \mathfrak{g} \tag{1.29}$$

3. *If the certifying equilibrium occurs at time t^*, then*
$$p_t^i(\hat{g}) = p_{t^*}^i(\hat{g}) \qquad \forall t > t^* \tag{1.30}$$

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:*** $P^{NC} \leq U^H - (1 - \mathfrak{g})(U^H - U^L)$.

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.

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

To test this hypothesize, 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 $\mathcal{S}^{Pre}$ treatment since this is the only treatment where the information treatment should matter and estimate:

$$\begin{aligned} Risky_{i,s} &= \alpha_0 + \Sigma\alpha_s + \beta_{LA_i} I_{LossAverse} + \beta_{Beliefs} Beliefs + \\ &\quad + \beta_{Public}(I_{Public} * I_{\mathcal{S}^{Pre}}) + \beta_{Public*LN}(I_{LossNeutral} * I_{Public} * I_{\mathcal{S}^{Pre}}) + \epsilon \end{aligned} \tag{1.31}$$

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

Results

Evidence from the experimental sessions shows a clear effect of information on the proportion of uncertified trades in the environment. Figure 1-6 shows the proportion of risky trades undertaken by loss neutral and loss averse individuals divided between the $\mathcal{S}^{Pre}$ 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 postive 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 1.12.[18]

Figure 1-6: Effect of Information on the Propensity to Purchase Risky Units

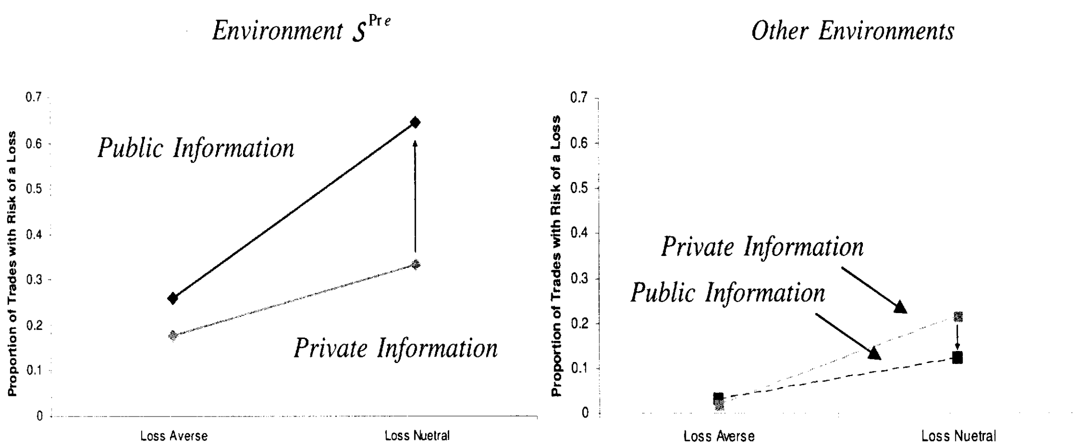


Table 1.12: Probit Regression of Risky Purchases on Information and Loss Aversion

Loss Aversion	–0.568***
	(0.069)
Public * $\mathcal{S}^{Pre}$	0.349***
	(0.130)
Loss Neutral * Public * $\mathcal{S}^{Pre}$	0.480***
	(0.151)
Beliefs	0.005
	(0.017)
Constant	–0.835***
	(0.119)
Psuedo Adj. R^2	.216
Observations	3353

[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.

1.7.2 Proofs

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

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

By Definition 2,

$$S(m^*, s) \neq 0 \Leftrightarrow m^* \in \arg\max_m \Sigma_b v(m, P^m, b, s) \quad \forall s.$$

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

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

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

It follows:

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

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

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

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

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

   A buyer is indifferent across all three markets if $P^{\mathcal{NC}} = U^H - (1 - \mathfrak{g})(U^H - U^L)$ and $P^{\mathcal{C}} = U^H$.

■

Proof.

1. When $\Delta P = U^H - U^L$:

   (a) By Lemma 1, $S(\mathcal{NC}, B) = M_B$, $S(\mathcal{C}, G) = M_G$, and $S(\mathcal{C}, C) = M_C$.

   (b) By Lemma 2, if $P^{\mathcal{NC}} = U^H, P^{\mathcal{C}} = U^L$, $D(\mathcal{C}, b_0) = [0, N_{b_0}] \in \mathbb{I}_+$, $D(\mathcal{NC}, b_0) = [0, N_{b_0}] \in \mathbb{I}_+$, $D(\varnothing, b_0) = [0, N_{b_0}] \in \mathbb{I}_+$ with $\Sigma_m D(m, b_0) = N_{b_0}$.

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

2. When $\Delta P > T$:

   (a) By Lemma 1, $S(\mathcal{NC}, B) = M_B$, $S(\mathcal{NC}, G) = M_G$, and $S(\mathcal{NC}, C) = M_B$.

   (b) By Lemma 2, a buyer is indifferent between all three markets if $P^{\mathcal{NC}} = U^H - (1-g)(U^H - U^L)$ and $P^{\mathcal{C}} = U^H$.

   If $P^{\mathcal{C}} - P^{\mathcal{NC}} = (1-g)(U^H - U^L) > T$, then $D(\mathcal{NC}, b_0) = M, D(\varnothing, b_0) = N_{b_0} - M$ 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. By Lemma 1, $S(\mathcal{C}, C) > 0 \to S(\mathcal{C}, G) = M_G$. It follows that for any competitive equilibrium where type-C sellers certify their good, all type-G sellers certify their good. Define

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

Since $u(\mathcal{C}, P^{\mathcal{C}}, b_0, G) + v(\mathcal{C}, P^{\mathcal{C}}, b_0, G) - T \leq u(\mathcal{NC}, P^{\mathcal{NC}}, b_0, G) + v(\mathcal{NC}, P^{\mathcal{NC}}, b_0, G)$, W is decreasing in $S(\mathcal{C}, G)$. Likewise, since $u(\mathcal{C}, P^{\mathcal{C}}, b_0, C) + v(\mathcal{C}, P^{\mathcal{C}}, b_0, C) - T \geq u(\mathcal{NC}, P^{\mathcal{NC}}, b_0, C) + v(\mathcal{NC}, P^{\mathcal{NC}}, b_0, C)$, W is increasing in $S(\mathcal{C}, C)$. 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. Let $\hat{\mathfrak{g}}$, $\hat{\mathfrak{c}}$, $\hat{\mathfrak{b}}$ be the prior beliefs about the proportion of good, conditional and bad agents in the market. When a non-certifying equilibrium exists, $P^{\mathcal{NC}} = U^H - (1-\mathfrak{g})(U^H - U^L)$. When a certifying equilibrium exists, only type-B sellers are in the uncertified market.

Define $S^{\mathcal{NC}}$ 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 $\hat{\mathfrak{g}}$, $\hat{\mathfrak{c}}$, $\hat{\mathfrak{b}}$ under the non-certifying and certifying equilibrium are as follows:

Type	Non-Certifying Equilibrium	Certifying Equilibrium
$\hat{\mathfrak{g}}$	$\frac{P^{NC}-U^L}{U^H-U^L}$	$\frac{\bar{\mathfrak{g}}}{\bar{\mathfrak{g}}+\bar{\mathfrak{c}}}(1-\hat{\mathfrak{b}})$
$\hat{\mathfrak{c}}$	$\frac{\bar{\mathfrak{c}}}{\bar{\mathfrak{c}}+\bar{\mathfrak{b}}}(1-\hat{\mathfrak{c}})$	$\frac{\bar{\mathfrak{c}}}{\bar{\mathfrak{g}}+\bar{\mathfrak{c}}}(1-\hat{\mathfrak{b}})$
$\hat{\mathfrak{b}}$	$\frac{\bar{\mathfrak{b}}}{\bar{\mathfrak{c}}+\bar{\mathfrak{b}}}(1-\hat{\mathfrak{c}})$	$S^{\mathcal{NC}}/M$

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

Proof. For a partial certifying equilibrium to exist, it must be the case that at $\Delta P = T$, $\Sigma_b D(\mathcal{C}, b) = \Sigma_s S(\mathcal{C}, s)$ and $\Sigma_b D(\mathcal{NC}, b) = \Sigma_s S(\mathcal{NC}, s)$. 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 λ_i and λ_{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:

$$\begin{aligned}
\widetilde{P}^{\mathcal{NC}}_{M-k} &= \frac{\widetilde{\mathfrak{g}}_{-k}U^H + (1-\widetilde{\mathfrak{g}}_{-k})\lambda_{M-k}U^L}{\widetilde{\mathfrak{g}}_{-k} + (1-\widetilde{\mathfrak{g}}_{-k})\lambda_{M-k}}, \\
\widetilde{P}^{\mathcal{C}}_{M-k} &= U^H, \\
\widetilde{\mathfrak{g}}_{-k} &= \frac{M_G - k}{M-k}
\end{aligned}$$

If $\widetilde{P}^{\mathcal{C}}_{M-k} - \widetilde{P}^{\mathcal{NC}}_{M-k} > T$, $D(\mathcal{NC}, \lambda_k) = 0$. 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 $\widetilde{P}^{\mathcal{C}}_{M-k} - \widetilde{P}^{\mathcal{NC}}_{M-k} > T$ and $\widetilde{P}^{\mathcal{C}}_{M-k+1} - \widetilde{P}^{\mathcal{NC}}_{M-k+1} \leq T$, 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.

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 $t \rightarrow \infty$, all periods leading up to the current period must all have had the non-certifying equilibrium form.

Step 2 Let $\mathbf{x} = (x_1, \ldots, x_T)$ 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 $p_0^i(\hat{g}) = \{p_0^i(\hat{g}_0), p_0^i(\hat{g}_1), \ldots, p_0^i(\hat{g}_{M-1})\}$ where $p_0^i(\hat{g}_k) > 0$ and $\Sigma_k p_0^i(\hat{g}_k) = 1$, the posterior $p_t(\hat{g}|\mathbf{x})$ converges almost surely to the true proportion as $t \rightarrow \infty$ as long as $\mathfrak{g} \in \hat{g}$ and

$$\Sigma_x q(x|\hat{g}_i) log \left[\frac{q(x|\hat{g}_i)}{q(x|\hat{g}_j)}\right] > 0. \tag{1.32}$$

Expanding condition (1.32) yields:

$$\hat{g}_i log\left(\frac{\hat{g}_i}{\hat{g}_j}\right) + (1 - \hat{g}_i) log\left(\frac{1 - \hat{g}_i}{1 - \hat{g}_j}\right) \tag{1.33}$$

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 (1.32) holds. Since $\mathfrak{g} \in \{0, \frac{1}{M}, \ldots, \frac{M-1}{M}\}$, convergence is guaranteed as $t \rightarrow \infty$.

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.

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 $t \rightarrow \infty$, there exists at least M buyers who get infinite draws. For these buyers, the willingness to pay for an uncertified unit converges to $\mathfrak{g}U^H + (1 - \mathfrak{g})U^L$. If a buyer has a higher value than $\mathfrak{g}U^H + (1 - \mathfrak{g})U^L$, 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 $P^{NC} = \mathfrak{g}U^H + (1 - \mathfrak{g})U^L$. ■

1.7.3 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.

   <u>**Your offer to sell:**</u>
   - 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".

   <u>**The offers from buyers:**</u>
   - 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

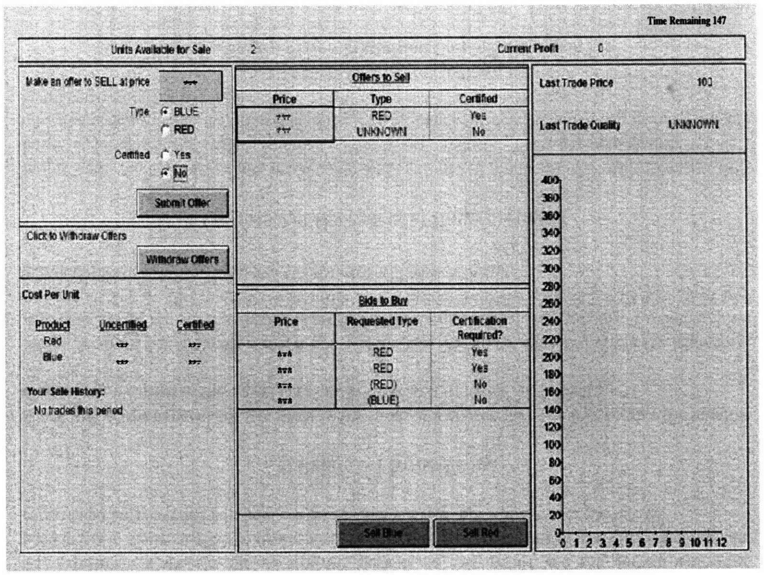
During the experiment you will enter your decisions using the computer. In the following instructions, all the functions will be explained in detail.

<u>**1. The Trading Phase**</u>

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 **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 three 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:

Quality	Costs without certification	Costs with certification
RED	30	30+60 = 90
BLUE	50	50 + 60 = 110

Quality	Costs without certification	Costs with certification
RED	80	80 + 60 = 140
BLUE	50	50 + 60 = 110

Case 2b, BLUE Quality is cheaper to produce

Quality	Costs without certification	Costs with certification
RED	130	130 + 60 = 190
BLUE	50	50 + 60 = 110

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 ≤ Price ≤ 400

- You have to specify the quality:

Quality = RED or BLUE

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

Certificate = Yes or No
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:

Quality	Price	Certified
RED	400	Yes
BLUE	50	No

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:

Quality	Price	Certified
RED	**350**	Yes
BLUE	50	No

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:

Quality	Price	Certified
RED	350	Yes
RED	**250**	No

To withdraw offer 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.

History

On the bottom left side of the screen, you will see your personal history. There you will see detailed information about the products you have sold so far during the respective period. For every product purchased you will see:

- the quality
- whether the product was certified
- the price you got
- the resulting earnings

On the right side of the screen you will see the market history. On the top you will find the information of the last traded good. Below you find a chart with all the trades of the period.
On the axis to the right you will find the amount of products traded. On the other axis you will find the price that has been paid for the product. Depending on the quality and certification of the product, the entry is of a different color:

- RED certified products appear in **red**
- BLUE certified products appear in **blue**
- Uncertified products appear in **black**

2. The bonus phase

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 showupfee will be paid out in cash at the end of the experiment.

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

Exercises

The experiment starts only after all participants are fully accustomed with the experiment. To ensure this, we ask you to solve the exercises on this page.

Please also write down intermediary steps.

After these exercises you will have the possibility to get to know the trading screen before the first period starts. The options you have will be presented again in detail and you can do some trial trades.

For these exercises please use the following cost structure:

	Cost without certification	Cost with certification
ROT	80	140
BLAU	60	120

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:

Offer number	Price	Quality	Certificate requested
1	220	BLUE	Yes
2	180	RED	No

Through which sale can you make the higher earnings?

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

1.7.4 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 1.13: Summary Statistics for $\mathcal{S}^{Pre}$ Environment

Session	Uncertified Price	Certified Price	Uncertified High Quality	Uncertified Low Quality	Certified High Quality	Number Loss Averse
1	141	194	22	12	28	3
2	170	203	19	13	40	1
3	103	193	1	17	53	4
4	145	202	19	12	39	2
5	143	197	34	18	19	3
6	176	209	40	11	16	2

Table 1.14: Summary Statistics for $\mathcal{H}^{Post}$ Environment

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

Table 1.15: Summary Statistics for $\mathcal{H}^{Pre}$ Environment

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

Table 1.16: Summary Statistics for $\mathcal{S}^{Post}$ Environment

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

Chapter 2

Protecting Antiquities: A Role for Long-Term Leases?

Abstract

Most countries prohibit the export of certain antiquities. This practice often leads to illegal excavation and looting for the black market, which damages the items and destroys important aspects of the archaeological record. We argue that long-term leases of antiquities or sales contracts with an option to buy the object back at a prearranged price would raise revenue for the country of origin while preserving national long-term ownership rights. By putting antiquities into the hands of the highest value consumer in each period, allowing leases would generate incentives for the protection of objects.

2.1 Introduction

As part of an effort to preserve their cultural heritage, 140 countries ban the export of certain antiquities. One side effect of these export bans is a black market in antiquities. Artifacts often have a greater monetary value outside their country of origin, especially if that country is poor.[1] Because of absent legal markets and weak enforcement, owners turn to black markets to sell objects abroad.

Illegal trade is surreptitious, and actions that conceal antiquity trade often destroy archeological sites, damage objects, and reduce economic value. Looters use fast methods of excavation such as bulldozers, dynamite, and pneumatic drills.[2] They work to keep site locations secret and often disguise the origin of objects by intentionally damaging sites to camouflage their activities and breaking objects into fragments to pass international borders.[3] When objects are traded illegally, and therefore surreptitiously, it is difficult to both search for and extract rent from the highest value buyer. The value to many potential buyers may be reduced because of limitations on the ability to display the object and because of danger of detection and prosecution. These factors reduce the price for the object relative to what sellers would obtain under legal trade.[4]

We argue that compared to complete export bans, allowing lease markets could raise revenue for artifact-rich countries and create incentives for maintenance and preservation, while keeping long-term ownership rights for the country of origin. By putting the object in the hands of the highest value consumer at each point in time, leases would reduce the size of the black market and generate funds that could be used for the legal excavation of at-risk sites or other needs. Since future ownership rights are preserved, a country could manage its cultural heritage without restricting objects from flowing to highest value use.

As one example of an environment where leases are likely to be useful, consider

[1] An Italian antiquities trafficker was recently caught offering Hellenistic marble statues of Marsyas and Apollo for $850,000. The statues were originally purchased from a Turkish farmer for $7,000. See Bagli (1993) and Borodkin (1995).

[2] See Coggins (1972), Bator (1982), and Prott & O'Keefe (1990) for many examples.

[3] Ross (1995) estimates that over 50 percent of archaeological sites in Mali have been severely damaged or destroyed by illegal looting. Archeologists, who rely on the stratification of objects to make inferences, have limited access to pristine sites and no access to objects already extracted illegally. Owners of artifacts have no legal channel by which they can return objects back to the public domain, leading to limited knowledge concerning the number of objects still existing and complicating arrangements that might lead to repatriation.

[4] Christie's Auction House estimates that the original owners of artifacts typically receive 2 percent of the object's final sale price. See Beech (2003).

the case of Nigeria. In the last three decades, at least seven of the major museums in Nigeria have been the victims of major robberies; the most notable occurring at the Ife museum in 1994 with estimated loses of 200-250 million dollars.[5] Illegal excavation of archeological sites has also escalated with the most significant losses occurring from 'Nok' sites in northern Nigeria. Nigeria's museum system has struggled to generate funds to maintain security and preserve objects. Tourism revenue for major exhibits is limited and the total 2008 public budget for museum and monument preservation was under 16 million dollars. We believe that a lease program that rotates a portion of Nigeria's collections internationally would generate revenue and reduce the potential of theft by moving at risk objects abroad.

In this paper, we look at the antiquities market from the perspective of a hypothetical Minister of Culture, a government agent who is attempting to manage the cultural heritage of his country. Our main approach is to first explore potential reasons why a benevolent Minister of Culture might choose to impose an export ban when faced with a choice set consisting only of an export ban or free market. We then expand the set of possible government policies to show that many of the governments objectives may be better achieved by restricting transactions to leases or sales contracts with an explicit option to repurchase.

We begin by considering the optimal policy for the Minister of Culture who is attempting to maximize social surplus when an object is in the hands of a private individual who has property rights over the object but must invest in maintenance to prevent its destruction. When the object is intact and within the nation's borders, the object generates a positive externality which the owner of the object does not take into account. We show that under free trade, a rich owner of an object will have incentives to use an object locally in all periods, while a poor owner may have incentives to sell the object to a foreign collector outright. For owners with a moderate level of initial wealth, the optimal policy is to share usage rights intertemporally with a foreign collector, which can be achieved by a lease contract. Under free trade, owners will invest in maintaining the object while under export bans, owners with sufficiently low initial wealth will not invest in maintenance leading to the object's destruction.

Given the actions of the owner, export bans may be effective at realigning incentives for wealthier countries, but may lead to inadequate maintenance, black markets, and the permanent loss of art in places where owners are poor. When taxes can be

[5] See Akinade (1999) and Shyloon (2000).

imposed costlessly, it is possible to obtain Pareto optimal allocations by using subsidies to keep antiquities intact and in the country. However, when taxes are inefficient, quantity constraints that limit the amount of time an object can leave the country may be second best. Allowing for intertemporal sharing through a lease contract may increase home usage relative to a pure export ban by generating income and strengthening maintenance incentives.

Next we examine a model with asymmetric information regarding the value that foreign collectors and the Minister of Culture assign to an object. If a country is initially poor but may become rich later, it may be optimal to initially transfer usage rights to a foreign collector, but for the artifact to return to the country of origin if the home country becomes wealthier. If the Minister of Culture is fairly certain that it will want the object in the long run, but its future value is private information, sale and repurchase contracts may be inefficient, since attempts by foreign collectors to extract surplus from the government may prevent efficient transactions. Either leases or a sales contracts with an option to buy the object back at a prearranged price may help avoid this hold up problem. In a world without credit constraints, leases dominate both sale and repurchase contracts and option contracts since negotiation occurs after the resolution of uncertainty and with the home country in control of the auction. With credit constraints, allowing for options can increase liquidity of the private owners thereby increasing their overall utility.

Finally, we introduce the probability that the Minister of Culture in each generation is corrupt and extracts value from antiquities by selling them abroad at the expense of future generations. We show that in this environment, laws imposing export bans may be preferable to free trade. In an effort to constrain the bad types, a benevolent Minister of Culture may create legislation which limits both their actions and those of future generations. For reasonable parameter values, allowing leases may be preferable to either free trade or complete export bans. Leases prevent the expropriation of value from future generations while still granting freedom to optimally allocate usage rights today.

Since the influential paper by Merryman (1986), the debate over cultural property has often been expressed in terms of cultural nationalism and cultural internationalism. Cultural nationalists, such as Greenfield (1996), emphasize the importance of local ownership and alienation of objects sold abroad. Cultural internationalists, such as Appiah (2006), stress the value of international access and cultural awareness. We see the lease approach as helping fulfill many of the goals of both viewpoints while reducing looting and increasing maintenance incentives.

Writing from a legal perspective, Bator (1982) and Borodkin (1995) advocate the use of legal markets to reduce looting in developing countries. These papers concentrate on the increased information that is revealed when markets are legal and when individuals have private information and partial ownership claims over the location of objects still in the ground. As a way of generating stronger incentives for individuals to report found sites, Wendel (2007) advocates the use of "possessory estates,"a shared trust set up by the government that grants a proportion of proceeds to an individual who reports the location of an antiquity site. Our paper is complementary to these analyses, focusing on the efficiency and preservation incentives generated by different contract structures.

Lease contracts have been briefly mentioned in press by Butcher & Gill (1990), Asgari[6], and Gerstenblith (2001). In all three of these articles, leases are proposed as a way to move objects between museums in order to decrease demand for new pieces from foreign countries. We believe our paper is the first to formally model the effects of export bans and lease markets and to suggest leases and option contracts as a broad alternative to export bans.

The idea that export bans can constrain dictators is related to Kremer & Jayachandran (2006) and Pogge (2001), who address the potential of dictators to expropriate wealth from future generations by entering debt contracts or selling natural resources. Our results on the optimal contract structure is related to Hart & Tirole (1988) and Dewatripont (1989), who study short- and long-term lease contracts when future valuations are known but private.

There is precedent for art being leased to cross international borders. The King Tut exhibit which circulated in the United States and London from 2005-2008 was leased to a private company in order to generate proceeds for Egypt.[7] Leases have also recently been used to resolve disputes over ownership. The Menil collection in Houston negotiated with the Church of Cyprus a 25-year lease of two 13th century Byzantine frescoes it recovered in 1982 from sources with disputed claims. More recently, the Metropolitan agreed to return a collection of objects believed to be looted from Italy in exchange for a long-term loan of objects with similar value.

The paper is organized as follows. Section 2.2 lays out the notation used in the

[6]Asgari argues that ten year leases may be used between major museums to reduce incentives to purchase illicit artifacts. See Erdem (1993).

[7]The lease agreements for the King Tut exhibit specified transportation, display, and storage conditions in order to reduce moral hazard. Egypt charged a flat fee of $5 million dollars per city and required insurance of roughly $1 million dollars per city. The exhibit was valued at $650 million dollars. See Boehn (2005).

paper and the assumptions of the model. Section 2.3 develops our benchmark model in an environment where there is an externality for home use. In Section 2.4, we relax the information assumptions to study how potential holdup may lead to inefficiency in sales contracts, which can be resolved via lease contracts. Finally, in Section 2.5 we study how leases may be superior to sales contracts and export bans in an environment with intergenerational corruption. Section 2.6 concludes.

2.2 The Model

Our model considers a single antiquity, referred to as the object. There are $N + 2$ actors — the domestic owner of the object, the Minister of Culture, and N foreign collectors. Initially the antiquity is in the hands of the domestic owner who must decide whether to maintain the object at home, sell it abroad, or let it be destroyed. The Minister of Culture cannot directly control the actions of the owner, but has influence over his decision by setting domestic policy.

Preferences—At each time $t \in \{0, 1, \ldots, T\}$ (with the possibility that $T = \infty$), the owner of the object gets utility from non-antiquity consumption, c_t, and domestic antiquity usage x_t. To distinguish between the two types of goods without additional qualifiers, we refer to c_t as "consumption" and x_t as "usage" throughout the paper. We assume that the agent's utility function for consumption and usage is additively separable, $U_t(c_t, x_t) = u(c_t) + D_O x_t$, where D_O is the usage value of the object to the domestic owner, and $x_t \in \{0, 1\}$ is a binary variable that is 1 when the object is held by the owner. We assume that the marginal utility of consumption is decreasing with $u'(c) > 0$,$u''(c) < 0$ and that $u'(0) = \infty$, $u'(\infty) = 0$.

The Minister of Culture is a benevolent government official that would like to maximize social welfare for the domestic country but can not directly affect the actions of the owner. Social welfare in each period is defined as $S(c_t, x_t) = u(c_t) + [D_O + D_{E,t}]x_t$ where $D_{E,t}$ is the domestic externality for an object to remain intact and within the country's borders. We assume that $D_{E,t}$ is stochastic and that $D_O + D_{E,t}$ be bounded between bounded between $\underline{D}$ and $\overline{D}$ and distributed according to a cdf $H_t(.)$ which has an increasing hazard rate.

There are $i = 1, \ldots, N$ foreign collectors who are interested in legally using the object. Each foreign collector has a private value for art consumption F^i bounded between $\underline{F}$ and $\overline{F}$ and distributed according to a cdf $A(.)$ which is constant over time, and has an increasing hazard rate. All foreign collectors are rich so that their utility for consumption is approximately linear for the potential changes in consumption we are considering. We thus assume that $V_t^i(c_t, z_t) = c_t + F^i z_t$ where $z_t \in \{0, 1\}$ is a binary variable that is 1 when the foreign collector legally keeps the object abroad. Without loss of generality, we assume that the buyers are ordered in ascending value. Thus F^N and F^{N-1} represent the highest and second highest values respectively.

All actors in our model share a common intertemporal discount rate of $\delta \in [\frac{1}{2}, 1]$ and are subject to identical interest rates R. We assume $\delta R = 1$. This ensures that the domestic owner of an object prefers flat consumption over time and is indifferent in the combination of periods for which domestic usage of the antiquity takes place.

Maintenance—Preserving art requires expenditure M at the beginning of each period to maintain the object. We consider M to be a reduced form parameter that includes the cost of preventing damage and theft by looters. While in reality, M is best represented by a continuous variable that influences the probability and severity of loss, we make the stark assumption that M is binary and that if it is not paid, the artifact is immediately destroyed.[8] When an object is destroyed, both home usage, x_t, and legal foreign usage, z_t, are equal to zero. Since destruction is permanent, we require that $x_{t+1} + z_{t+1} \leq x_t + z_t$ for all t. We assume that $\underline{F} > M$ so that all foreign collectors value the object more than its maintenance cost.

In the infinite horizon representation of our problem, it will be convenient to represent some of our results in terms of $\pi_D = [1 - \delta] \sum_{t=0}^{\infty} \delta^t x_t$, the proportion of time that an object is used domestically after adjusting for the discount rate. It is shown in the appendix, that when $\delta \geq \frac{1}{2}$, there exists a combination of $x_0, \ldots, x_\infty$ that corresponds to any value of π_D in $[0, 1]$. Reformulating problems using this equivalent representation allows us to solve problems in terms of the continuous variable π_D instead of an infinite set of binary variables $x_0, \ldots, x_\infty$.[9]

Budget Constraints—The owner of an object has initial assets ω_0 and receives a deterministic endowment shock of ω_t in all other periods. We take wealth as exogenous, but it is worth noting that depending on how well markets function inside the country, the object may wind up in the hands of the highest value domestic owner.

[8]Extending the model to a more general moral hazard environment is straightforward and does not substantively change the analysis. Part of the reason we have chosen a binary representation is that with a minor redefinition of variables, M may also include opportunity costs such as the revenue passed up from not selling an object to a smuggler. In this case, the cost of maintenance is not likely to be convex and the optimal maintenance decision will inherently be binary in nature.

[9]As $\delta \to 1$, we can also get arbitrarily close to a continuous $[0, 1]$ variable using the first T^* periods instead of an arbitrary set of periods in $x_0, \ldots, x_\infty$. When we study export bans, we will use this formulation in order to avoid violating the maintenance constraints.

This will typically be someone rich.[10] Letting b_t be the amount borrowed in period t and P_t being the price if an object is legally traded abroad, the owners budget constraints are:

$$\begin{aligned} c_1 + x_1 M &= b_1 + z_1[P_1 - M] + \omega_1 & for\ t = 1, \\ c_t + x_t M &= b_t - Rb_{t-1} + z_t[P_t - M] + \omega_t & for\ \forall t > 1. \end{aligned} \tag{2.1}$$

where

$$b_t \leq \sum_{s=t}^{T} \delta^{t-s+1} \big[\omega_s + z_s[P_t - M]\big]. \tag{2.2}$$

We assume that all foreign collectors have wealth in each period greater than $\sum_{t=0}^{\infty} max(\overline{F}, \overline{D})$ so that their credit constraints never prevent them from purchasing the object.

2.2.1 Timing

Timing of our model occurs as follows. Prior to period 0, the Minister of Culture decides on a policy which is written into law and binding for all future periods. In each section, we initially assume that the Minister of Culture has a choice only between allowing free trade or passing an export ban which restricts $z_t = 0$ for all t. We then explore the alternative of allowing for leases and sales with an option to buy the object back at a prearranged price.

Each subsequent period is comprised of four stages: a subsidy/tax stage, a sales stage, maintenance, and consumption:

Subsidy/Taxation: In the subsidy and taxation stage, the Minister of Culture can decide on a subsidy schedule which is paid to the owner when an object is held intact domestically and a tax schedule which is collected when an object is auctioned abroad. We assume that the domestic owner cannot be taxed if an object is stolen. This is similar to a limited liability constraint in a standard moral hazard problem.

Sales: After subsidy and tax schedules have been set, the owner of the antiquity chooses whether to sell the object or not. We make two key assumptions. First, we assume that whoever owns the object gets to decide on the selling procedure. Second, we assume that the laws passed by the Minister of Culture and the tax/subsidy schedules do not apply to objects that are owned abroad. Thus, if a foreign collector

[10] Poor owners of unregistered art may wish to hold on to objects as a way of keeping an informational advantage over the government and increasing their chance of being able to sell an object to a smuggler moving the object overseas.

owns the object, the foreign collector chooses the selling procedure optimal to him. This will typically be an optimal auction. When the domestic owner has future ownership rights, he will choose the selling procedure that maximizes his surplus subject to the legal and tax environment.

Maintenance: The party holding an object after the sales stage must immediately pay M to prevent the destruction of the object.

Consumption: All players consume for the period and play continues to the next period.

2.3 Aligning Incentives through Leases

As is typical with moral hazard problems, when taxation is frictionless and there is no asymmetric information, a tax system that can use lump sum transfers for redistribution purposes and taxes and subsidies to resolve the externality can reach the first best. In particular, the externality can be internalized by a subsidy to the owner of $D_{E,t}$ for every period the object remains intact and in the country. This will induce optimal maintenance and export decisions.

In reality, governments may be reluctant to subsidize wealthy owners of antiquities, estimation of the externality may be difficult, and bureaucrats in charge of taxes and subsidies may be corrupt. If potentially corrupt bureaucrats must judge the value of revealed objects, an asymmetric information model suggests that a subsidy program may be inefficient.[11]

In this section we examine the optimal policy when taxes and subsidies are completely inefficient. A dollar taxed from the individual is consumed by bureaucracy and leads to no social welfare. Likewise a dollar in subsidy requires infinite government funds. We begin with a simple two period model that illustrates how the owner of the object responds to export bans and then sketch how this model extends to the infinite horizon case.

[11]See the appendix for a simple model of this in the case in which high (H) and low (L) quality objects are random across a large number of potential domestic owners and the government wants to create an incentive program for people to reveal these antiquities. Under these conditions an incentive program for people to reveal local antiquities would expand to become far larger than the optimal program size, and would be partially comprised of bribes which the government may view as wasteful.

2.3.1 Two Period Case

Taking the most simple case of our model, let $t \in \{0,1\}$, $\delta = 1$, $R = 1$, and $N \to \infty$. Further assume that there is a perfect credit market so that

$$c_0 + c_1 + [x_1 + x_2]M = \omega_1 + \omega_2 + [z_1 + z_2][\overline{F} - M] \tag{2.3}$$

where $\overline{F} - M$ is the per period price generated by auctioning the object abroad.

We begin by studying the owner's optimal actions under a free market and an export ban when no subsidies and taxes are used.[12] Beginning with the case of a free market, since credit markets are perfect and the owner's utility for consumption is concave, the owner will borrow so that his consumption is constant over time. Using the fact that credit markets are perfect, the owner's problem simplifies to:

$$\max_{x_1,x_2} \quad 2u\left(\frac{\omega_2 + w_1}{2} + (\overline{F} - M) - \frac{x_2 + x_1}{2}\overline{F}\right) + D_o[x_1 + x_2]. \tag{2.4}$$

As can be seen in this equation, the proportion of time that the object stays at home, $\frac{x_1+x_2}{2}$, is increasing monotonically in the owner's total wealth. For low wealth levels, the marginal utility of consumption is high leading to a sale of the object in period 1. For intermediate wealth levels the owner shares usage, for instance by selling the object in period 2. For high wealth levels the owner keeps the object domestically in both periods.

The owner of an object facing an export ban faces a similar problem with the added constraint that foreign sales are illegal and thus $z_t = 0$. Again equating marginal utility across periods, an owner's problem becomes:

$$\max_{x_1,x_2} \quad 2u\left(\frac{\omega_2 + \omega_1}{2} - \frac{x_2 + x_1}{2}M\right) + D_o[x_1 + x_2]. \tag{2.5}$$

subject to $x_2 \leq x_1$.

Equations 2.4 and 2.5 differ in two respects. First, the inability to sell abroad reduces the total potential income of the owner by $\overline{F} - M$. Since consumption is concave, the decrease in income leads to an increase in the the marginal utility of consumption for individuals under an export ban which reduces maintenance incentives and potentially decreases home usage. On the other hand, an owner facing an

[12]It is interesting to note, that even though taxes are inefficient, they still may be used as punishment to dissuade owners from selling an object. The optimal tax schedule, however, is isomorphic to the sharing rules developed in the next section and thus is omitted from formal analysis.

export ban considers the tradeoff between keeping an object and letting it be destroyed rather than the tradeoff between keeping an object and selling it. For a given wealth level, this substitution effect will increase the amount of time an object stays intact in the country of origin.

Theorem 6 *For high wealth levels of the private owner, an export ban leads to an increase in home usage relative to a free market. For low wealth levels, an export ban leads to the destruction of objects.*

Proof. Define an allocation profile as $\{x_1, x_2\}$. Under a free market $z_t = 1 - x_t$ while under an export ban $z_t = 0$. Letting $W = \omega_1 + \omega_2$ and writing out the maximal utility for each possible allocation profile generates Table 2.1.

Table 2.1: Utility by Allocation Profile $\{x_1, x_2\}$

	$\{0,0\}$	$\{0,1\}$	$\{1,0\}$	$\{1,1\}$
Export Ban	$2u\left(\frac{W}{2}\right)$	$\varnothing$	$2u\left(\frac{W}{2} - \frac{M}{2}\right) + D_o$	$2u\left(\frac{W}{2} - M\right) + 2D_o$
Free Market	$2u\left(\frac{W}{2} + \overline{F} - M\right)$	same as $\{1,0\}$	$2u\left(\frac{W}{2} + \frac{\overline{F}-M}{2} - \frac{M}{2}\right) + D_o$	$2u\left(\frac{W}{2} - M\right) + 2D_o$

The marginal utility of u is decreasing in W and thus the choice of $x_1 + x_2$ is increasing monotonically in W. For each trade policy, there exists a unique wealth, $\underline{W}$, such that the owner is indifferent between selecting allocation $\{0,0\}$ and $\{1,0\}$ (or $\{0,1\}$ in the case of a free market) and a unique wealth $\overline{W}$ such that the owner is indifferent between selecting allocations $\{1,0\}$ and $\{1,1\}$.

Notice that the gross utility for choosing allocation $\{1,1\}$ is the same under both an export ban and a free market. Since $\frac{P-2M}{2} > -\frac{M}{2} \rightarrow \overline{W}^{Ban} < \overline{W}^{Free}$.

The ordering of $\underline{W}^{Ban}$ and $\underline{W}^{Free}$ is ambiguous. As with the upper threshold, the difference in consumption is larger in the free market than under an export ban. However, there is now a difference in income for allocation profile $\{1,0\}$ so that the concavity of u needs to be taken into account. In general, an export ban may lead to both a decrease in welfare and a decrease in the amount of time an object stays in the country of origin if u is very concave or $\overline{F} - M$ is large. ■

2.3.2 Infinite Horizon

The two period model can be readily extended to infinite periods. Assume that the period length gets small so that $\delta \rightarrow 1$ from below and assume $\delta R = 1$. Recall that under these conditions, it is possible to construct $\pi_D = [1-\delta]\sum_{t=0}^{\infty}\delta^t x_t$ which is continuous over $[0,1]$ using a combination of the discrete variables $\{x_0, x_1, \ldots, x_\infty\}$. Maintaining the assumption of perfect credit markets, we can simplify the infinite horizon problem under free markets to:

$$\max_{\pi_D}\ u\Big((1-\delta)W + [\overline{F} - M] - \pi_D\overline{F}\Big) + \pi_D D_O \tag{2.6}$$

where $\pi_D \in [0,1]$ is the discount adjusted proportion of time the object is used by the domestic owner and $W = \sum_{t=0}^{\infty}\delta^t\omega_t$. We can similarly rewrite the maximization problem under an export ban as:

$$\max_{\pi_D}\ u\Big((1-\delta)W - \pi_D M\Big) + \pi_D D_O. \tag{2.7}$$

Just as in the two period case, the creation of an export ban leads to an income and substitution effect. An owner facing an export ban considers the tradeoff between the utility gained from using the object and the consumption lost from maintaining the object. An owner under a free market considers the tradeoff between using the good and the value gained from selling it. As shown graphically on the right hand side of Figure 2-1 by the gap between $\overline{W}_{Ban}$ and $\overline{W}_{Free}$, this substitution effect leads to an increase in usage for high wealth levels.

On the other hand, an owner faced with an export ban cannot increase wealth through partial usage. For any wealth level where $\pi_D^{Ban} < 1$, the marginal utility for consumption is higher under an export ban than under a free market. As seen in the slopes of the two lines, this income effect leads art usage to decline faster in response to a decrease in wealth under an export ban than under a free market.

2.3.3 The Minister of Culture's Problem

Returning to the policy makers problem, we analyze the simplest case of the model where the domestic externality is constant with value D_E.[13]

[13]Our main results of this section are about the reachable set of contracts. Since $D_{E,t}$ does not affect this dimension, we avoid the added notation necessary to characterize solutions with this additional dimension.

Figure 2-1: Discounted proportion of time object remains intact and in the country of origin as a function of owner's wealth

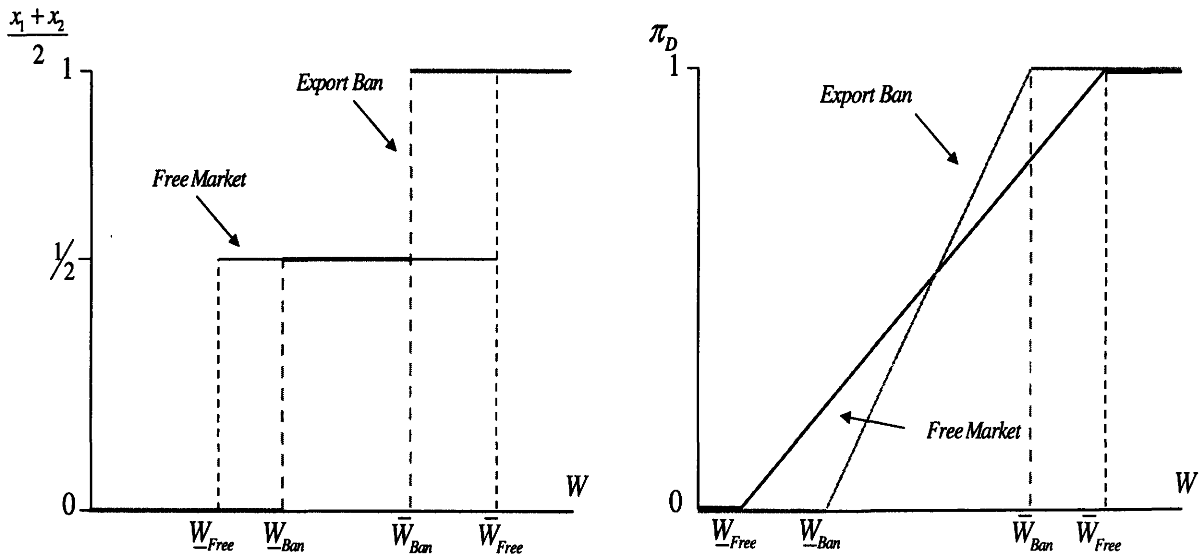


Figure 1a: Two Period Model　　*Figure* 1b: Infinite Horizon Model

A Rationale for Export Bans

An export ban may be welfare improving if it keeps the object in the country of origin for an extended period of time and the externality for domestic usage, D_E, is large. Export bans are a constraint imposed on the owner of an object. If the foreign collector places more value on the object, the constraint is binding and fundamentally reduces the possible utility of the owner. From a social perspective, an export ban will be preferable to free trade if the increased utility from usage is greater than the decrease in consumption utility. In the infinite horizon case, this will be true if:

$$(\pi_D^{Ban} - \pi_D^{Free})(D_o + E[D_E]) \geq u(c^{Ban}) - u(c^{Free}) \tag{2.8}$$

where

$$\begin{aligned} u(c^{Free}) &= u\Big((1-\delta)W + [\overline{F} - M] - \pi_D^{Free}\overline{F}\Big) \\ u(c^{Ban}) &= u\Big((1-\delta)W - \pi_D M\Big). \end{aligned} \tag{2.9}$$

Our results show why export bans may have significantly different effects in rich versus poor countries. In a country such as Italy where artifacts are typically in the hands of the affluent and where the average income is high, export bans may increase the amount of time an object stays in the country without increasing the theft and destruction rate of objects. In a poor country such as Mali, however, many objects are buried in areas of high poverty with the location of objects known only by the local population. At least in the model, this combination leads to a deterioration of protection when bans are put into place in poor locales.

Optimal Policy

As before, define π_D as the discounted proportion of time that an object is maintained domestically. The social planner's first best solution is to choose π_D that solves

$$\max_{\pi_D \in [0,1]} u\big((1-\delta)W + P - M - \pi_D P\big) + \pi_D[D_O + D_E]. \tag{2.10}$$

Taking the FOC of Equation 2.10 let π^* be the value of π that solves

$$u'\big((1-\delta)\hat{W} + P - M - \pi\big) = \frac{D_O + D_E}{P} \tag{2.11}$$

for wealth level $\hat{W}$. From the bounds on π_D, it follows that

$$\pi_D^{FB} = \begin{cases} 1 & \pi^* > 1 \\ \pi^* & \pi^* \in [0,1] \\ 0 & \pi^* < 0 \end{cases} \tag{2.12}$$

We take the viewpoint that the government cannot directly regulate π_D but does have the ability to limit, π_F, the proportion of time that the object can leave the country legally. We restrict our attention to regulation that can restrict the percentage of time an object can leave the nation but does not regulate the exact periods. Legislation that requires an object to stay in the country for a certain amount of time before export increases the reachable set of domestic usage in this simple model but suffers in more complicated environments with credit constraints, stochastic endowments, or asymmetric information.

Theorem 7 *When taxes and transfers are completely inefficient, the government can achieve the second best solution by setting a constraint on the amount of time an object can leave the country.*

Proof: Proof is in the appendix.

While we leave the formal proof of Theorem 7 to the appendix, we show the key ideas from the proof graphically. In the interesting case of our model, the first best level of π_D is not implementable since it does not leave enough consumption to the agent to satisfy his maintenance constraints. The government must therefore look for $\pi_D^{Reach}(\tilde{W})$, the maximum reachable π_D given the agents initial wealth $\tilde{W}$.

Figure 2-2: Construction of the optimal sharing rule

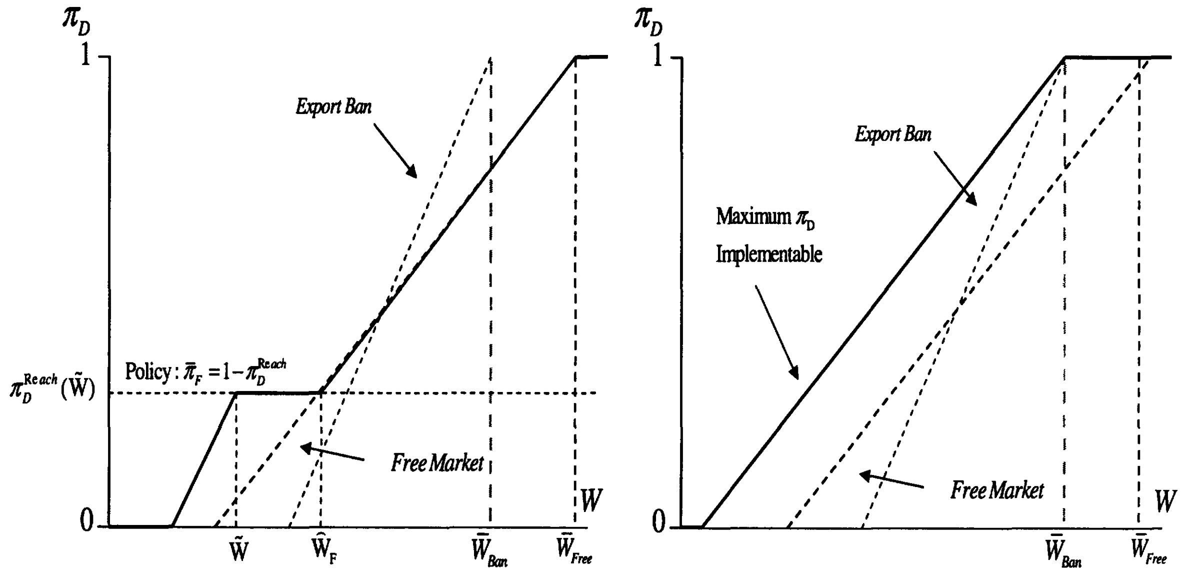


Figure 2a: Policy that maximizes π_D at $\tilde{W}$ *Figure* 3b: Maximum Reachable π_D

As shown in the left hand side of Figure 2-2, $\pi_D^{Reach}(\tilde{W})$ is found by restricting the share of foreign usage to a level $\bar{\pi}_F$ such that an owner with slightly less wealth would begin to allow the object to be destroyed. The logic for such a law is as follows. Under any sharing rule, domestic usage is increasing in wealth. Thus, since the destruction of the object leads to a decrease in wealth, the optimal policy never creates incentives for destruction.[14] At the same time, after an object has been moved abroad for a proportion of time equal to $\bar{\pi}_F$, the owner of the object faces a defacto export ban. For any wealth level where an agent who faces an export ban chooses to maintain the object, domestic usage under an export ban is weakly higher than under any other policy. Thus, by using the substitution effect of the export ban and the income effect of the free market, a sharing rule that limits the time an object can be shared maximizes π_D for a given wealth $\tilde{W}$.

Following this logic over all possible wealth levels, the maximal reachable π_D is parallel to the contracts attainable with a free market but is shifted to the left by

[14]By a similar logic, the optimal policy never charges a positive tax for foreign usage.

$\overline{W}^{Free} - \overline{W}^{Ban}$. This set is shown on the right hand side of Figure 2-2.

Theorem 8 *Let π^{**} be the solution to*

$$u'\big((1-\delta)\hat{W} + P - M - \pi\big) = \min\left\{\frac{D_O + D_E}{P}, \frac{D_O}{M}\right\}. \tag{2.13}$$

For a given owner wealth of $\hat{W}$, The optimal policy that is incentive compatible is to constrain the amount of time an object can leave the country to $\overline{\pi}_F = 1 - \pi_D^{SB}$ where

$$\pi_D^{SB} = \begin{cases} 1 & \pi^{**} > 1 \\ \pi^{**} & \pi^{**} \in [0,1] \\ 0 & \pi^{**} < 0 \end{cases}. \tag{2.14}$$

Wealth and the ratio of value to maintenance costs both play an important role in defining the optimal policy. In a relatively wealthy country, a straight export ban may approximate the optimal second best policy. In a developing country, D_O may be small and the wealth level of owners may be low. Under these circumstances, a straight export ban can lead to a much lower level of domestic usage than a policy that allows for the sharing of usage rights.

It is worth noting that the optimal sharing rule developed in this section can also be implemented using tax policy. Again using the assumption that taxes are completely inefficient, the optimal tax policy uses a non-linear tax schedule which has zero tax for an object that are shared abroad for more than time $\overline{\pi}_F$ and imposes the maximum possible tax subsequently. This tax policy is identical to a law that restricts foreign use to $\overline{\pi}_F$.

2.4 Transaction Costs and the Role of Leases

In this section we explore how asymmetric information and uncertainty about the future shape the optimal contract for sharing objects abroad. When a country is initially poor but may want objects back in the future, it must decide how best to move the object and secure its return. When the externality for domestic usage is unknown, selling the object today and attempting to repurchase it in the future may be inefficient, since attempts by foreign collectors to extract surplus from the government may prevent efficient transactions. Keeping ownership rights to an object in the home country via a lease or fixing prices in future negotiations with an option allows the government to exert influence on future negotiations and eliminate this hold up problem.

To focus on asymmetric information, we present a case polar to the prior section in which the government's tax technology is completely efficient. The government can use lump sum transfers so that the marginal utility of consumption is equal to the marginal value of money for the government. Further, it can use subsidies and taxes on legal transactions to align the incentives of the owner with that of the government.

Relative to the size of the country's total budget, the value of the object is assumed to be small. As such, we simplify the problem by assuming a linear marginal utility of consumption. Initially we also assume that the country is not credit constrained and normalize $u(c_t) = c_t$. In section 2.4.2, we relax this assumption in order to study optimal contracting in an environment where the shadow cost of money decreases over time.

2.4.1 Asymmetric Information

Let $D_t = D_O + D_{E,t}$ for $t = \{0, \ldots, \infty\}$. We assume that the country is poor initially so that $D_0 = \underline{D} < \underline{F}$. However, there is potential for it to become rich in the future. Let D_1 be drawn from distribution $H_1(.)$ and assume that all exogenous variation is resolved at this point so that $D_t = D_1, \forall t > 1$. We assume that at time D_0, the the value D_1 is unknown to all parties but the distribution $H_1(.)$ is common knowledge. At time 1, D_1 is realized, but is private information known only to the Minister of Culture.

In our model, ownership of an object gives the owner the power to choose the mechanism that is used in future exchange. Thus, in the first period, the government has the power to select the mechanism by which an object is sold. If the object is sold to the foreign buyer in the initial period, then that foreign seller selects the mechanism in future negotiations. We assume that both the government and foreign buyers have the power to commit to any mechanism they choose.

As before foreign collectors have a private value for art consumption F^i which is independently drawn from cdf A(.) which is time invariant. Keeping the mnemonics $D = Domestic$, $F = Foreign$, $H = Home$, $A = Abroad$ in mind, the problem can be though of as a two period allocation problem.

$$\begin{array}{lcc} & \underline{\text{Period 0}} & \underline{\text{Period 1+}} \\ \text{Domestic} & D_0 = \underline{D} & D_1 \sim H_1(.) \\ \text{Foreign} & F_0^i \sim A(.) & F_1^i = F_0^i \end{array} \quad (2.15)$$

In the first period, all of the foreign collectors have greater value for object than the home country and thus the object should be moved abroad in a first-best world. In the second period, the home country learns its new domestic valuation and may value the object more than foreigners.

To avoid corner solutions, we assume that H has sufficient variance such that $H_1(\underline{F}) < 1$ and $H_1(\overline{F}) > 0$. Asymmetric information surrounding the government's willingness to pay for an object is most likely more acute than foreign collectors' valuations. In most government decisions, an official is assigned to deal with repatriation and must estimate the net present value of future utility that citizens of its country would get from the object. Given the highly subjective nature of this estimation it may be difficult for a foreign buyer to accurately gauge the government's willingness to pay.

Sales and Repurchase Contracts

We first analyze the efficiency of selling an object in period 0 and attempting to repurchase the object back from the foreign buyer in period 1. In our model, the foreign buyers (i) have independent valuations, (ii) are risk neutral, (iii) are not credit constrained, and (iv) are symmetric with valuations drawn from the same distribution functions. Under these characteristics, if an object is sold using either the efficient or optimal auction, the winner of the object will be the foreign collector with the highest intrinsic valuation for the object F^N.

Given this observation, analysis of the sale and repurchase contract occurs in two steps. We first study the optimal mechanism for the foreign buyer who is attempting to sell the object back to the Minister of Culture in period 1. Using the expected revenue from this resale auction, we then return to period 0 to analyze efficiency.

Step 1 - Repurchase: For readability we define

$$\tilde{P} = P + M \quad (2.16)$$

and designate the hazard rate of distribution H and A as

$$\psi_H(\tilde{P}) = \frac{h(\tilde{P})}{1 - H(\tilde{P})}, \psi_A(\tilde{P}) = \frac{a(\tilde{P})}{1 - A(\tilde{P})}. \tag{2.17}$$

The foreign collector in our problem has commitment power and thus selects the mechanism that maximizes his expected utility. To avoid the decrease in monopoly power associated with intertemporal price discrimination, the foreign buyer commits to offering the object back to the Minister of Culture exactly once in period 1. The optimal price is found by solving the standard per period monopoly problem

$$P_t = \arg\max_{P_t}[P_t - (F^N - M)][1 - H(\tilde{P}_t)], \tag{2.18}$$

and then aggregating up over all periods:

$$P^{Resale} = \sum_{t=1}^{\infty} \delta^{t-1} P_t, \tag{2.19}$$

Taking the first order condition of Equation 2.18, the solution takes on the familiar monopoly solution:

$$P^{Resale} = \frac{\delta}{1 - \delta}\left[F^N - M + \frac{1}{\psi_H(\tilde{P}^M)}\right]. \tag{2.20}$$

Step 2 - Sale: Returning to the auction in stage 1, a foreign collector with value F^i incorporates the monopoly rents into his original value. Thus, the value of an object to a foreign collector with value F^i is:

$$V^i(F^i) = \frac{1}{1 - \delta}(F^i - M) + \frac{\delta}{1 - \delta}\frac{1 - H(\tilde{P}^M)}{\psi_H(\tilde{P}^M)}. \tag{2.21}$$

Consider the case where $\lim_{N \to \infty}$. When the optimal reserve price is less than $V(\overline{F})$, the optimal and efficient auction will both generate revenues equal to $V(\overline{F})$ and the foreign collectors expected value goes to zero. However, the monopoly power of the foreign collector in the second period creates residual inefficiencies in the initial auction. By attempting to extract rents from the domestic owner, the foreign collector offers an inefficiently high price in period 1. While these rents are recaptured by the domestic owner in period 0, the allocation in the future is inefficient which leads to a permanent loss of possible total utility.

Since the collector's utility is always zero and his utility is linear, maximizing domestic utility subject to the foreign collector's IR constraint yields the socially optimal price:

$$P^{FB} = F^N - M. \tag{2.22}$$

In a sale and repurchase scheme, $P^{Resale} \geq P^{FB}$ and thus there exists possible realizations of D_1 in which an object is misallocated to the foreign collector when the object has a greater value at home.

Rationale for an Export Ban

Whereas objects that leave the country in the future stay in foreign hands too long, inefficiency may also cause the home country to block sales in the initial period if the government can only sell objects. As a simple example, continue to assume that $N \to \infty$ so that $F^N = \overline{F}$. Comparing the revenue that the home country expects to receive in period 1 when it has control of the auction mechanism versus when it does not, the home country expected loss from a sale is:

$$S^{FB} - S^{Sale} = \frac{\delta}{1-\delta} \underbrace{[H_1(\tilde{P}^{Resale}) - H_1(\overline{F})]}_{\substack{\textit{Probability of} \\ \textit{Inefficient Trade}}} \underbrace{[E(D_2 | \overline{F} \leq D_2 \leq \tilde{P}^{Resale}) - \overline{F}]}_{\textit{Expected Loss}}. \tag{2.23}$$

Intuitively, the magnitude of this inefficiency is equal to the probability of an inefficient trade multiplied by the expected loss when such an event occurs. When the discount rate is small, the home country may have incentive to hold the object until uncertainty is resolved.

Leases

At its core, the problem with selling an object and trying to repurchase it in the future is a contractual one. Both the foreign collector and the domestic owner have an incentive to distort prices and consumption in order to increase rents in the second stage. However, since these rents are already priced into the initial auction, strategic action leads to pure efficiency losses without any change in the overall share of profits.

Leases diminish the effects of asymmetric information by leaving the choice of mechanism in both periods to the government, which can use auctions to significantly reduce the asymmetric information about the valuations of the foreign collectors. Consider a lease auction in which the government leases the object to the foreign agent in the first period but retains future ownership rights. As is well known from the auction literature, an agent running multiple auctions cannot improve his final outcome by using information about the winning bid from the first auction in later auctions. We thus assume the agent constrains the information generated in the auction by running an optimal English auction in stage 1 in which the second highest bid price but not the true value of the winner is revealed.

Given the information revelation of the initial auction, the domestic owner knows the value of the second highest agent F^{N-1} and the density function of the highest bidder:

$$a^N(F) = \begin{cases} \frac{a(F)}{1-A(F^{N-1})} & F > F^{N-1} \\ 0 & otherwise \end{cases} \tag{2.24}$$

The home country attempting to maximize profit in the second period solves:

$$\max_{p} \quad [1 - A^N(\tilde{P})]P + A^N(\tilde{P})[D_2 - M]. \tag{2.25}$$

Noting that since:

$$A^N(\tilde{P}) = \frac{A(\tilde{P}) - A(F^{N-1})}{1 - A(F^{N-1})}, \tag{2.26}$$

$A(F^{N-1})$ drops out of the FOC leaving an optimal period 2 offer price of:

$$P = max\left(D_2 + \frac{1}{\psi_A(\tilde{P})}, F^{N-1}\right) - M. \tag{2.27}$$

What is most interesting with this result is that the optimal pricing rule from equation (2.27) is equivalent to running an English auction in each period with a reservation price dependent only on the home countries value and the initial distribution. As the number of bidders goes toward infinity, such a lease auction converges to the socially efficient price $P = max(D_2, F^N) - M$.

Theorem 9 *When the government's utility function is linear and the buyers' valuation are independent and identically distributed, the optimal allocation mechanism is to lease objects each period using an anonymous English auction with reservation price:*

$$P_t^{RES} = D_t - M + \frac{1}{\psi_A(\tilde{P})} \tag{2.28}$$

Proof: Proof is in the appendix.

As indicated in the theorem, the optimal lease contract generalizes readily to a more complicated environment where D_t grows over time.[15] Since contracting in each forward period is done once uncertainty about the home countries value is resolved, the reserve price in each period extracts the maximum expected value. Further, since the true valuation of the highest valued bidder stays private, there is no incentives for individuals to strategically manipulate their bids.

2.4.2 Credit Constraints

In the previous section, access to a perfect credit market allowed efficiency of the contract to be separated from intertemporal consumption. In each period the government had the ability to borrow against future leases and thus smooth consumption over time.

In many environments where objects are threatened, the government is likely to be credit constrained. This section looks at the case where the total value of an object is small relative to the nation's budget, but where the marginal utility of money differs over time. If the government expects to get richer in the future, the government would prefer to extract and spend as much surplus as possible from the object in period 0 as long as the inefficiencies caused in the art market are not too large. A sale auction with an option to buy the object back at a prearranged price resolves the holdup problem of the sales contract while maximizing expected revenue in the first period.

Option Contracts

Suppose that the government is credit constrained so that the marginal utility of money is decreasing over time. We assume that the value of the object relative to the total budget is small and represent the utility of consumption in each period as $u(c_t) = (1+\gamma_t)c_t$. The country is poor today and is credit constrained so that $\gamma_0 \geq \gamma_t$ for all t. As such, the Minister of Culture prefers to (i) consume the value of the object in period 0 and (ii) prevent holdup in future periods.

An alternative to a lease in this circumstance is to sell an object to the foreign buyer with an option to repurchase the object in the future at a fixed price r. Determining the optimal reserve price in future periods is complicated by the fact that the actual value of the object is unknown ex ante. The home country has two instruments — the reserve price P_t^{RES} and the option price r at which it could repurchase the

[15] That is $D_t = D_{t-1} + Z_t(.)$ where $Z_t(.)$ is bounded below at zero.

object in period 2. In order for foreign collectors to be willing to enter the auction, the option price of an object must be greater than or equal to the sale price adjusted for interest. Thus $r_t \geq R^t P_t$. Under the assumption of independent and identically distributed values, the following theorem gives the optimal option contract:

Theorem 10 *When the government's utility function is linear and the buyers' valuations are independent and identically distributed, the optimal mechanism using a sale and option to repurchase is an English auction that sells ownership rights for each period independently with a reservation price of:*

$$P_t^{RES} = \delta^t \left[\frac{1}{1+\gamma_t+\kappa(\gamma_0)} \left[E(D_t | D_t < \tilde{P}_t^{RES} R^t (1+\gamma_t)) - M \right] + \frac{1}{\psi_A(\tilde{P}_t^{RES} R^t)} \right]. \quad (2.29)$$

where

$$\kappa(\gamma_0) = \frac{\gamma_0 - \gamma_t}{H(\tilde{P}_t^{RES} R^t)} \quad (2.30)$$

The optimal option price r is set equal to the sale price in each period, multiplied by the rate of return R^t.

Proof: Proof is in the appendix.

Like the lease contract, the optimal option contract eliminates the holdup problems of the sale contract by settling all negotiations ex-ante. The $\kappa(\gamma_0)$ term in this equation is a term that adjusts the reservation price downward when (i) the marginal value of money is higher today and (ii) it is likely that the country will repurchase the object in the future. In both these cases, the object acts like collateral in a loan with the reservation price being used to maximize loan size.

When there are no credit constraints and the marginal utility is normalized to one, the optimal option contract auctions the object using a reservation price equal to:

$$P_t^{RES} = \delta^t \left[E(D_t | D_t < \tilde{P}_t^{RES} R^t) - M + \frac{1}{\psi_A(\tilde{P}_t^{RES} R^t)} \right]. \quad (2.31)$$

This equation is analogous to the reservation price of a lease, shown in Equation 2.28, with the difference being that the auction is done in period 1 prior to the true valuation of D_2 being known. In an environment in which there is no corruption and the government knows its future valuation, the efficiency of the lease and option contracts are the same. However, in the case where future valuations are unknown and credit constraints do not exist, leases dominate both option contracts and sale and repurchase contracts.

Theorem 11 *When future valuations of the home country are unknown to all parties and there are no credit constraints, a lease auction maximizes social welfare of the home country.*

Proof: Proof is in the appendix.

When credit constraints exist and the country is expected to get wealthier over time, options may dominate leases since the country has the ability to consume more than its per period valuation. In our model, this will occur if $\gamma_0 > \gamma_t$ for all t and the future valuations of $D_{E,t}$ are known but private.

Theorem 12 *When future valuations of the home country are known but private information and credit constraints bind, an auction that sells the object with an option to repurchase maximizes social welfare of the home country.*

Proof: Proof is in the appendix.

Option auctions differ from lease auctions in that they can generate a much larger surplus in the initial period for consumption. In an option contract, the object is sold in the first period, allowing a country the ability to consume more than F^{N-1} in period 1. In cases where the government is credit constrained, the flexibility in consumption that option auctions allow may be advantageous.

2.5 Corruption and Intergenerational Conflict

In this section we argue that if there is a probability of a corrupt Minister of Culture in each period who seeks to appropriate the value of the object from future generations, laws restricting international art transactions may be optimal. When the only laws that can be passed is an export ban, a benevolent Minister of Culture may choose to enact such a law in order to constrain bad agents at the cost of restricting good agents from acting optimally. For reasonable parameter values, less draconian export restrictions that allow one period leases are superior to both free trade and complete export bans.

Recall from the timing of the model that the Minister of Culture can pass a law prior to period 0 that binds itself and all future regimes. In this section, we assume that decisions in subsequent periods are made by a sequence of new Ministers of Culture who act as a social planner with probability $1 - \epsilon$ but which maximizes their own consumption with no regard for current or future generations with probability ϵ. Bad ministers will allow objects to be exported and consume the proceeds.

As in section 2.4 we assume that the value of $D_O + D_{E,t}$ is stochastic with independent and identically distributed shocks and CDF $H_t(.)$. We continue to assume that tax and subsidy are perfect under a benevolent Minister of Culture and that the country faces no credit constraints so that the marginal utility of consumption is normalized to 1.

A good leader who has no constraints on his action allows an object to be used by the foreign collector any time $D_o + D_E < P$ and keeps the object local otherwise.[16] Under an export ban, the object always stays in the country resulting in a value of $D_o + E[D_E]$.

Theorem 13 *If the only law available to a benevolent Minister of Culture is an export ban or free trade, as $\delta \to 1$, there exists a probability that a bad minister will come into power, $\epsilon^* \in (0,1)$, such that if $\epsilon < \epsilon^*$ the government maintains a free market and if $\epsilon \in (\epsilon^*, 1 - \epsilon^*)$ the government passes an export ban.*

Proof: Proof is in the appendix.

In this model, export bans act as a very blunt tool to constrain bad future ministers from acting in a malevolent way. By attempting to reduce the ability of future corrupt leaders to steal funds, the government limits the ability of good actors to make welfare improving trades.

Leases act as a way of balancing concerns of corruption with efficiency. Such leases may achieve a good balance of restricting the long-term damage that corrupt officials can do while still giving benevolent ones the ability to make Pareto-improving short-term trades.

To see this, note that with free trade, the country gets $\max(P, D_o + D_E) - M$ each period before the first bad leader arrives. Afterwards it gets nothing. The net present value of this stream is:

$$\mathbb{E}_{D_E} \sum_{t=0}^{\infty} \delta^t (1-\epsilon)^t \big(\max(P, D_o + D_{E,t}) - M\big). \tag{2.32}$$

[16]A benevolent dictator may also sell an object and distribute the earnings during his rule to prevent future bad dictators from expropriating this value. We assume that $P < D_o + E(D_E)$ so that such preemptive distribution is always dominated by an export ban.

Under an export ban that successfully binds bad bureaucrats, the country receives a maximum NPV of:[17]

$$\sum_{t=0}^{\infty} \delta^t (1-\epsilon)^{min(t,1)} \big(D_o + \mathbb{E}[D_{E,t}] - M\big). \tag{2.33}$$

Under a constitution or international treaty that permits one period leases but not sales, the NPV of the stream is:

$$\mathbb{E}_{D_E} \sum_{t=0}^{\infty} \delta^t (1-\epsilon)^{min(t,1)} \big(\max(P, D_o + D_{E,t}) - M\big). \tag{2.34}$$

This implies that for any positive ϵ, allowing leases but not sales dominates free trade. If P exceeds $D_o + D_{E,t}$ in some state of the world, then allowing leases but not sales is preferable to a complete export ban.

As noted in section 2.4.2, one caveat is that allowing leases but not sales dominates free trade only if there are no credit constraints. In a model with credit constraints, it may be desirable to transfer long-run claims on the object in exchange for higher consumption in the short run. A benevolent Minister of Culture, who believes that they would like the object in the future, will always want to sell the object with the option to repurchase in order to avoid holdup problems. A bad Minister of Culture would like to eliminate these repurchase clauses since the sale price in the future is higher without the option. As with leases, putting some restriction on the type of contract being signed will ensure efficiency and reduce the damage that corrupt officials can have on future generations.

2.6 Conclusion

Debates between cultural nationalists and internationalists have focused on the desirability of export bans. We argue that it may be appropriate to consider a broader class of contracts, including leases and perhaps sales with options to repurchase. Under three of the potential rationales for export bans we consider — externalities from keeping the object intact and within the country, the possibility that corrupt rulers or bureaucrats will expropriate the value of the national patrimony, and the difficulty

[17]We implicitly assume that bad leaders will still consume the usage value of an object in the period that they are in power, but that the export ban is otherwise successful. We view this as an upper bound of social value with export bans since in reality ban laws are typically porous.

of repurchasing objects once sold — leases or sales with options to repurchase may perform as well or better than export bans while generating more revenue for the country and improving maintenance incentives.

The simple models we examine here may abstract from important issues. First, objections to the sale of important cultural items may relate to unwillingness to alienate objects from the nation or distaste for "commodification" of antiquities. In this case, sales with repurchase options may be unacceptable, but leases may still be acceptable. If leases for general revenue are not acceptable, leases which dedicate revenue to the preservation of antiquities may be more acceptable.

Second, we do not address the case of disputed ownership (including when objects are overseas). It seems likely that in such cases, such as between Greece and Britain concerning the Elgin marbles, leases may be a way to effectively split ownership and thus avoid legal costs without declaring the value of the object.

Third, we focus on the case when the government knows the object exists and where it is. The appendix discusses a case in which the government needs to create incentives for citizens to reveal the existence of objects and argues that offering them lease rights for a set number of years could create such an incentive and may be robust to corruption problems in valuing objects that might make a cash reward system untenable.

It is worth noting that lotteries, as well as lease arrangements, could allow the value of objects to be split without declaring their value and could create incentives to reveal objects without a need for the state to estimate objects' value, but under lotteries the parties bear more risk and, as argued above, lease arrangements that give ownership rights to the state may achieve preferable intertemporal allocation.

As with any policy that increases the domestic value of antiquities, there is potential that the introduction of lease markets can exacerbate looting by giving individuals incentive to grab objects today. The ability to grant leases of varying length depending on the method of extraction may partially mitigate this incentive problem by creating differentiated products between legally and illegally excavated antiquities. The rules of such a policy are likely to vary on a country by country basis depending on the relative weight placed on recovering objects that have already been looted and preserving pristine sites.

Fourth, we have assumed that the valuation of foreign buyers does not change with respect to their ownership rights. It may be the case that for some private collectors, ownership itself generates value which is lost under lease and option contracts. In such circumstances, the sale of fractional ownership rights may be preferable to other contracts where a portion of ownership is sold to the foreign collector. Fractional ownership allows for both sharing and foreign ownership, but requires more sophisticated contracting which may be difficult in an international setting. Nonetheless, given the large amount of looted objects moving to private hands, such contracts may be worth exploring.

We have noted that leases are likely to be preferable in the presence of corruption or asymmetric information while option contracts are likely to be optimal in an environment with credit constraints. Limitations on the types of contracts that may be signed are a general constraint imposed on the action of the owner. The choice between allowing options and leases ultimately depends on the constraints facing the country and the relative weights given to the value of domestic ownership.

We have not modeled the optimal length of leases. If transaction costs are substantial, relatively longer leases may be desirable. We also do not address moral hazard in maintenance and return of the object by the receiving country. Based on existing experience with loans between museums, our sense is that these issues could be adequately addressed contractually, as long as the legal system in the receiving country is sufficiently well-functioning.

Finally, although this analysis has focused on markets for antiquities, it is worth noting that parts of the analysis may have implications for other contracting situations. In particular, the argument in Section 2.4 may help explain other patterns of asset ownership.

2.7 Appendix

2.7.1 Appendix 1: Using Leases as an Information Rent

We consider an economy with high (H) and low (L) quality objects which are distributed randomly across a large number of potential domestic owners. Each owner in the economy has no intrinsic value for her object ($D_o^H = D_o^L = 0$) but is in contact with a smuggler who will pay $\frac{\delta}{1-\delta}V^H$ or $\frac{\delta}{1-\delta}V^L$ in exchange for a high or low quality object. For simplicity, we assume that $u(c_t) = c_t$ so that each owner maximizes the expected value of her action.

The government accurately estimates the proportion (p) of high quality objects in the economy and the externality to its constituents for domestic use ($\frac{\delta}{1-\delta}D_E^H$ and $\frac{\delta}{1-\delta}D_E^L$). It has de jure rights to all domestic objects in the economy but lacks information in two dimensions. First, the government does not know the location of objects and must provide an information rent to domestic owners in order to convince them to reveal their objects. Second the government cannot distinguish between H and L quality units without the use of a bureaucrat who estimates its value and generates a report. Some bureaucrats are corrupt and may adversely alter a report to the government by reporting a low quality object as high quality. Assume that a proportion b of low quality objects are passed through the hands of corrupt officials who will accept a bribe B as compensation for deception.

The government would like to design an incentive mechanism that creates incentives for individuals to report their goods but is hampered by the inefficiencies that are generated by using bureaucrats as a part of their mechanism. Reliance on bureaucrats to carry out interim action may make the cost of a program prohibitively expensive or lead to allocation inefficiencies that may swamp the actual value of the program.

The timing of the game is as follows:

- Stage 0: The government decides upon an incentive mechanism and purchase rule.
- Stage 1: Individual owners decide whether to publicly disclose their object or sell them to the smuggler.
- Stage 2: Publicly disclosed objects are randomly assigned to bureaucrats:

- Stage 3: If an individual owner is assigned to a corrupt bureaucrat, the owner chooses whether to offer a bribe

- Stage 4: Bureaucrats generate their reports and the governments incentive mechanism is implemented.

Suppose that $D_E^H \geq V^H$ and $D_E^L \geq V^L$. In this case, the government would like to retain all objects for home use and thus must create an incentive structure that induces all agents to reveal their objects. Suppose first that the only mechanism available to the government is to provide cash transfers that are contingent on bureaucratic reports. Let T^H and T^L be transfers made to owners whose objects are reported as high and low respectively. Setting the IR constraint for individuals holding both high and low quality goods to the value of their outside option, we have

$$\begin{aligned} T^H &= \frac{\delta}{1-\delta}V^H, \\ (1-b)T^L + b(T^H - B) &= \frac{\delta}{1-\delta}V^L. \end{aligned}$$

Rearranging in terms of T^H and T^L yields:

$$\begin{aligned} T^H &= \frac{\delta}{1-\delta}V^H, \\ T^L &= \frac{\delta}{1-\delta}V^L - \frac{b}{1-b}\left(\frac{\delta}{1-\delta}\Delta V - B\right). \end{aligned}$$

The total cost of a project that purchases both high and low quality goods is

$$[p + (1-p)b]T^H + (1-p)(1-b)T^L.$$

Plugging in for T^H and T^L yields:

$$p\frac{\delta}{1-\delta}V^H + (1-p)\frac{\delta}{1-\delta}V^L - (1-p)bB.$$

In the case of $D_E^H \geq V^H$ and $D_E^L \geq V^L$ the possibility of bribes reduces the cost to providing incentives for low quality objects which reduces the cost of the overall program. Bribery generates additional transfers to bureaucrats which must ultimately be paid by the government. Such bribes may make the total cost of the program prohibitive, especially in cases in which the government views bribes as a form of pure waste in the economy.

The effect of bribery is exacerbated in the more likely case where $D_E^H \geq V^H$ and $D_E^L < V^L$. Here, the government would like to retain objects with a large externality and allow other objects to be moved out of the country. Corruption generates inefficiency both by making the total size of the program larger than it should be and by generating inefficiency through the misallocation of some low quality objects to the domestic market. The net allocation gain for the program is $\frac{\delta}{1-\delta}$ times

$$p[D_E^H - V^H] - (1-p)b[V^L - D_E^L]$$

If p is small, the overall efficiency gain for the program may be negative. The gross cost for the program is $\frac{\delta}{1-\delta}$ times

$$[p + (1-p)b]V^H$$

with $(1-p)bB$ of these proceeds going to pay for bribes to intermediaries. This program size is potentially far larger than the optimal program and is partially comprised of bribes which the government may view as wasteful.

2.7.2 Appendix 2: Loss Aversion

The social psychology literature suggests that many collectors' valuations of an object increase after taking possession of it. Such attachment creates inefficiencies any time the owner and private collector renegotiate display rights with changed ownership.

We model the attachment effect as stemming from loss aversion, though political economy constraints on foreign institutions limiting their ability to resell objects will generate near identical results. Assume that there exists a foreign collector who is loss averse with a reference dependent utility function that values art consumption x_t and non art consumption c_t in each period. Following Koszegi & Rabin (2006) we assume that the foreign collector's utility function is composed of two separable parts: the pure consumption utility of a bundle and a reference dependent gain-loss value of a bundle. A buyer's utility is given by:

$$U(c, x | r_c, r_x) = \sum_{t=1}^{2} \delta^{t-1} \nu(c_t, x_t | r_{c_t}, r_{x_t}) \tag{2.35}$$

where

$$\nu(c_t, x_t | r_{c_t}, r_{x_t}) = c_t + x_t F + \mu(c_t - r_{c_t}) + \mu(Fx_t - Fr_{x_t}). \tag{2.36}$$

Define $\mu(x)$ as a "universal gain-loss function" with the following properties:[18]

(1): $\mu(x)$ is continuous for all x, twice differentiable for $x \neq 0$ and $\mu(0) = 0$

(2): $\mu(x)$ is strictly increasing

(3): $\mu''(x) = 0$ for all $x \neq 0$

(4): if $y > x > 0$ then $\mu(y) + mu(-y) < \mu(x) + \mu(-x)$

(5): $\mu'_-(0)/\mu'_+(0) \equiv \lambda > 1$

$\mu(x)$ is linear for positive and negative values of x but has a steeper slope for losses than gains. Let $\mu(x) = \eta x$ for $x > 0$. By (3) and (5), this implies that when $x < 0$, $\mu(x) = \lambda\eta x$ with $\lambda > 1$. Agents who internalize a loss in one consumption dimension and a gain in the other will be reluctant to move away from the reference point even if consumption utility remains constant.

Consider first the case in which an agent's reference point is based on his initial state. When an agent does not own the object, as in Figure 1a, his reference point is $r_t = (c_{Init}, 0)$. A reduction in non art consumption to c in exchange for an increase in art consumption to x yields a change of utility of:

$$\nu(c, x|r_1) - \nu(r_1|r_1) = (1 + \lambda\eta)[c - c_{Init}] + (1 + \eta)Fx. \tag{2.37}$$

A loss averse agent's indifference curve passing through the reference point r_1 is all (c, x) pairs such that:

$$\frac{1 + \eta}{1 + \eta\lambda} Fx + c = c_{Init}. \tag{2.38}$$

If the agent is loss neutral ($\lambda = 1$), a foreign collector's utility function reduces to the simple utility function. As loss aversion increases, the agent requires a larger change in ownership to compensate for the losses felt in consumption. This drives down the price a foreign collector is willing to pay to buy an object.

Conversely, an agent who owns the object and has a reference point equal to the status quo $r_2 = (c_{Init} - P^{Sale}, 1)$ will be averse to moving away from owning the object.

[18]The universal gain-loss functions are identical to Bowman, Minehart & Rabin (1999) where the gain-loss function is assumed to be linear in sensitivity. See also Kahneman & Tversky (1991).

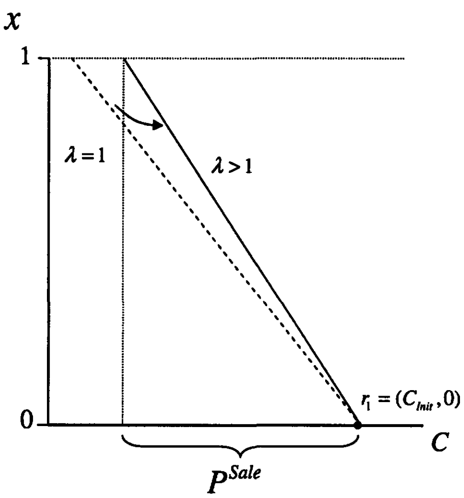


Figure 1a: Indifference curve with reference at r_1 for $\lambda = 1$ and $\lambda > 1$

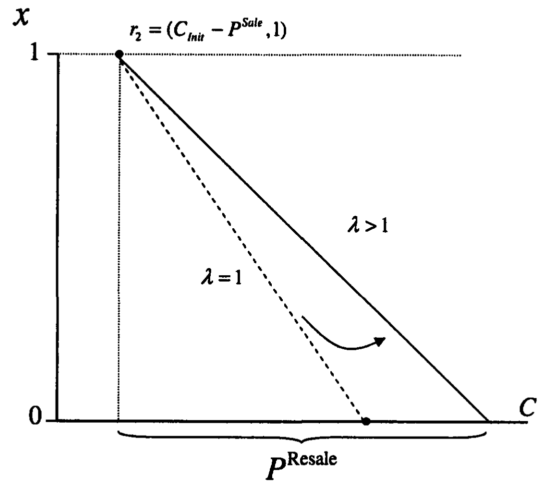


Figure 1b: Indifference curve with reference at r_2 for $\lambda = 1$ and $\lambda > 1$

As in Figure 1b, the indifference curve from r_2 is all points such that:

$$c - \frac{1+\eta\lambda}{1+\eta} F x = C_{init} + \frac{1+\eta\lambda}{1+\eta} F. \tag{2.39}$$

As λ increases, the level of c necessary to compensate for the lost art consumption x increases. This drives up the price that the domestic owner must pay to repurchase the object from the foreign collector.

In the case of sale and repurchase contracts, negotiations in the sale and resale phase are likely to occur at two reference points. This may drive a wedge in the price paid for an object and the price at which an object may be repurchased. Leases and options mitigate the effect of loss aversion by fixing the repurchase agreement in the original reference state.

Loss Aversion and Asymmetric Information

In isolation, both loss aversion and asymmetric information will lead to inefficiency in sale and repurchase contracts. In this section, we briefly comment on how these two forces compound.

Much of the current literature on reference dependent preferences centers on how individuals gain and update their reference point. The effects of loss aversion on auctions, for instance, depend strongly on assumptions made on self knowledge (whether individuals understand that they are loss averse) and at what speed agents update their reference points.[19] Side stepping some of the less empirically studied issues that arise with loss aversion in an auction setting, we concentrate on the resale portion of the sale and resale contract to illustrate how loss aversion and asymmetric information might compound.

Continuing to use the utility function presented in Koszegi & Rabin (2006), we assume that an agents reference point is his beliefs about outcomes. To accommodate for probabilistic beliefs, we allow for the reference point to be a probability measure G over consumption ($\mathbb{R}$) and usage ($\{0,1\}$). When the actual consumption bundle $C = (c_2, x_2)$ is drawn according to a probability measure F, utility is given by

$$U(F|G) = \int\int u(c|r)dG(r)dF(c). \tag{2.40}$$

For simplicity, we assume that preferences are linear in probabilities.

Given that the foreign buyer decides his offer price long before the domestic agent accepts or rejects it, it is likely that the foreign buyer will become acclimated to the risk associated with his offer. That is, the choice of an offer price today will determine his reference point by the time that the domestic agent chooses to accept or reject the offer. As a solution concept, therefore, we use the choice-acclimating personal equilibrium proposed by Koszegi & Rabin (2007):

Definition 3 *For any choice set D, $F \in D$ is a choice-acclimating personal equilibrium (CPE) if $U(F|F) \geq U(F'|F')$ for all $F' \in D$.*

A foreign buyer offering an object back to the country of origin takes into account the gain and losses that are associated with his offer price. Adding in the gain/loss terms to the standard monopoly maximization problem yields the following maximization:

$$\max_P \underbrace{[1 - H(\tilde{P})]P + H(\tilde{P})(F^N - M)}_{\textit{Consumption Utility}} - \underbrace{\eta(\lambda-1)H(\tilde{P})[1 - H(\tilde{P})][P + (F^N - M)]}_{\textit{Gain-Loss Utility}} \tag{2.41}$$

The first two terms of this maximand are identical to a simple monopoly pricing model with reservation price $F^N - M$. The last term is the loss of utility due to loss aversion. As the probability of changing hands moves toward $\frac{1}{2}$, loss aversion has a

[19] See Dodonova & Khoroshilov (2005) for a brief discussion of these issues in relation to revenue equivalence.

stronger effect since there the agent feels a strong loss both when the object is sold and held.

Theorem 14 *A foreign buyer with loss aversion parameter $\lambda > 1$ sets:*

$$\begin{aligned} P^{LA} &= (F^N - M) + \frac{1}{\psi_H(\tilde{P}^{LA})} \\ &\quad + \eta(\lambda - 1)\left[[2H(\tilde{P}^{LA}) - 1][P^{LA} + (F^N - M)] - \frac{H(\tilde{P}^{LA})}{\psi_H(\tilde{P}^{LA})} \right] \end{aligned} \tag{2.42}$$

This price is increasing in λ if the last term is positive:

The first two terms on the right hand side of Equation 2.42 are the standard optimal pricing optimum in the case of a monopolist with reservation price $F^N - M$ who faces asymmetric information and can set only one price for all consumers. The last term can be thought of as the increase in price generated by loss aversion. As long as the foreign buyer reasonably expects to keep an object, attachment for the object will drive the repurchase offer upward.

As a simple example assume that the domestic valuation for the home country, $D_2 \sim U[0,1]$. A foreign buyer will increase prices due to loss aversion any time $F \geq .2613$. For prices to be decreasing in λ, the future expected value of the object to the home country would need to be approximately double that of the foreign buyer.

As can be seen with this simple example, the trade mechanism in a sale and repurchase contract is inefficient for three main reasons:

1. The reference state prior to the first auction incorporated the uncertainty of losing the auction to all future states, depressing the prices paid by the agents.

2. The foreign collector had bargaining power in the second stage which created inefficiency as he tried to extract surplus from the domestic owner.

3. The foreign collector was loss averse in the resale phase leading to a higher offer price and an exacerbation of asymmetric information.

Leases are likely to improve upon sales and repurchase plans for all three reasons. In a lease contract, bargaining in the second stage onward occurs before the discovery that the agent is the highest bidder is realized. This depressed the reference point of the agent leading to a reduction in bids from all agents. By contrast, in a lease auction the identity of the highest bidder is known prior to the second stage negotiation. This increases his willingness to pay.

Unlike the resale case in which the loss averse agent is making take it or leave it offers, loss aversion may mitigate the effects of asymmetric information when it is on the responder side of the market. A government attempting to extract money from the foreign buyer by setting a reservation price in subsequent auctions must take into account how this affects the expectations that the foreign collector has about keeping the object. Provided that the foreign buyer correctly predicts the distribution of prices and her behavioral reaction to these prices, the foreign buyer's willingness to pay is decreasing as the offer distribution shifts upward.

Finally, if the home country is also loss averse and has expectations of keeping an object in the future, leases remove future considerations from current ones. By retaining an option to repatriate the object in the future, the reservation price imposed on objects can be reduced increasing short-run allocation efficiency.

2.7.3 Proofs

THEOREM 1: Proof shown in main text

THEOREM 2: We first prove the following lemma:

Lemma 3 *For any* $\delta \in \left[\frac{1}{2}, 1\right]$, $\exists\{x_0, x_1, \ldots, x_\infty\}$ *such that* $\sum_{t=1}^{\infty} \delta^t x_t = a$ *for any* $a \in \left[0, \frac{1}{1-\delta}\right]$

Proof. Let a be an arbitrary value in $\left[0, \frac{1}{1-\delta}\right]$ and consider the following algorithm. Define:

$$S_0 = \begin{cases} 1 & a > 1, \\ 0 & otherwise \end{cases} \qquad\qquad (2.43)$$

$$S_N = \begin{cases} S_{N-1} + \delta^N & a - S_N > \delta^N, \\ S_{N-1} & otherwise \end{cases}$$

When $0 \leq a - s_N \leq \frac{\delta}{1-\delta}\delta^N$ this implies that $0 \leq a - s_{N+1} \leq \frac{\delta}{1-\delta}\delta^{N+1}$ since $\delta^N \leq \frac{\delta}{1-\delta}\delta^N$ when $\delta \in (\frac{1}{2}, 1]$. Thus, by induction, $S_N \to a$ since $\delta^N \leq \frac{\delta}{1-\delta}\delta^N$ converges to zero. ∎

Since a is continuous in $\left[0, \frac{1}{1-\delta}\right]$, $\pi_D = (1-\delta)a$ is continuous from $[0,1]$. Note also, that as $\delta \longrightarrow 1$, can get arbitrarily close to $\pi_D \in [0,1]$ by selecting a T^* such that

$$x_t = \begin{cases} 1 & t < T^* \\ 0 & otherwise \end{cases}. \tag{2.44}$$

In cases where there is the potential that an object can be destroyed, we use this convention in order to avoid violating the maintenance constraints.

Analogous to π_D, define $\pi_F = [1-\delta]\sum_{t=1}^{\infty}\delta^t z_t$. When there is no destruction of the object, it follows that $\pi_F = 1 - \pi_D$. Let $T(\pi_F)$ be a (possibly nonlinear) tax imposed on the owner of an object based on the amount of time he legally leases an object out of the country. Since taxes are inefficient, the penalty is completely wasteful and any money taken from the owner is destroyed. The owner solves:

$$\max_{\pi_D,\pi_F} u\Big((1-\delta)\hat{W} + \pi_F[\overline{F} - M] - \pi_D M - T(\pi_F)\Big) + \pi_D D_O \tag{2.45}$$

subject to $\pi_F + \pi_D \leq 1$

Let π_D^*, π_F^* be the domestic and foreign usage when π_D^* is maximized and define:

$$w(\pi_F^*, T(\pi_F^*)) = (1-\delta)\hat{W} + \pi_F[\overline{F}^* - M] - \pi_D M - T(\pi_F^*) \tag{2.46}$$

By concavity π_D^* is increasing in $w(\pi_F^*, T(\pi_F^*))$ and thus $T(\pi_F^*) = 0$. However, π_D^* is increasing as $\frac{\partial w(\pi_F^*, T(\pi_F^*))}{\partial \pi_F}$ gets smaller. Thus $T(\pi_F^* + \epsilon) = \infty$. The optimal policy is thus a bang-bang tax policy of the form:

$$T(\pi_F) = \begin{cases} 0 & \pi_F \leq \overline{\pi}_F \\ \infty & otherwise \end{cases} \tag{2.47}$$

<u>THEOREM 3:</u> We first determine the largest reachable π_D at wealth $\hat{W}$ facing a tax policy $T(\overline{\pi}_F)$. Given the bang-bang nature of $T(\overline{\pi}_F)$, we can represent the effect of $T(\overline{\pi}_F)$ as a constraint $\pi_F \leq \overline{\pi}_F$. Using this representation, the owner solves:

$$\begin{gathered} \max_{\pi_D,\pi_F} u\Big((1-\delta)\hat{W} + \pi_F[P - M] - \pi_D M\Big) + \pi_D D_O \\ \textit{Subject To: } \pi_F + \pi_D \leq 1, \pi_F \leq \overline{\pi}_F, \pi_F, \pi_D \in [0,1]. \end{gathered} \tag{2.48}$$

Ignoring the boundary conditions and taking the derivative with respect to π_F, π_D

yields:

$$
\begin{aligned}
&(1)\tfrac{\partial}{\partial \pi_F} : [P-M]u'\Big((1-\delta)\hat{W} + \pi_F[P-M] - \pi_D M\Big) = \lambda_{\overline{\pi}_F} + \mu \\
&(2)\tfrac{\partial}{\partial \pi_D} : Mu'\Big((1-\delta)\hat{W} + \pi_F[P-M] - \pi_D M\Big) = -\mu + D_O
\end{aligned}
\tag{2.49}
$$

where $\lambda_{\overline{\pi}_F}$ is the lagrangian from the $\pi_F \leq \overline{\pi}_F$ constraint and μ is the lagrangian from the $\pi_F + \pi_D \leq 1$ constraint. Substitution for μ yields:

$$
[P]u'\Big((1-\delta)\hat{W} + \pi_F[P-M] - \pi_D M\Big) = \lambda_{\overline{\pi}_F} + D_O. \tag{2.50}
$$

When $\lambda_{\overline{\pi}_F} = 0$, we know that $\mu > 0$ and thus $\pi_F + \pi_D = 1$. Under these conditions:

$$
u'\Big((1-\delta)\hat{W} + \pi_F[P-M] - (1-\pi_F)M\Big) = \frac{D_O}{P}. \tag{2.51}
$$

If $\lambda_{\overline{\pi}_F} > 0$, then

$$
\begin{aligned}
&(1) : u'\Big((1-\delta)\hat{W} + \overline{\pi}_F[P-M] + \pi_D M\Big) = \tfrac{D_O}{M} - \mu. \\
&(2) : u'\Big((1-\delta)\hat{W} + \overline{\pi}_F[P-M] + \pi_D M\Big) = \tfrac{D_O}{M} - \tfrac{\mu}{M}.
\end{aligned}
\tag{2.52}
$$

These two equations can only both be satisfied when $\mu = 0$. Define $\pi_D(\hat{W}, \overline{\pi}_F)$ as the solution to

$$
u'\Big((1-\delta)\hat{W} + \overline{\pi}_F[P-M] + \pi_D M\Big) = \frac{D_O}{M}. \tag{2.53}
$$

$\pi_D(\hat{W}, \overline{\pi}_F)$ is increasing monotonically in $\overline{\pi}_F$. Thus, there exists a value $\overline{\pi}_F^*$ such that $\overline{\pi}_F + \pi_D(\hat{W}, \overline{\pi}_F^*) = 1$. At this point $\mu = 0$ since the constraint is just satisfied. Further, since $\frac{D_O}{P} < \frac{D_O}{M}$ and u is concave, π_D in the constrained problem will be larger than the level from the unconstrained problem.

Using the fact that $\overline{\pi}_F + \pi_D(\hat{W}, \overline{\pi}_F^*) = 1$ when π_D is maximized, let π_D^{**}, be the solution to

$$
u'\Big((1-\delta)\hat{W} + [P-M] + \pi_D M\Big) = \frac{D_O}{M}. \tag{2.54}
$$

It follows that

$$
\pi_D^{Reach}(\hat{W}) = \begin{cases} 1 & \pi^{**} > 1 \\ \pi^{**} & \pi^{**} \in [0,1] \\ 0 & \pi^{**} < 0 \end{cases}. \tag{2.55}
$$

If $\frac{D_O + D_E}{P} < \frac{D_O}{M}$ the first best is reachable and thus can be implemented. Otherwise, the optimal contract is the contract where π_D is as large as possible. This will be $\pi_D^{Reach}(\hat{W})$.

THEOREM 4: This result stems directly from Myerson's optimal mechanism. Let $\phi(F_i) = F_i - M + \frac{1}{\psi_A(F_i+M)}$ be the virtual valuation for buyer i. The optimal mechanism (Q, M) is an allocation rule Q and payment rule M such that:

$$\begin{aligned} Q_i &= \begin{cases} 1 & if\ \phi(F_i) > \max_{j\neq i} \phi(F_j)\ and\ \phi(F_i) > \phi(D_2) \\ 0 & otherwise \end{cases} \\ M_i &= max\{\phi^{-1}(D_2), max_{j\neq i} F_j\} \end{aligned} \tag{2.56}$$

An English auction with reservation price $P = D_2 - M + \frac{1}{\psi_A(F_i+M)}$ is an optimal auction which does not disclose the true valuation of the highest bidder in each period.

THEOREM 5: We show the proof for the simple case where $\delta = R = 1$ and $\gamma_0 = \gamma_t = 0$. The proof of the more complicated version is straightforward. The owner of an option, attempting to maximize his expected value selects a price P_{res} and reservation value r which maximizes:

$$\max_{P_{res},r} \underbrace{[1 - A(\tilde{P}_{res})]}_{\substack{\textit{Probability of} \\ \textit{Initial Sale}}} \Big[\underbrace{(1 - H(\tilde{r}))}_{\substack{\textit{Probability of} \\ \textit{Exercise}}} \underbrace{(P_{res} + E(D_2|D_2 > \tilde{r}) - \tilde{r})}_{\substack{\textit{Expected Value Conditional} \\ \textit{on Option Exercise}}} + H(\tilde{r}) P_{res} \Big] + A(\tilde{P}_{res}) \underbrace{[E(D_2) - M]}_{\substack{\textit{Value Conditional} \\ \textit{on No Initial Sale}}} \tag{2.57}$$

Subject to $r \geq P^{Res}$.

Taking the FOC with respect to the option price r and the inequality constraint we have:

$$\begin{aligned} \tfrac{\partial L}{\partial r} &: [1 - A(\tilde{P}_{res})] \big[- h(\tilde{r}) E(D_2|D_2 > \tilde{r}) - \tilde{r} \\ &\qquad + (1 - H(\tilde{r}))[\tfrac{d}{dr} E(D_2|D_2 > \tilde{r}) - \tilde{r}] + \tfrac{\lambda}{1 - A(\tilde{P}_{res})} \big] = 0 \\ \tfrac{\partial L}{\partial \lambda} &: (r - P_{res}) \geq 0, \lambda \geq 0, \lambda(r - P_{res}) = 0. \end{aligned} \tag{2.58}$$

Noting that

$$- h(\tilde{r}) E(D_2|D_2 > \tilde{r}) - \tilde{r} + (1 - H(\tilde{r}))[\frac{d}{dr} E(D_2|D_2 > \tilde{r}) - \tilde{r}] = 0, \tag{2.59}$$

Equation 2.58 reduces to:

$$- r[1 - A(\tilde{P}_{res})](1 - H(\tilde{r})) + \lambda = 0. \tag{2.60}$$

Since $a(\tilde{r}) > 0$, the SOC of $[1 - A(\tilde{r})]$ is positive. Thus $r = \infty$ is a global minimum which implies $\lambda > 0$ and $r = P_{res}$. The home country doesn't gain anything in leaving a separation in the price that it sells an object to a foreign collector and the price that it can rebuy the object in the future. The home countries problem thus reduces to:

$$\max_{P_{res}} \ [1-A(\tilde{P}_{res})]\left[(1-H(\tilde{P}_{res}))(E(D_2|D_2 > \tilde{P}_{res})-M)+H(\tilde{P}_{res})P_{res}\right]+A(\tilde{P}_{res})[E(D_2)-M] \tag{2.61}$$

Noting that $E(D_2) = E(D_2|D_2 \geq \tilde{P}_{Res})[1 - H(\tilde{P}_{Res})] + E(D_2|D_2 < \tilde{P}_{Res})H(\tilde{P}_{Res})$ this reduces to:

$$\max_{P_{res}} \ [1 - A(\tilde{P}_{res})][H(\tilde{P}_{res})]\left[(E(D_2|D_2 \leq \tilde{P}_{res}) - M + P_{res}\right] \tag{2.62}$$

Taking the FOC and using the simplification from Equation 2.59 we find:

$$P_{Res} = E(D_2|D_2 < \tilde{P}_{Res}) - M + \frac{1}{\psi_A(\tilde{P}_{Res})}. \tag{2.63}$$

<u>THEOREM 6:</u> With no credit constraints, the marginal utility of money is constant over time and thus the country is indifferent to the periods in which the contract generates payment. Leases with an English auction are surplus maximizing for the home country. Since the reservation prices for the option contract differ, it follows that leases are more efficient that options. The difference between the optimal lease and the optimal option is that under options, contracting is done before the home country knows its valuation D_2. This can generate situations in which the option price results in no trade in future periods even though, upon realization of D_2 trade would be optimal.

<u>THEOREM 7:</u> Since the future valuations of the objects are known, the reservation price of the option contract can replicate those of the lease auction perfectly. Further, since $\gamma_0 > \gamma_t$ for all t, the option contract dominates the lease auction by guaranteeing that all payments are received in period 0.

<u>THEOREM 8:</u> Without constraints, a generation that is reached without a bad Minister of Culture that is serviced by a good Minister of Culture gets expected value

$$max[P, Do + E[D_E]] = [1 - H(\hat{P})][D_o + E(D_E|D_E \geq \hat{P})] + H(\hat{P})P), \tag{2.64}$$

where $\hat{P} = P - D_o$ and H is the CDF of possible home valuations. The NPV of an object with a free market is:

$$\frac{1}{1-\delta(1-\epsilon)}\left[[1-H(\hat{P})][D_o + E(D_E|D_E \geq \hat{P})] + G(\hat{P})P\right]. \tag{2.65}$$

The NPV of an export ban is

$$\frac{1-\delta\epsilon}{1-\delta}[D_o + E(D_E)]. \tag{2.66}$$

The home country prefers an export ban if Equation 2.65 is less than equation 2.66. This condition is equivalent to requiring that

$$\hat{P} \leq E(D_E|D_E \leq \hat{P}) + \frac{\delta^2\epsilon(1-\epsilon)}{1-\delta}\frac{[D_o + E(D_e)]}{H(\hat{P})} \tag{2.67}$$

or

$$\underbrace{[H(\hat{P})\hat{P} + [1-H(\hat{P})]E(D_E|D_E \geq \hat{P}) - E(D_E)]}_{\textit{Per Period Gain From Flexibility}} \leq \delta \underbrace{\frac{\delta\epsilon(1-\epsilon)}{1-\delta}[D_o + E(D_E)]}_{\substack{\textit{Per Period Expected Loss} \\ \textit{From One Bad Dictator}}}. \tag{2.68}$$

At $\epsilon = 0$, the RHS of (2.67) is $E(D_E|D_E \leq \hat{P})$ which is less than $\hat{P}$ for $G(\hat{P}) > 0$. Thus, with no corruption, free trade is optimal. As $\delta \rightarrow 1$, for $\epsilon \in (0,1)$ the right hand side of (2.68) goes to infinity implying that an export ban is always optimal. Intuitively, the more patient a country is, the more it values the losses that occur if an object is stolen. As $\delta \rightarrow 1$ the losses that occur if an object is ever stolen weighs heavily in making a decision. This leads to a larger set of ϵ for which an export ban is optimal.

<ins>THEOREM 9:</ins> This result is the FOC of Equation 2.41.

Chapter 3

Experiments in Authority, Delegation, and Incentive Conflict

Abstract

A large portion of the organization literature has centered on how control rights are distributed in a setting with incentive conflicts. In this experiment, we use a simplified version of Aghion & Tirole (1997) to study the interplay between formal control rights, effort, and delegation. We find that individual effort choices and beliefs are consistent with the theoretical Nash Equilibrium but that principals retain control rights even when it is strongly in their interest to delegate. We also find differences in how males and females allocate and respond to the allocation of authority. As agents, women appear to have strong fairness preferences resulting in diminished effort in asymmetric treatments but higher effort in symmetric ones. As principals, women are more likely to transfer authority when it is efficient to do so. We show that a significant portion of these gender effects are due to differences in beliefs, a result consistent with a growing experimental literature documenting a differential response to incentives between genders.

3.1 Introduction

Economists have become increasingly interested in the interaction of individuals in environments where explicit incentives are absent, but where control rights may be granted to one party or the other to make a final decision. While a large theoretical literature has formed analyzing these incomplete contracting environments, empirical analysis has been limited by the complexities inherent in data from informal settings.

This paper uses experimental methods to study how individuals respond to authority and incentive conflicts. We propose a new experimental design based on the formal and real authority model developed in Aghion & Tirole (1997). As in their model, a principal and an agent must decide on the implementation of one project among a large list of potential candidates. The payoffs to the principal and agent for selecting a task is unknown ex-ante and effort must be exerted in order to learn the valuations of each project. One of the tasks is best for the principal while a different task is best for the agent. This misalignment of valuation leads to inefficient effort by the agent since, in cases where both parties gain information, her suggestions may be overruled. Delegation of authority to the agent increases the agent's effort but reduces the control that the principal has in selecting his optimal project.[1]

By using experiments, we generate a unique data set which includes typically unobservable primitives such as beliefs, counterfactual actions, and risk attitudes. Since these primitives are the basis on which theoretical models are based, our experiments allow for careful study of theoretical predictions and systematic capture of violations. Experimental methods also allow us to study the behavioral forces present. Given the one-on-one nature of the incomplete contracting problem, these behavioral forces are predicted to be acute based on findings from previous experiments and the psychology and behavioral organization literatures.

Our experiment consists of two players, a principal and an agent. The principal and the agent are shown a set of 36 cards representing potential projects. One of these cards has a small positive payout for both the principal and the agent and is placed face up representing the outside option. The remaining thirty-five cards are shuffled face down and include two cards with large positive values for both the principal and the agent, and thirty-three cards with zero payoffs. One of the high valued cards is best for the principal while the other card is best for the agent creating an incentive

[1] The literature on authority often assumes that monetary incentive contracts can not realign incentives because of either infinite risk aversion by subjects, non-pecuniary payoffs, or non-verifiability of the value of the selected project. We largely side step these issues by restricting the actions of the principal to delegation and effort.

conflict.

Play of the game is conducted in four main stages: (i) The principal decides whether to delegate formal authority to the agent over the future choice of projects; (ii) the principal and agent privately and simultaneously gather information about the face down projects' payoffs; (iii) the party who does *not* have formal authority recommends a project to the controlling party; and (iv) the controlling party picks a project or the outside option on the basis of his or her information and the information communicated by the other party.

We study the effect of incentive conflict on both effort choice and incentives for delegation by changing the degree of payoff alignment across treatments. In our low alignment treatment, the payoffs for the other parties preferred project is small leading to a large degree of incentive conflict and no incentive to delegate. In the high alignment treatment, the payoffs for the other parties preferred project is large leading to low amounts of incentive conflict and a small incentive to delegate. Finally, in the asymmetric treatment, the agent's preferred project has a high payoff for the principal while the principal's preferred project has a low payoff for the agent. This treatment allows us to test the effects of congruence under different authority structures and ensures that there is a large incentive to delegate within the treatment.

We find that conditional on keeping control rights, the principal's effort choices and beliefs are similar to the theoretical Nash Equilibrium but that principals retain control rights even when it is strongly in their interest to delegate. This reluctance to delegate is surprising since, based on the principals own beliefs, he would have a higher expected earnings delegating authority instead of keeping control.

Actions of the agent are less consistent with the Nash Equilibrium, with significant heterogeneity across individuals. A significant portion of agents put in zero effort when the principal maintains control rights suggesting a behavioral response to being the subordinate. When delegated to, agents put in more effort than is predicted by the Nash equilibrium while the principal puts in significantly less effort. These differences from theory leads to a significantly higher value to delegation than the initial experimental design.

We also find differences in how males and females allocate and respond to the allocation of authority. As agents, women appear to have strong fairness preferences resulting in diminished effort in asymmetric treatments but higher effort in symmetric ones. As principals, women are more likely to transfer authority when it is efficient to do so. We show that these gender effects are primarily due to differences in beliefs, a result consistent with a growing experimental literature documenting a differential response to incentives between genders.

This paper is related to three distinct literatures — the theoretical literature on authority and control, the experimental literature on behavioral contracting, and the experimental literature on gender and incentive systems. The theoretical literature on authority has moved steadily into unifying authority with the broader study of repeated games, communication, and organization. Baker, Gibbons & Murphy (1999) studies how the inability of the principal to grant formal authority to an agent can create incentives to renege on delegation in a static game. They stress that information structures generate different decisions due both to the information that is generated and to the different dynamic incentives for reneging on past promises. Rantakari (2008) and Alonso, Dessein & Matouschek (2008) study how governance structures differ in terms of the informativeness of information and efficiency when both adaptation and coordination are important for the principal and agents involved. We view our experiments as being an important and necessary first step to experimentally studying these more complex models. Our main contribution is in validating the core building blocks of these models while pointing out the importance of incorporating behavioral forces into the analysis.

Our finding of reduced effort by agents when the principal retains formal authority is similar to a newer experimental literature in behavioral contracting. Falk & Kosfeld (2006) find that in a game where a principal can constrain the actions of an agent, agents put in higher levels of effort, on average, when free to choose an action than when constrained. In their paper, a principal's distrust in the voluntary performance of an agent has a negative impact on the agent's motivation to perform well. In our paper, keeping authority leads to lower average effort than the Nash prediction while giving up authority leads to higher relative effort. While in their experiment the principal largely grants freedom to the agent, in our experiment the principal rarely relinquishes formal control. This result is surprising, especially in the asymmetric treatment, where giving the agent control results in a much higher expected value for the principal. Our result suggests that the cost of control may be overshadowed in many situations by the reduction in uncertainty and non-pecuniary value of having power.

Finally, our paper sheds further light on the differential effect of incentive systems by gender and status. In an experiment on competition, Gneezy, Niederle & Rustichini (2003) find that women have a weaker response to performance incentives than men in mixed groups, but that the difference in response is diminished in single sex designs. In an experiment of competition in India, where the effects of caste were studied, Hoff & Pandey (2005) find that when there is scope for discretion in rewarding performance, the announcement of caste leads to less competition among lower castes. One hypothesis that conforms to results of these experiments, and the newer work by Shurchkov (2008) and Niederle & Vesterlund (2007), is that a priori beliefs about the environment differs by gender and leads to different responses to incentives. Our findings largely support this hypothesis.

In Section 3.2, we quickly review a simplified version of the incentive conflict model of Aghion & Tirole (1997) and review its theoretical predictions. Section 3.3 details our experimental design. Section 3.4 reports the main results of our experiment and is separated into three parts. In section 3.4.1 we study how the effort choices of the principal and agent respond to the level of inherent incentive conflict. In section 3.4.2 we concentrating on differences in effort choices by gender and in section 3.4.3 we study delegation decisions by the principals. Section 3.5 concludes. We include summary statistics of all our experiments in the appendix.

3.2 The Model

This section describes a simplified version of the incentive conflict model of Aghion & Tirole (1997) with a few minor modifications to better fit a laboratory environment. The model is as faithful to the original model as possible with only minor changes in notation and is reproduced here for the convenience of the reader.

Consider a world in which a principal (he) and agent (she) are organized in a hierarchical structure and must decide to implement one or zero project out of a set of $n \geq 3$ potential projects. With each project $k \in \{1, \dots, n\}$, there is an associated noncontractible gain of P_k for the principal and a private benefit A_k for the agent. The agent's private benefit includes perks on the job, acquisition of human capital, the chance to signal ability, promotion consideration, or the disutility of implementing the project. If no project is implemented, the profit and private value of the project are both equal to a known outside value of P_0 and A_0 respectively.

Ex ante, all objects are identical and information must be collected about the value of the projects. For each party, the expected value for guessing randomly is less than their respective outside options. Thus, under the assumption of risk aversion or risk neutrality, the agent prefers to recommend inaction rather than guessing randomly. Similarly, an uninformed principal would never choose unilaterally to undertake a project.

The principal's preferred project yields known profit P_p to the principal and A_p to the agent where $P_p > P_0$ and $A_p > A_0$. Likewise, the agent's preferred project yields known benefit P_a to the principal and A_a to the agent with $A_a > A_0$ and $P_a > P_0$. We assume that for the principal, the principal's preferred project yields a strictly higher value to the agent's preferred project ($P_p > P_a$) and define the Principal's congruence parameter $\alpha = \frac{P_a}{P_p}$. Likewise, we assume that $A_a > A_p$ and define the Agent's congruence parameter $\beta = \frac{A_p}{A_a}$. Note that both congruence parameters belong to $(0, 1]$.

The principal is considered risk neutral and has utility $B_k - w$. The agent is also risk neutral and has utility $b_k + w$. Since the principal's benefit is noncontractible, the agent's wage is a constant and is used only to satisfy her participation constraint.[2]

We assume that the nature of projects' payoffs are initially unknown to both parties. The principal and agent acquires information in a binary form. At private cost $g_A(e)$, the agent learns her payoffs to all candidate projects with probability e. With probability $1 - e$, the agent learns nothing and still views the objects as identical. Similarly, the principal chooses how much time or effort to put in to learning payoffs. At private cost $g_P(E)$, he becomes perfectly informed about the payoffs with probability E and learns nothing with probability $1-E$. We assume that effort choices are made simultaneously, that both g_A and g_P are strictly convex,satisfy $g_i(0) = 0$ and $g_i'(0) = 0$ for $i = A, P$, and that $A_a - g_A(1) \leq A_0$, and $P_p - g_P(1) \leq P_0$.

Information in our model is *soft* so that information passed between parties can not be verified. Communication between the parties, therefore, must be interpreted as a pure suggestion for a project choice.

Formal authority represents the right to make the final decision. We analyze two cases, P-Formal authority in which the principal maintains formal authority and an A-Formal authority structure in which the principal delegates authority to the agent who can not be overruled. In the P-Formal case, a principal may always overrule the agent. He indeed does so if he is informed and if the agent's recommendation is not "congruent". Otherwise, he (optimally) rubber-stamps the agent's proposal any time he is not informed since $P_a > P_0$. Under A-Formal authority, the principal delegates authority to the agent. In this case, the agent now has formal authority and has the irrevocable right to make this particular decision.

[2]More complicated versions of this model may be considered in which monetary incentives do play a role. See, for instance section VB of Aghion & Tirole (1997). Our goal is to create as simple a model as possible in order to study how individuals respond to incentive conflicts.

The timing is as follows: (i) The principal proposes a contract that allocates formal authority to him or to the agent over the future choice of projects; (ii) the parties privately and simultaneously gather information about the n projects' payoffs; (iii) the party who does *not* have formal authority recommends a project to the controlling party; and (iv) the controlling party picks a project or the outside option on the basis of his or her information and the information communicated by the other party.

Under P-Formal authority, the utilities of the principal and agent are

$$u_P = E\hat{P}_p + (1-E)e\hat{P}_a + P_0 - g_P(E), \tag{3.1}$$

$$u_A = E\hat{A}_p + (1-E)e\hat{A}_a + A_0 - g_A(e), \tag{3.2}$$

where

$$\hat{P}_i = P_i - P_0, \ \ for\ i \in \{a,p\}, \tag{3.3}$$

$$\hat{A}_i = A_i - A_0, \ \ for\ i \in \{a,p\}. \tag{3.4}$$

Under A-Formal authority, the utility of the principal and agent are

$$u_P^d = (1-e)E\hat{P}_p + e\hat{P}_a + P_0 - g_P(E), \tag{3.5}$$

$$u_A^d = (1-e)E\hat{A}_p + e\hat{A}_a + A_0 - g_A(e), \tag{3.6}$$

where the d in the A-Formal utility functions stands for the mnemonic *delegation*.

3.2.1 Analysis and Theoretical Implications

Starting with the analysis of P-Formal authority, the reaction curves for the principal and agent become, respectively,

$$\hat{P}_p - e\hat{P}_a = g_P'(E) \tag{3.7}$$

and

$$\hat{A}_a - E\hat{A}_a = g_A'(e). \tag{3.8}$$

We assume that $\hat{A}_a\hat{P}_a < g_P''(E)g_A''(e)$ so that the system of equations has a unique, stable intersection (e^{NE}, E^{NE}). Since the agent's reaction curve is downward sloping, effort by the principal and agent are strategic substitutes. Thus, if the agent's effort increases, the principal's effort decreases.

Holding $\hat{A}_a$ and $\hat{P}_p$ constant, the principal's effort is decreasing in $\hat{P}_a$. The more value that the principal receives for the agents best action, the less incentive he has to interfere with the agent and overrule. This leads to an increase in effort by the agent since it is more likely that the project she suggests will be implemented. Notice also that neither the principal nor the agents best response correspondence depends on $\hat{A}_p$, the agents valuation under the principals best project. This implies that under P-Formal authority, effort choices are independent of the agents payment when the principal acts on his own information. Recalling that $\alpha = \frac{P_a}{P_p}$ is the congruence parameter for the principal and $\beta = \frac{A_p}{A_a}$ is the congruence parameter for the agent, we have the following prediction:

Theoretical Prediction 15 *Under P-Formal authority, effort of the principal is decreasing in α while the effort of the agent is increasing in α. Effort levels of the principal and agent are unaffected by β.*

Repeating the exercise under A-Formal authority, the reaction curves of the principal and agent are

$$\hat{P}_p - e\hat{P}_p = g'_P(E), \tag{3.9}$$

$$\hat{A}_a - E\hat{A}_p = g'_A(e). \tag{3.10}$$

Under similar stability and uniqueness criterion, there exists an interior intersection of $(e^{d^{NE}}, E^{d^{NE}})$ which constitutes the Nash equilibrium. Given the symmetry in the best response functions, it is unsurprising that under A-Formal authority, $\hat{P}_a$ does not affect the equilibrium. Mirroring prediction 1, we have:

Theoretical Prediction 16 *Under A-Formal authority, effort of the principal is increasing in β while the effort of the agent is decreasing in β. Effort choices of the principal and agent are unaffected by changes in α*

Finally, mapping the reaction functions in (e, E) space, where e is the horizontal axis, the slope of the reaction function for both the principal and agent are flatter under A-Formal authority than P-Formal authority. Since the intersection points with the origins are identical for the two functions, it must be the case that effort by the principal is less under delegation than under P-Formal authority.

Theoretical Prediction 17 *Delegation results in less effort by the principal and more effort by the agent. Delegation is optimal when $EV_P^d(E^{d^{NE}}, e^{d^{NE}}) > EV_P(E^{NE}, e^{NE})$.*

3.2.2 Behavioral Forces

Based on current experimental evidence, behavioral forces appear to be most acute in settings with a small number of parties. The ultimatum game, for instance, leads to a high level of rejection for low offers with a single principal offering contracts to a single agents but has a low rejection rate when the number of agents increases to two or more.[3]

Given that our game of interest is inherently a one-on-one interaction, the effect of non-standard preferences are likely to be large. In this section, we briefly present two behavioral effects which we believe a priori will influence the outcome of the experiment. We will return to these predictions in the analysis of section 3.4.

Efficiency versus Equality - Inequity Aversion

There is now a large body of experimental evidence that individuals are often willing to sacrifice social efficiency in order to normalize payments across parties. Fehr, Naef & Schmidt (2006), for instance, shows that third parties who have the choice between a socially efficient but unequal allocation and a socially inefficient but equal allocation often choose the later.

In our experiment, equity concerns will lead the principal and the agent to alter effort choices so that their expected earnings converge. In all treatments we consider, this implies that the party with the lower expected payment will decrease their effort level while the agent with the higher expected earnings will increase effort.

Behavioral Prediction 18 *Let E^{NE} and e^{NE} be the Nash Equilibrium of the model without behavioral effects. When $EV_P(E^{NE}, e^{NE}) > EV_A(E^{NE}, e^{NE})$ and Inequity Aversion is present, then $E > E^{NE}$ and $e < e^{NE}$. The difference $E - E^{NE}$ is growing in the difference in $EV_P(E^{NE}, e^{NE}) - EV_A(E^{NE}, e^{NE})$.*

Hidden Cost of Control & Gift Exchange

Positive and Negative reciprocity have been shown in a number of environments to be a strong force in principal-agent games. Falk & Kosfeld (2006) finds that in a game where the principal can reduce the influence that the agent has on the principal's

[3]See, for instance Roth, Prasnikar, Okuno-Fujiwara & Zamir (1991) who shows that behavior in a market version of the ultimatum game with 9 responders is radically different than the one-on-one version. See also Charness, Corominas-Bosch & Frechette (2007) for a study of bargaining over a network where the effects of social preferences appear to be mitigated.

payments, the agent punishes this form of control with less effort. Based on these results, we suspect that the act of keeping control may be seen as a hostile act while the transfer of authority may be seen as friendly. We thus predict:

Behavioral Prediction 19 *Let $e^{d^{NE}}$ and e^{NE} be the Nash Equilibrium effort levels of the agent without behavioral effects. Positive and Negative reciprocity would predict $e^d > e^{d^{NE}}$ and $e < e^{NE}$.*

3.3 The Experiment

In order to study how individuals respond to incentive conflicts and authority, we developed an experimental design based on the theoretical model summarized above. Our experiment consisted of two players, a principal and an agent. The principal and the agent were shown a set of thirty-six cards representing potential projects. One of these cards had a small positive payout for both the principal and the agent and was placed face up representing the outside option. The remaining thirty-five cards were shuffled face down and included a red and a blue card with large positive values for the principal and agent and thirty-three white cards with zero payoffs to both parties. Incentive conflict was generated by making the payout to the red card highest for the principal and the payout to the blue card highest for the agent. The white cards ensure that guessing had a lower expected return than selecting the outside option.

We studied the effect of incentive conflict on both effort choice and incentives for delegation by changing the degree of payoff alignment across treatments. In our low alignment treatment, the payoffs for the other parties preferred project was small leading to a large degree of incentive conflict and no incentive to delegate. In the high alignment treatment, the payoffs for the other parties preferred project was large leading to low amounts of incentive conflict and a small incentive to delegate. Finally, in the asymmetric treatment, the agent's preferred project had a high payoff for the principal while the principal's preferred project had a low payoff for the agent. This treatment allowed us to test the effects of congruence under A-Formal and P-Formal authority and also to ensure that there was a large incentive to delegate authority to the agent.

3.3.1 Protocol

All our experiments took place in May and June 2007 at the experimental laboratory of the Institute for Empirical Research in Economics at Zurich University. In total, 140 subjects participated in our experiments, divided into 5 sessions and 14 matching groups. A matching group consisted of 5 principals and 5 agents.[4] Within each matching group, principals and agents were randomly matched in each of the 10 periods played. Hence, within a matching group, each principal was matched with each agent twice.[5] Each subject maintained the role of either a principal or an agent throughout the experiment. To avoid framing, all instructions referred to the principal as player A and the agent as player B. Our sample pool consisted primarily of students at Zurich University and the University of Zurich and we excluded economics and psychology students in order to prevent potential confounds.

An experimental session consisted of an instructions phase, a single player stage game, the main experiment, and a final questionnaire. In the instruction phase, subjects were given a set of written instructions followed by a small quiz concerning payments. After all subjects had successfully completed the quiz, oral instructions summarizing the stage game were read aloud.

All subjects next played a 7 period single player variant of the main game. In the single player variant, players chose a search intensity, received information probabilistically based on their search intensity, and chose a project. We used search intensities from the paid periods of this single player experiment (periods 3-7) as a proxy for risk aversion and other idiosyncratic preferences throughout the analysis. The Nash equilibrium level of search was identical across all treatments of our experiment so that there is no predicted variation in this variable across treatment groups.

For the main experiment, subjects were exposed to a single treatment which remained constant across the 10 periods. We chose not to vary the payments within a session to minimize confusion and framing.

At the end of the experiment, subjects were paid based on their earnings in the last 5 periods of the single player variant of the game and all 10 periods of the main experiment. Subjects were then asked to take a short questionnaire in which we recorded gender, demographic information, and risk preferences. On average, an experimental session lasted 50 minutes with an average payment of 28 CHF ($25.50

[4]Due to recruitment problems, one group had 8 subjects and one group had 12 subjects.

[5]Three matching groups were run simultaneously to increase the perception of randomness and limit dynamic incentives.

at the time of the experiment).

3.3.2 The Principal-Agent Game

The stage game of our experiment was designed to capture as much information as possible about the participants effort choices and beliefs. The stage game was split into six stage, summarized in Figure 3-1:

Stage 1 - Delegation: In the delegation phase, the principal was asked whether he wished to retain control rights or to delegate them to the agent. Control rights were termed "the right to make the decision" to minimize framing.

Stage 2 - Search: In the search phase, the principal and agent were asked to choose their search intensities. Search intensities were bounded between $[0, 100]$ and corresponded to the probability that the player would be able to view all the cards. For the principal, we asked only for his search intensity conditional on his delegation choice. For the agent, we employed the strategy method in which the agent was asked for her effort choice for the case in which the principal kept control rights or delegated them.[6]

Stage 3 - Beliefs: In the beliefs stage, the principal and agent were asked their beliefs about the effort choices of the other party. We employed the strategy method for both the principal and the agent. For the principal, we first asked for his beliefs about the effort of the agent given his actual delegation decision. We then asked for his beliefs about the effort of the agent if he had chosen the opposite control right scheme. We did not use monetary incentives for beliefs to prevent potential hedging between search intensities and belief reports. Further, we did not capture beliefs about the delegation decision of the principal for fear of creating frames of resentment when delegation was kept.

Stage 4 - Resolution: In the resolution stage, the agent learned the delegation decision of the principal and the agents reported strategy for this authority structure

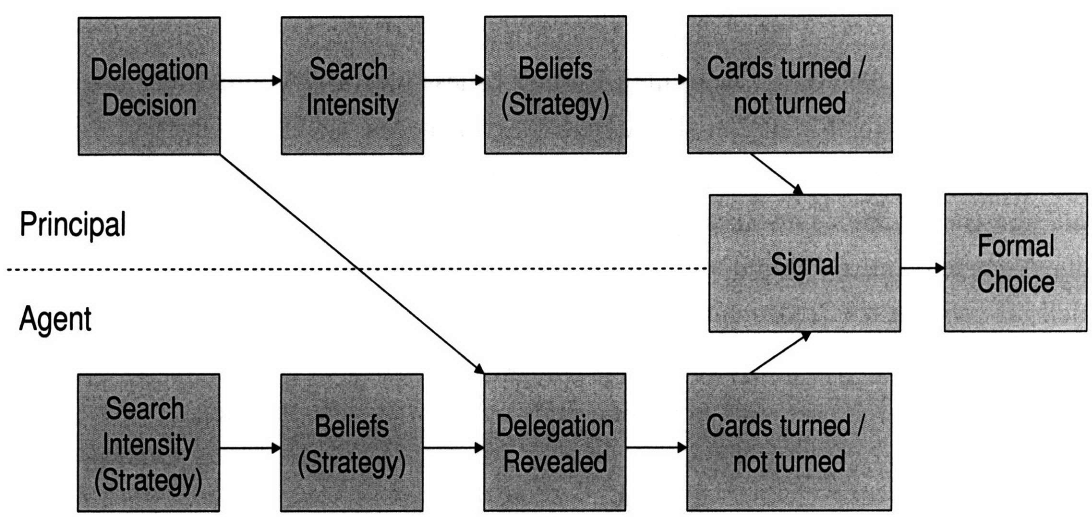
[6]The strategy method has the advantage of capturing a complete strategy profile for the agent but may lead to a greater level of self control and consistency than would exist if delegation was revealed prior to search. For instance, if the agent believes that there is a very low probability that the principal does not delegate, the cost of committing to zero effort in this case may be very small. Once the actual delegation decision has been realized, however, this cost is large. Future experiments will be run using the direct method to ensure this bias is not strongly influencing our results.

was carried out. If a player's search intensity was greater than a random draw from U[0,100], that player privately learned the location of all the cards. This information was private and not shared with the other player.

Stage 5 - Signal: The party who did not have formal authority was next given the ability to mark a single card as a recommendation to the other party. This marker, but not the location of the cards, was shown on the screen of the other player.

Stage 6 - Project Selection: If the principal retained authority, he was shown the card marked by the agent and allowed to select a project. Otherwise, the agent was shown the card marked by the principal and allowed to select the final project. Payment for the round was based on the card selected and the search intensity of the player.

Figure 3-1: The Stage Game



3.3.3 Treatments and Predictions

Experimental treatments were designed to study 1) how individuals responded to incentive conflict and 2) how individuals delegated authority. We divided experiments into three treatments that varied in P_a, the principal's valuation for the agent's preferred project, and A_p, the agent's valuation for the principal's preferred project.

Recall that in the set of 36 projects, there was one card representing the outside option, one card which was the agent's preferred project, one card which was the principal's preferred project, and 33 cards that yielded zero gross profit for both parties. The payments for each of the projects across our three treatments — Low, High, A.High — are recorded in Table 3.1.

Table 3.1: Overview of Project Payoffs

	High		**Low**		**A.High**	
	Principal	**Agent**	**Principal**	**Agent**	**Principal**	**Agent**
Principal's Preferred	40	35	40	20	40	20
Agent's Preferred	35	40	20	40	35	40
Outside Option	10	10	10	10	10	10
Unsuccessful Project	0	0	0	0	0	0

We studied the effect of incentive conflict on both effort choice and incentives for delegation by changing the degree of payoff alignment across treatments. In our low alignment treatment, congruence was low for both parties. Given Nash equilibrium search intensities, the principal's expected value for keeping control rights was higher than delegating. In the high alignment treatment, the congruence parameters for both the principal and agent were high leading to less incentive conflict and a small incentive to delegate. Finally, in the A.High treatment, the agent's preferred project had a high payoff for the principal while the principal's preferred project had a low payoff for the agent. This treatment allowed us to test the effects of congruence under A-Formal and P-Formal authority and ensured that there was a large incentive to delegate within the treatment.

Across treatments, we kept the cost of searching constant. For both the principal and the agent, search cost was increasing and convex with cost function

$$C(E) = 25E^2, \tag{3.11}$$

where a search intensity of E represented the probability of having the cards revealed to the player. Due to the quadratic cost function, the reaction curves of both the principal and agent were linear over effort. In the domain of possible effort choices, the Nash equilibrium predictions and expected values for each authority structure were:

Looking at the table, the expected value for retaining authority by the principal is greater than giving up authority in the Low alignment treatments but greater in the High and A.High treatments. We thus expect:[7]

[7]As we will see in the results, the actual increase in value for delegation in the high treatment is

Table 3.2: Predicted Efforts and Expected Values

	P-Formal				A-Formal			
	E^{NE}	e^{NE}	EV_P	EV_A	$E^{d^{NE}}$	$e^{d^{NE}}$	EV_P^d	EV_A^d
Low	55	25	20.1	17.3	25	55	17.3	20.1
High	45	35	23.3	24.0	35	45	24.0	23.3
A.High	45	35	23.3	17.2	25	55	25.6	20.1

Hypothesis 7 *Principal's will keep authority in the Low treatment and delegate authority in the High and A.High treatments*

In terms of effort choices, conditional on keeping authority, our theoretical model predicts that the principal's effort is decreasing in his payment for the agent's preferred project and should be unaffected by changes in payments to the agent for her preferred project. Theory Prediction 1 would predict, therefore, that under P-Formal authority, E in the low treatment should be significantly different from E in the High or A.High treatment. It also predicts that E in the High and A.High treatments should be the same. Since these are best reply functions, it follows that the agents effort levels have the same pattern with less effort under the Low treatment than the other two. We thus predict:

Hypothesis 8 *Conditional on the principal keeping control, we predict:*

$$\begin{aligned} E_{Low} &> E_{High} = E_{A.High}, \\ e_{Low} &< e_{High} = e_{A.High}. \end{aligned} \tag{3.12}$$

The predicted point estimates are given in Table 3.2.

Similarly, conditional on giving up authority, our theoretical model predicts that the agent's effort is decreasing in her payments when the principal's preferred project is selected. Based on our treatments, we predict that under A-Formal authority, the agent will put in lower effort under the High treatment than the Low and A.High treatments. It also predicts that the effort level between the Low and A.High treatments will be the same.

Hypothesis 9 *Conditional on delegation, we predict:*

$$\begin{aligned} E^d_{Low} &= E^d_{A.High} < E^d_{High}, \\ e^d_{Low} &= e^d_{A.High} > e^d_{High}. \end{aligned} \tag{3.13}$$

large. Based on the principals' beliefs, delegation in the high effort case would result in roughly a 20% increase in expected profit over keeping control.

The predicted point estimates are given in Table 3.2.

3.4 Experimental Results

Analysis of the experiment is divided into three parts: Effort Choices, Gender Differences, and Delegation Decisions.

1. **Effort Choices**: In section 3.4.1, we compare effort choices of the principal and agent to our Nash Equilibrium predictions. We find that the pattern of effort choices across treatments are consistent with our predictions in the Low and High treatments but are violated in the A.High treatment. We also find that when the principal keeps control rights, his effort choices can be explained in part by his beliefs about the actions of the agent. When the agent is delegated authority, however, the agent's effort choices are best predicted by their play in the single player version of the game. This difference can partially be explained by the fact that principal's who delegate often put in zero effort.

2. **Gender Differences**: In section 3.4.2, we study the differences in search intensities by gender with particular attention paid to the A.High treatment. We find that the violation of the Nash prediction in the A.High treatment can be primarily attributed to women. Differential preferences for efficiency and fairness across genders are consistent with all of our data and is a likely candidate for observed behavior.

3. **Delegation Decisions**: In section 3.4.3, we study the decision to delegate. We find that principal's retain authority even when it is strongly in their interest to delegate. As with effort, there is a gender difference in delegation, with women delegating authority closer to the optimal levels. This difference in delegation can partially be explained by a difference in beliefs about the agents actions when delegation does and does not occur.

While most of the summary statistics for each of the experiments can be found in the Appendix, it is worth noting two dimensions of heterogeneity in our data that we control for throughout the analysis. First, the average search intensity from the single player game is much larger for the principal in the Low alignment treatment than in the other two. Since individual level fixed effects would eliminate variation in gender, we instead use average effort in the single player game as a control for idiosyncratic preferences for search. For completeness, we show one version of each regression without this control variable.

Second, there is a slight downward trend in the agent's search intensity across periods. To control for this, we introduce period level fixed effects that are used throughout the analysis. Elimination of these fixed effects, or use of a linear trend, does not significantly alter the results.

As a simple check of rationality and whether subjects understood the incentives of the game, we look for situations where individuals made choices that appear to be suboptimal. There were 153 out of 700 cases in which the subject without decision rights searched successfully. Out of these, in 149 cases the subject recommended their preferred project. In the other 4 cases, the preferred project of the agent who had control rights was recommended. It was never the case that an unsuccessful project or the outside option was chosen. In the 547 periods in which the subject without control rights was uninformed, the outside option was chosen 533 times while a card was selected at random in 14 cases.

Subjects holding the decision right were informed in 374 out of 700 cases. In all 374 cases where the individual with decision rights was informed, he selected his preferred project. In 312 of the remaining 326 observations, the subject followed the recommendation of the agent. The outside option was selected instead of the recommended card 6 times while a face down card was chosen at random 8 times. We view these error rates as being relatively low and do not further pursue them in the analysis.

3.4.1 Effort Choices and the Nash Equilibrium

The theoretical model presented in section 3.2 yielded three predictions on the effort choices of the principal and the agent. First, the model predicted that P_a, the principal's valuation for the agent's preferred project, should affect effort choices of both the principal and agent under P-Formal authority, but should not influence effort choice under A-Formal authority. The theoretical model also predicted that A_p, the agent's valuation for the principal's preferred project, should affect effort levels under A-Formal authority, but not under P-Formal authority.

Before a direct analysis of these predictions, it is useful to get an overview of the data. Table 3.3 shows the average search intensities for the principals and agents, divided into treatments and authority structures. For ease of comparison, we include the Nash Equilibrium predictions from Table 3.2, the total number of observations within a treatment, and the percentage of periods in which authority was delegated.[8]

Table 3.3: Average Search Intensity Across Treatments

	P-Formal				**A-Formal**				**Delegation**	
	E	e	E^{NE}	e^{NE}	E^d	e^d	$E^{d^{NE}}$	$e^{d^{NE}}$	% Delegation	Obs_A
Low	62.6	16.3	55	25	14.2	71.4	25	55	15.2%	210
High	46.0	27.5	45	35	17.3	58.1	35	45	35.0%	340
A.High	56.2	18.6	45	35	22.5	62.8	25	55	37.3%	150

While a more careful analysis follows, a number of useful observations can be made by studying the averages reported in the table. First, the point estimates for the Low and High treatment are close to the theoretical predictions, with a slightly higher search intensity from the party who has control rights and slightly lower search intensity for the agent who does not.

Second, search intensities chosen under A-Formal authority appear to be farther away from the Nash Predictions than under P-Formal authority. In the A-Formal authority structure, effort levels for the Agent are near or above the single player optimum. This suggests that the Agent may either view search by the principal as unlikely, or is rewarding the principal for being given final power.

Finally, the search intensities in the A.High treatments do not fit the theoretical prediction in terms of magnitude or pattern. For P-Formal authority, search looks similar to the Low treatment. Under A-Formal authority, search intensities look similar to the High treatment. These outcomes are opposite of what the theory would predict, a topic that we will return to in section 3.4.2.

Search Decisions with Control Rights

To better understand the factors influencing search decisions, we split analysis between parties who have control rights and those that do not. Turning first to the situation in which the principal retains control rights, we start with a simple regression looking at search levels across treatment with controls only for idiosyncratic search preferences (via the single player game) and the time dimension. For principals

[8]Since the agent's effort was elicited using the strategy method, the total number of observations for the agent under both P-Formal and A-Formal authority equals the number of observations within the treatment. Since the principal's search intensity is conditional on his delegation decision, the number of observations for the principal under P-Formal authority is $Obs_A * (1 - \%Delegation)$. The number of observations for the principal under A-Formal authority is $Obs_A * (\%Delegation)$.

who did not delegate authority, we estimate:

$$E_{i,t} = \alpha_0 + \sum \alpha_t + \beta_{Single} Single + \beta_{High} I_{High} + \beta_{A.High} I_{A.High} \tag{3.14}$$

where $E_{i,t}$ is the search intensity of principal i in period t, α_t are the period fixed effects, *Single* is the demeaned average search intensity of the paid single player periods, and I_{High} is an indicator variable for treatments in the high alignment. Our a priori theoretical predictions is that $\beta_{High} = \beta_{A.High} < 0$.

Result 1 *The principal's search intensity is largely consistent with the theoretical Nash Equilibrium in the symmetric High and Low versions of the game, but not consistent with the Nash Equilibrium in the A.High treatment. Heterogeneity in beliefs explains a large portion of the deviations across treatments.*

Column (1) of Table 3.4 shows the results of this regression with period 1 as the omitted category for the time fixed effects. As can be seen from the first two rows, the high treatment is significantly different than the Low treatment with a magnitude in line with our theoretical predictions. The A.High treatment looks similar to the Low treatment in contrast to our a priori hypothesis.

Table 3.4: Principal's Search Intensity with Decision Rights[a]

	(1)	(2)	(3)
High	–13.813***	–4.898	-9.229**
	(3.612)	(3.664)	(4.579)
A.High	–3.102	2.714	-1.508
	(4.359)	(4.979)	(5.314)
Best Response to Beliefs		0.588***	0.506**
		(0.214)	(0.219)
Single	0.305*	0.366**	
	(0.167)	(0.173)	
Constant	58.901***	28.535**	34.940***
	(3.169)	(11.753)	12.17
Period Fixed Effects?	Yes	Yes	Yes
Adj. R^2	0.176	0.240	0.206
Observations	493	493	493

[a]Standard errors in parenthesis. Errors are clustered by individual.

To better understand the difference between the theoretical predictions and actual search decisions, we turn to our data on beliefs. Recall that in each period, we asked the principals to predict the search intensity of the agent in the case of kept control and delegation. Using this information, we construct the theoretical best response to beliefs across all three treatments:

$$E_{BR} = \frac{100\hat{P}_p - e^{belief}\hat{P}_a}{50}, \tag{3.15}$$

where e^{belief} is the principal's belief of the agents search intensity, $\hat{P}_p$ is the difference between the payment he receives when his preferred project is selected and the outside option, and $\hat{P}_a$ is the difference between the payment he receives when the agent's preferred project is selected and the outside option.

In our theoretical model, the principal's action should be perfectly predicted by a best response to his beliefs. In order to test this prediction, we estimate:

$$E_{i,t} = \alpha_0 + \sum \alpha_t + \beta_{BR}E_{BR} + \beta_{Single}Single + \beta_{High}I_{High} + \beta_{A.High}I_{A.High}. \tag{3.16}$$

We predict that $\beta_{BR} = 1$ and that all variation between treatments are subsumed into this variable ($\beta_{High} = \beta_{A.High} = 0$).

Column (2) of Table 3.4 presents the search intensity regression from Equation 3.16. While our theoretical prediction of $\beta_{BR} = 1$ is rejected, much of the variation across treatments is explained by a best response of actions to the principals beliefs. Note that even with beliefs taken into account, actions in the single game still have a large effect on the final outcome. This may in part be due to heterogeneity, in which a subset of our subjects either (i) do not update their actions in response to their beliefs or (ii) put a high probability weight on the agent putting in zero effort.

One concern with studying the search intensity of the principal is that the composition of individuals who chose to delegate across treatments may differ. An advantage of using the strategy method for the agents, is that there is no selection in their search decisions. As such, we turn to analyzing their search decisions next.

Result 2 *For the agent with control rights, the difference in search intensity between the High and Low alignment treatments match the theoretical predictions in both sign and magnitude. In contrast to the theory, the search intensity in the A.High treatment are closer to the High treatment rather than the Low treatment.*

Column (1) and (2) of Table 3.5 report identical regressions for the agent as those done for the principal in Table 3.4 above. As with the principal, the difference in search intensity between the Low and High treatments are both the right sign and

Table 3.5: Agents's Search Intensity with Decision Rights[a]

	(1)	(2)	(3)
High	–11.485***	–10.215***	-10.166*
	(3.210)	(3.025)	(5.446)
A.High	–9.446***	–9.495***	-8.708
	(3.090)	(3.134)	(6.731)
Best Response to Beliefs		0.126	0.304
		(0.114)	(0.205)
Single	0.812***	0.806***	
	(0.064)	(0.067)	
Constant	68.901***	61.955***	53.039***
	(2.066)	(6.561)	(11.725)
Period Fixed Effects?	Yes	Yes	Yes
Adj. R^2	0.549	0.551	0.093
Observations	700	700	700

[a]Standard errors in parenthesis. Errors are clustered by individual.

magnitude, but the search intensity of the A.High treatment is similar to the High treatment instead of the Low treatment.

In the Principal's regression, beliefs about the action of the agent largely influence the Principal's effort. For the agent, however, it is idiosyncratic differences in their actions in the single player game that is the main driver of heterogeneity across agents. The next section provides some evidence as to why both beliefs and the search levels of the single player game both factor into final search intensities.

Search Decisions without Control Rights

An immediate question that arises from the previous analysis is whether the beliefs from the principal and the agent with control rights are consistent with the true actions of the agent without control rights. We answer this question by looking at the actions of the principal or agent without control rights, which we refer to as the subordinate for the remainder of the analysis. Likewise, we refer to the individual who has decision rights as the superior.

The left hand side of Table 3.6 reports the average search intensity of the subordinate under P-Formal and A-Formal authority as well as the percentage of time zero search intensity was selected. Striking in this chart is the relatively large percentage of subordinates that put in zero search in each period. Recall that the cost for search is convex with $C'(0) = 0$. Since incremental search is nearly costless, zero search is never predicted by the Nash Equilibrium.

Result 3 *A large number of individuals without control rights put in zero search intensity. This is never a best response to the Superior's search efforts.*

Table 3.6: Search Intensities of Subordinates and Beliefs of the Superior

	Action of Subordinates				**Beliefs of Superiors**			
	P-Formal		**A-Formal**		**P-Formal**		**A-Formal**	
	Average	*%Zero*	*Average*	*%Zero*	*Average*	*%Zero*	*Average*	*%Zero*
Low	16.3	43.8%	14.2	59.4%	29.1	26.4%	21.8	31.0%
High	27.5	27.1%	17.5	39.5%	40.3	9.0%	29.1	22.7%
A.High	18.6	40.0%	22.5	46.4%	29.7	19.2%	19.6	38.7%

Given the large amount of time that the Principal chooses zero search after delegating, it is not a huge surprise that the decisions in the single player version of the search game is such a strong predictor of search in an A-Formal authority structure. If the Agent correctly estimates that there is a high chance of matching with a principal who puts in zero search, she selects a search level that is optimal in the single player game.

While our beliefs data appears to underestimate the proportion of agents who put in zero effort, it is important to note that we ask for a single point estimate rather than a range or distribution. Thus, depending on how the individual interprets the beliefs question, beliefs may be biased upward. This would be the case if, for instance, the agent answered the belief question with her belief about the principals effort conditional on effort being greater than zero.

3.4.2 Gender Differences in Search

Contrary to our theoretical predictions, search intensities in the A.High treatment tracked closer to the Low treatment when the principal kept control and the High treatment when control rights were delegated. In this section we show that a large portion of this effect is driven by gender. We find that as agents, the search intensities of women in the A.High treatment violates our theoretical prediction while the search intensities of men are consistent. These systematic differences also exist in the beliefs of the principal, suggesting that men and women view the incentive conflict in the A.High treatment differently.

As will be pointed out in the next section, the decision to delegate differed significantly between genders. As such, the sample of male and female principals who are making effort choices under P-Formal and A-Formal authority are likely to differ. We therefore concentrate solely on the actions of the agent and the beliefs of the principal, both captured via the strategy method.

Table 3.7 shows search intensities of the agent by gender in the case of the principal keeping control and delegating. As can be seen in the P-Formal columns, females put in significantly more effort in the High treatment than in the other two treatments, while males put in more more effort in the High and A.High treatments. As seen in the A-Formal columns, males put in lower effort levels in the High treatment while females put in lower effort in both the High and A.High treatments. For men, the pattern of effort choices across both P-Formal and A-Formal authority are consistent with the predictions from Hypothesis 2 & 3. For women, effort in the A.High treatment is lower than predicted.

Table 3.7: Search Intensities of Agents by Gender

	P-Formal		**A-Formal**		**Observations**	
	Male	*Female*	*Male*	*Female*	*Male*	*Female*
Low	15.2	17.6	68.4	74.7	110	100
High	23.5	31.4	55.2	61.0	170	170
A.High	21.0	16.4	65.9	60.1	70	80

Result 4 *The response by agents in the A.High treatment varies by gender. Men respond to the A.High treatment with search intensities consistent with the Nash Equilibrium. Women respond to the A.High treatment with lower effort than predicted both when the principal keeps authority and when it is delegated.*

A large part of our data, especially under P-Formal authority, has values in the bottom quartile of the variable range. This creates the potential that a small number of outliers may be driving our results. To formally test for differences across treatments, therefore, we employ a clustered version of the Rank-Sum Test developed by Datta & Satten (2005). This test is a direct extension of the Mann-Whitney-Wilcoxen test to clustered data without the need for averaging prior to analysis. Analysis using the median search intensity by individuals yields similar (slightly stronger) results.

Table 3.8 reports the two tailed probability that the pairwise treatment groups differ. While our sample size is small when the data is segmented by both treatment

Table 3.8: Pairwise Rank-Sum Tests of Search Intensity by Treatment

P-Formal[a]			A-Formal		
	Male	*Female*		*Male*	*Female*
Low vs. A.High	.435	.911	Low vs. A.High	.748	.214
High vs. A.High	.760	.129	High vs. A.High	.129	.946

[a]Clustered Rank-Sum Test using the procedure developed in Datta & Satten (2005)

and gender, there is some evidence that men and women respond to the asymmetric treatment differently.

There are a few possible explanations for this pattern of search intensities. One possibility, based purely on selection, is that the random sample of females in the A.High treatment is not representative. While a larger sample will reduce this possibility, beliefs data from the principals in the experiment shows that this explanation is unlikely. Table 3.9 shows the beliefs of the principal if he keeps authority and if he delegates. As with the actual actions of the agent, women believe that effort choices in the A.High treatment will be similar to the Low treatment under P-Formal authority and similar to the High treatment under A-Formal authority. This is the reverse of what is predicted from our theoretical model.

Table 3.9: Beliefs of Principal by Gender

	P-Formal		**A-Formal**		**Observations**	
	Male	*Female*	*Male*	*Female*	*Male*	*Female*
Low	34.1	23.3	62.0	68.3	120	90
High	38.1	31.1	57.7	52.4	180	160
A.High	33.6	22.3	66.8	55.6	80	70

A more likely reason for the gender difference is heterogeneous preferences for fairness and efficiency. In the A.High treatment, the difference in payments between the principal and agent under both P-Formal and A-Formal authority is greater in the A.High treatment when compared to both the Low and High treatments. If preferences between fairness and efficiency differ between genders, with females putting more weight on fairness, theory would predict that the search intensity of females would be reduced under both authority structures in the A.High treatment. This theory would also predict a larger change in search intensities between the Low and High treatment for women, a result also consistent with our data.

3.4.3 Delegation Decisions

Our initial experimental design predicted that with optimal search decisions, the principal has incentive to keep authority in the Low treatment and delegate authority in the High and A.High treatments. In this section, we study the delegation decision in detail. We start by confirming that given the actual search intensities of the agent, a principal who is maximizing expected value would choose to keep authority in the Low treatment and delegate in the High and A.High treatments. We then analyze the delegation decisions of the principal and relate these findings back to gender and beliefs.

Optimal Delegation

In order to verify that the delegation predictions of our experimental design hold in the actual experiment, we look at the best response of the principal given that he knows the average search intensities of the agent within the treatment. Table 3.10 shows the actual average search intensity of the agent, $\hat{e}$ and the best response to that search intensity, $E_{BR}(\hat{e})$, under P-Formal and A-Formal authority. As can be seen from the right hand columns, the expected value for keeping authority is higher in the Low treatment and lower in the High and A.High treatments.

Table 3.10: Principal's Best Response to Actual Agent Effort Levels

	P-Formal		**A-Formal**		**Expected Values**	
	$\hat{e}$	$E_{BR}(\hat{e})$	$\hat{e}^d$	$E_{BR}(\hat{e}^d)$	EV_P	EV_P^d
Low	16.3	55	71.4	15	**19.7**	17.9
High	27.5	45	58.1	25	22.2	**26.1**
A.High	18.6	50	62.8	20	21.1	**26.9**

While we have verified our experimental design using the optimal responses to the agents actions, it is interesting to note that our predictions are true even if we take into account (i) deviations from the best response after delegation and (ii) differences between beliefs and actual actions. Our predictions that delegation is optimal in the High and A.High treatments still hold if we repeat our analysis from Table 3.10 using any combination of the following changes:

1. Instead of the principal best responding to beliefs after delegation, we assume the principal puts in zero effort any time he chooses to delegate.

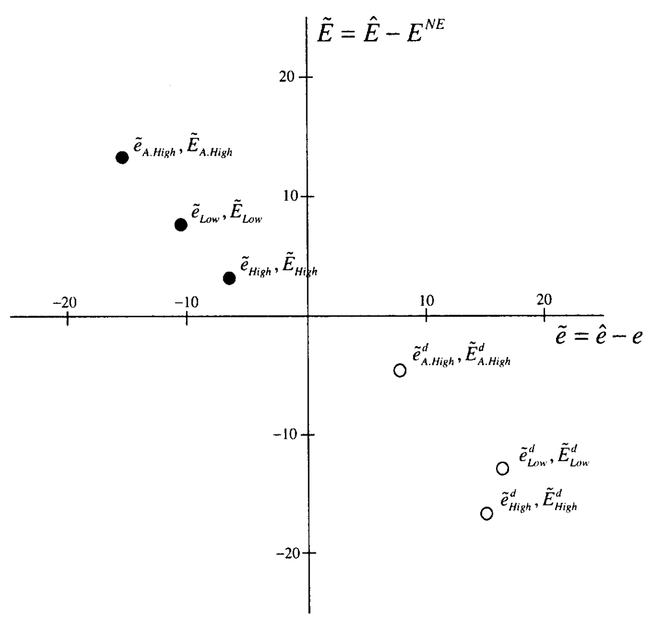
2. We use average beliefs of the principal instead of actual search intensities of the agent.

3. We use average beliefs of only the principals who do not delegate.

4. We use average beliefs of only males or females.

Figure 3-2, which shows the difference in search intensity observed versus the Nash Equilibrium, gives intuition as to why our delegation predictions are robust. Notice that when the control right is kept, $\hat{E} > E^{NE}$ and $\hat{e} < e^{NE}$. Under P-Formal authority, the decrease in $\hat{e}$ by the agent leads to a decrease in expected value for the principal. When control rights are delegated, however, $\hat{E}^d < E^{d^{NE}}$ and $\hat{e}^d > e^{d^{NE}}$. In the High and A.High treatment, even though the principal puts in too little effort when compared to the best response $BR_p(\hat{e}^d)$, the increase in effort from the agent more than compensates for this suboptimal response. In the Low treatment, high effort by the agent leads to a maximum of 20 points, which does not adequately compensate the principal for the loss of control.

Figure 3-2: Difference Between Actual Effort and Nash Equilibrium Prediction by Delegation Decision



The finding that $\hat{e} < e^{NE}$ and $\hat{e}^d > e^{d^{NE}}$ are consistent with our behavioral prediction of positive and negative reciprocity. When the principal keeps authority, reciprocity would predict that the agent punishes the principal with lower effort. Likewise, when the principal delegates to the agent, the agent rewards the principal with higher effort. Under this hypothesis, the increased punishment and decreased reward of the agent in the A.High treatment may be due to a difference in expectations about what the agent would do in the principal's position. In the high treatment, keeping control rights may be viewed as the expected action while delegation could be perceived as a friendly gesture. In the A.High treatment, delegating control rights may be seen as the expected action with kept control rights could be perceived as a hostile action.

An alternative hypothesis is that the act of delegation in some way signals the principals type. Given the high percentage of principal's who put in zero effort after delegation, it may be that the agent believes they will be matched with this type of principal if delegation occurs. This would lead to similar predictions as reciprocity, since different treatments may have a different subset of principals choosing to delegate and remain.[9]

Irrespective of the reason for the deviation from Nash equilibrium, the increase in effort by the agent after delegation leads to a higher expected earning for delegation in the High and A.High treatments. Deviations from our predicted pattern of delegation, therefore, are true violations of theory and not a direct result of violations in the subgame. To test for violations, we use a simple regression estimating the proportion of delegation across treatments. We estimate:

$$Delegation_{i,t} = \alpha_0 + \beta_{High} I_{High} + \beta_{A.High} I_{A.High}. \tag{3.17}$$

With a linear regression, our estimates would be $\beta_{High} = \beta_{A.High} = 1$ and the constant, representing the low treatment equal to zero. In a probit specification which takes into account the discrete nature of our dependent variable, estimates are in terms of deviations from the mean and thus we expect $\beta_{High} = \beta_{A.High} >> 0$ and the constant to be strictly below zero.

Actual Delegation

Column (1) and (2) of Table 3.11 reports the linear and probit regressions for Equation 3.17 respectively. While column (2) is the correct specification and will be used for extensions, the coefficients on the linear version of the regression are easier to interpret

[9] In future experiments we plan to isolate the effects by creating some randomness in the delegation decision. This will allow us to distinguish signals versus reciprocity.

and thus will be discussed here.

Table 3.11: Delegation Decisions by Principal

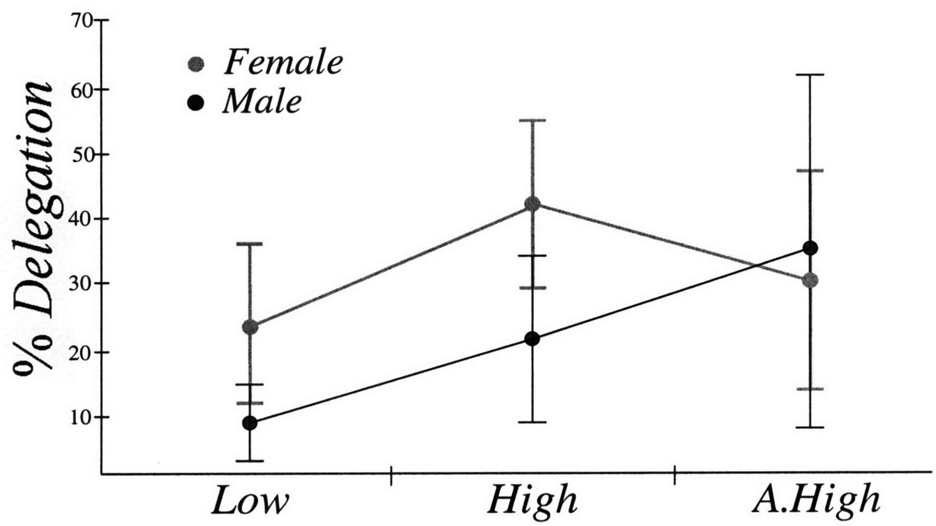
	(1) Linear	(2) Probit	(3) Probit	.(4) Probit
High	0.198***	0.641***	0.642***	0.778***
	(0.067)	(0.219)	(0.207)	(0.234)
A.High	0.221**	0.703**	0.713**	0.748**
	(0.099)	(0.290)	(0.292)	(0.311)
Female			0.376*	0.345*
			(0.206)	(0.208)
Belief under P-Formal				−0.009*
				(0.005)
Belief under A-Formal				0.009*
				(0.006)
Constant	0.152***	−1.026***	−1.212***	−1.572***
	(0.038)	(0.162)	(0.174)	(0.477)
Adj. R^2	0.040			
Psuedo-R^2		.038	.054	.087
Observations	700	700	700	700

Looking at the linear specification in Column (1), notice first that the constant in the linear specification is significantly greater than zero. In over 15% of our observations in the Low treatment, the principal delegated to the agent. This is surprising, given that even under an effort choice of 100 by the agent, the principal is worse off in delegating than his expected value playing a single player game. One possible interpretation of this result is that we have created an experimental demand effect in this treatment by giving principals an option that they never are expected to take. However, given that delegation differs by gender, we view this as unlikely.

The second interesting result, is that the delegation levels of the principal and the agent in the High and A.High treatments are far lower than predicted. In the High and A.High treatments, delegation occurred in only 35% and 37.3% of cases respectively. This low level of delegation is surprising, especially since the principal accurately predicts the effort levels of the agent under P-Formal and A-Formal authority.

To better understand the deviation of our prediction, we look at delegation decision by gender. Figure 3-3 shows the percentage of delegation divided by gender and treatment. Notice that in both the Low and High treatment, females are significantly more likely to delegate than males. This pattern is reversed in the A.High treatment, but the difference is not significant due to the size of our asymmetric sample. A joint test that the delegation decision of men and women are the same across all treatments is rejected at the .05 level.

Figure 3-3: Delegation by Gender



Result 5 *Women delegate more than men. This difference in delegation is due in part to more pessimistic beliefs about the effort choices of the agent under P-Formal authority.*

As with a growing literature on gender and competition, it appears that some of our difference in delegation can be explained by a difference of beliefs across genders. As we saw in table 3.9, females (correctly) predict that the agent will work less under P-Formal authority than predicted by theory. In Columns (3) and (4) of Table 3.11, we include gender and beliefs in our analysis of delegation. When beliefs are not taken into account, there is some evidence that females are more likely to delegate across treatments. Differences in beliefs across genders account for roughly 15% of this effect with the other 85% unexplained in our current data set.

3.5 Conclusion

We develop a new experimental design to study authority, delegation and incentives. Our results show that in many respects, the model of formal and real authority developed by Aghion & Tirole (1997) has some power in predicting how individuals respond to incentive conflicts. As predicted by theory, increased alignment of preferences leads to increased effort by agents and decreased effort by the principal, which translates into more influence for the agent in group decisions. Different from theory, an increase in congruence for the subordinate leads to an increase in the subordinates effort, especially for women. One potential reason for this effect is heterogeneity in the preference for fairness and efficiency, a hypothesis worth exploring in future work.

We have shown that the agents effort choices are strongly influenced by the delegation decision, with greater effort than predicted by the Nash equilibrium when delegation occurs and less effort than predicted when the principal maintains control. Consistent with the expanding literature on the hidden cost of control, these differences in effort across authority structures increase the value of delegating. Unlike other studies, we find that the principals in our experiment do not take advantage of these differences in effort and keep authority when it is optimal to delegate. This is even more surprising since the principal correctly forecast the agents' actions. Given the importance on the optimal transfer of control rights to the incomplete contracting literature, a more careful analysis of the forces generating inefficient delegation is warranted.

Finally, our results provide evidence of gender differences that is consistent to the larger experimental literature on gender and competition. We find that much of the gender difference in delegation is a result of different beliefs about the agents effort choices. In our experiment, men tend to overestimate the level of search conducted by the agent while women's beliefs tend to be close to actual levels. This difference in beliefs explains some, but not all, of the difference in delegation across gender.

3.6 Appendix

3.6.1 Instructions (Translated from German)

Instructions for Participant A

You are about to participate in a scientific study. A research foundation has provided funds for conducting this research. Please read the following instructions carefully. The instructions inform you about everything you need to know to participate in the study. If you do not understand something, please raise your hand and wait for an instructor, he will answer the question at your terminal.

For participating in this study and arriving on time, you have already earned a show up fee of **10 Swiss Francs.** During the study you may receive additional money by earning **Points.** The amount of points you earn during the study will depend on your decisions and the decisions of other participants.

All points that you earn during the study will be converted to Swiss Francs at the end of the experiment. The conversion rate is

10 Points = 1 Swiss Franc

At the end of the study you will receive the amount of money that you earned during the experiment plus the 10 Swiss francs show up fee.

During the study, it is strictly forbidden to communicate with each other. In addition, please use only the functions on the computer which relate directly to the study. Communication or using the computer in a way unrelated to the study will lead to exclusion from the study. If you have questions we are happy to assist you.

<u>This study consists of two parts:</u>

<u>Part 1:</u>
The first part of the study lasts for **7 periods.** The first two of these are practice periods and won't be paid. In this part of the study you make decisions on your own and your decisions have no consequences for other participants. Equally, the decisions of the other participants have no consequences for you.

<u>Part 2:</u>
The second part of the study lasts for **10 periods.** In the second part of the study, you and the rest of the participants have been divided into two groups: Participant As and Participant Bs.

<u>During the whole study you will be a Participant A.</u>

In each period you will be matched with a different participant from group B. At no time will you or any other participant be informed of the identity of the individuals that you match with.

The detailed instructions of part 1 and part 2 of this experiment are similar. For this reason, the instructions will first explain how to play the game in the second part of the study. At the end of the instructions, the specific details about the first part will be given.

Summary of Part 2

In this part of the study you will be matched with a different participant from group B in each of 10 periods. In each period, you and the other participant have the task of picking a single card out of 35 possible choices. Initially, all but one of these cards will be shuffled and turned faced down so that you and Participant B can not see their location. A single **Green Card** will remain face up.

The 34 cards that are face down contain:
1. The Red Card
2. The Blue Card
3. 32 Blank Cards

Each period has the following steps:

1. **Step 1 - Keep or Transfer the Decision Right:** In each period, as Participant A you have the right to decide the card chosen at the end of the period.

   - You may keep the right
   - Or you can transfer the right to Participant B. In this case, Participant B has the right to decide the card position chosen at the end of the period.

2. **Step 2 - Search:** After you either kept or transferred the decision right to Participant B, you and Participant B can, separately from each other and at a cost, search for the Red and Blue Cards.

   - If your search is successful, you will be informed about the positions of the Red and Blue card.
   - If your search is unsuccessful, all cards but the Green card remain covered.

   The same conditions apply to participant B. However, you will not be informed about the success of B's search. Equally, B will not be informed about the success of your search.

3. **Step 3 – Recommendation:** The participant without decision right can recommend a card position to the other participant. The recommendation will be transmitted to the participant who holds the decision right, before he makes his decision.

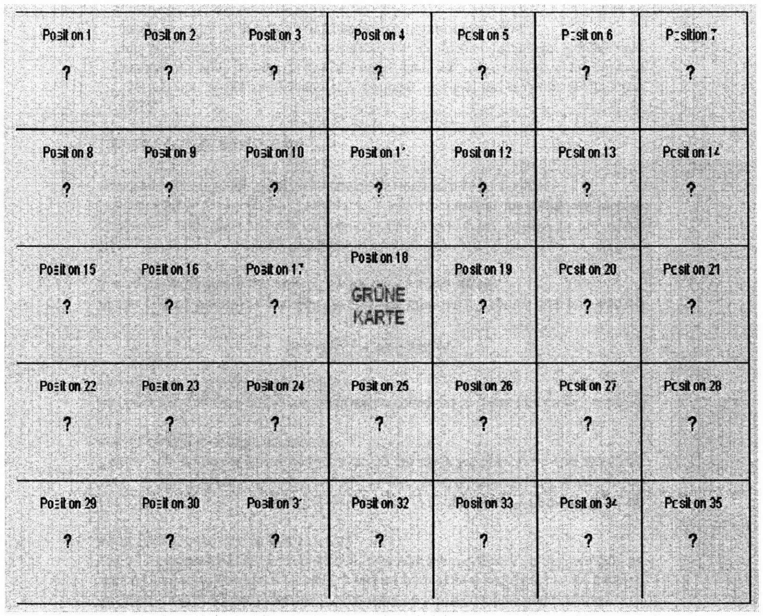
4. **Step 4 – Card Selection:** The participant with decision right chooses a card.

5. **Step 5 – Income:** You and participant B are informed about your incomes in this period.

Detailed Description

The Setup

The screen below shows the cards at the beginning of each period. 35 cards are reshuffled in each period and place at a random position. Only **the Green Card's location remains fixed at position 18** and is initially visible to you and participant B:



The Cards

In each round there are four kinds of cards: (1) The **Green Card,** (2) **The Red Card,** (3) **The Blue Card,** and (4) **32 Blank Cards.** At the start of each period, all but the green card are randomly shuffled and placed face down at random positions. At the end of each period, you or Participant B chooses one of these Card Positions. The card selected has payment consequences for you and participant B:

- Blanks: You and participant B get 0 points.
- The Blue Card: You get 35 points, Participant B gets 40 points
- The Red Card: You get 40 points, Participant B gets 35 points
- The Green Card: You get 10 points, Participant B gets 10 points

Card Overview

Card	Your Earnings	Earnings Participant B
Blue	35	40
Red	40	35
Green	10	10
Blank	0	0

Step 1: You can transfer or keep the decision right

Either you or participant B chooses a card position at the end of the period. The chosen card determines the earnings of you as well as participant B. At the beginning of each period you hold the decision right. You can

- Keep the decision right
- Transfer the decision right

If you keep the decision right, you make the final decision about the chosen card. If you transfer the right, Participant B makes this decision.

Step 2: The search for cards

In each period, you can search for the Blue and the Red cards. **If your search is successful, all cards will be turned and you will be able to see the positions of the Red and Blue Card.** If your search is unsuccessful you will, as before, only know the position of the Green Card. All other cards remain covered.

Participant B also has the possibility to search for the Red and the Blue Card. Your search and the search of Participant B are completely independent from each other. You will not be informed as to whether participant B searched successfully, and Participant B does not know whether you searched successfully.

If, for example, participant B searched successfully but you did not, only participant B is informed about the position of the Red and the Blue Card. You do not receive this information.

How does search work?

You and Participant B can independently **choose a search intensity between 0 and 100. The search intensity equals exactly the probability, with which all cards are turned.**

$$0 \leq \text{Search Intensity} \leq 100$$

A search intensity of 0 means that the cards will be NEVER turned. A search intensity of 100 means that the cards will be ALWAYS turned. For intermediate values it may happen that the cards are turned or not

The cost of search

The higher you choose the search intensity, the higher are your costs. The costs of participant B are identical to your costs. The following table shows the costs for every possible search intensity. It is only possible to choose search intensities in increments of 5.

Search intensity	0	5	10	15	20	25	30	35	40	45	50
Costs in points	0	.06	.25	.56	1	1.56	2.25	3.06	4	5.06	6.25

Search intensity	55	60	65	70	75	80	85	90	95	100
Costs in points	7.56	9	10.56	12.25	14.06	16	18.06	20.25	22.56	25

Please always consider this table when you choose search intensities. Your costs will remain the same throughout the experiment.

The higher you choose the search intensity, the more likely it is that the cards will be turned and you are informed about the position of the Red and the Blue Card. But your costs are also higher, the higher you choose the search intensity

The Success of Search:

Based on your chosen search intensity, the computer will determine whether search is successful as follows:

Your chosen search intensity is between 0 and 100. The computer then randomly draws one out of a hundred balls, which are numbered from 1 to 100. If the drawn number is smaller or equal your chosen search intensity, all cards will be turned. If the number is larger then your search intensity, no card will be turned. Hence, the search intensity equals exactly the probability with which all cards will be turned.

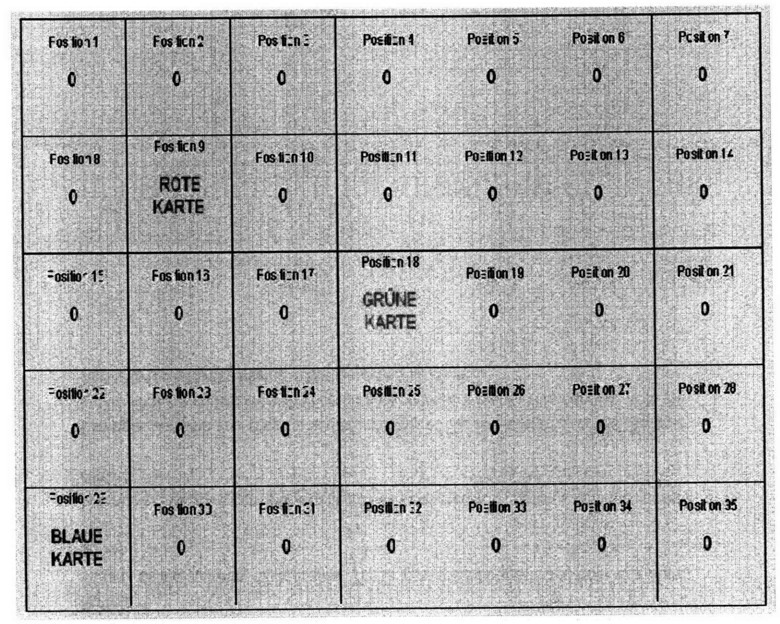
Examples:

1. You choose a search intensity of 15:
   If the randomly drawn ball has a number between 1 and 15 (=15 out of 100 balls) all cards will be turned. If the number is larger than 15 (16-100, and therefore 85 balls), no cards will be turned.

2. You choose a search intensity of 75:
   If the randomly drawn ball has a number between 1 and 75 (=75 out of 100 balls) all cards will be turned. If the number is larger than 75 (76-100, and therefore 15 balls), no cards will be turned.

At no point in time will you be informed about the search intensity of participant B. Equally, participant will ever be informed about your search intensity.

If your search was successful, all cards will be turned and you know the positions of the Red and the Blue cards. You will then see the following screen (example):



Step 3: The recommendation

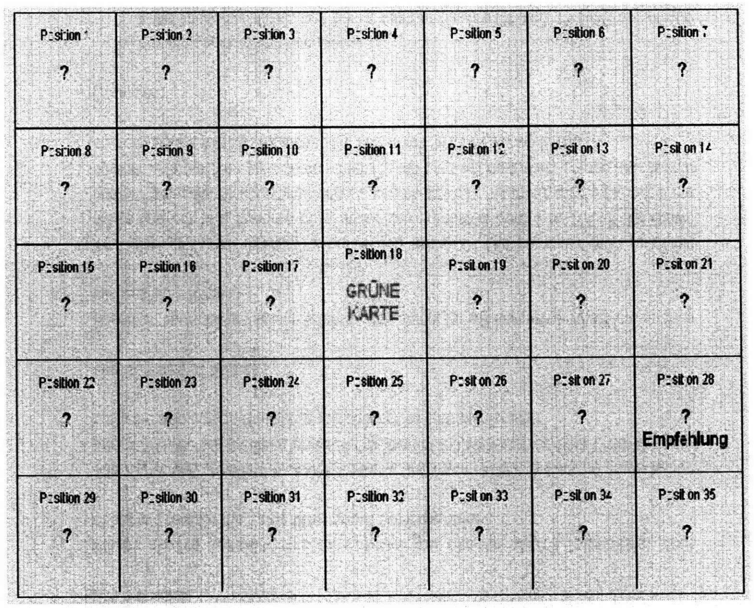
After the search, the participant without decision right can **recommend a card position** to the other participant.

- If you transferred the decision right, you can send a recommendation to Participant B.
- If you kept the decision right, Participant B will send a recommendation to you.

Depending on whether the recommender's search was successful, he will either be informed and know the location of all the cards or be uninformed and know the position of only the Green Card. Independent of this information, he can recommend any position.

The participant with decision right is only informed about the recommended card position, and not, which card it is.

Example: The following screen shows how the recommendation is transmitted to the participant with decision right. (In the shown example the search of the participant with decision right was unsuccessful):



Step 4: Choosing a card

The participant with decision right can decide at the end of each period, which card is chosen.

- If you kept the decision right, you make this decision.
- If you transferred the decision right, participant B makes this decision.

If the search of the participant with decision right has been successful, he knows the position of the Red and the Blue card. If search was unsuccessful, he does not know these positions. In addition, he knows the recommendation of the other participant.

The participant with decision right can then, with this information, choose a card and the participants earn the points associated with that card minus the costs of search.

Step 5: The Incomes

The income of both participants is determined by the following two parts:

- The income associate with the chosen card
- Minus the costs of search

Income in a period = Income from chosen Card – Costs of search

Notice:
In each period you may make losses! These losses will be subtracted from your show up fee.

You make a loss in a period if your search costs exceed your earnings from the chosen card.

Example: Assume you chose a search intensity of 50. This costs 6.3 points. Assume further, that the participant with decision right chooses a Blank. Your earnings from the card is 0. Hence, your income in that period would be -6.3.

Summary of one Round

1) At the beginning of every period you decide, whether you want to transfer the decision right to participant B or whether you want to choose a card yourself.

2) Thereafter you and participant B can, independent of each other and at a cost, search for the position of the Red and the Blue Cards.

3) Both participants are informed about the success of their own search. They do not get any information about the search and the success of the other participant.

4) The participant without decision right can recommend a card position to the participant with decision right

5) The participant with decision right receives the recommendation of the other participant. Thereafter he can choose the final card.

6) The earnings associated with the chosen card are realized, and you and participant B are informed about your incomes.

7) You are randomly matched with another participant B, and a new round starts.

The first part of the study:

1) In the first part of the study, you are NOT in a group with a participant B. Hence, you always have the right to decide, which card is chosen. You cannot transfer it.

2) In each period you choose a search intensity, to find the positions of the Blue and the Red card. The cost of search and the earnings associated with the cards are identical to the second part of the study.

3) You are informed about the success of your search and you can choose a card. Since there is no second participant, there is also no recommendation.

4) The earnings associated with the chosen card are realized. Only the payment to you (participant A) is relevant. Since there is no participant B, only your points are paid out.

5) The first two periods are for practice purposes only, so you can get familiar with the program. You earnings from these two periods will not be paid out. Your income from the following 5 periods, together with your income from the 10 periods in the second part of the study, will be paid out to you at the end of the study.

Control Questions:

Please answer the following control questions. Write down all calculations you make. If you have questions, please contact an instructor.

1. You kept the decision right and you chose a search intensity of 80. Your search was successful. Participant B recommended to you to choose the Green Card (position 18). You decided to choose the RED Card.

What are your search costs? ....
What is your final income? ....

2. You kept the decision right and you chose a search intensity of 30. Your search was unsuccessful. Participant B recommended to you to choose position 32. You decided to choose position 32. It is the BLUE Card.

What are your search costs? ....
What is your final income? ....

3. You kept the decision right and you chose a search intensity of 30. Your search was unsuccessful. Participant B recommended to you to choose position 24. You decided to choose position 28. It is a BLANK.

What are your search costs? ....
What is your final income? ....

4. You transferred the decision right and you chose a search intensity of 40. Your search was successful. You recommend position 23 to participant B (the RED card). Participant B chooses position 27. It is the BLUE Card.

What are your search costs? ....
What is your final income? ....

5. You transferred the decision right and you chose a search intensity of 40. Your search was successful. You recommend position 7 to participant B (the RED card). Participant B chooses position 7. It is the RED Card.

What are your search costs? ....
What is your final income? ....

3.6.2 Summary Statistics

1. **Session:**
   (a) Sessions 1-4 are Low treatments.
   (b) Sessions 5-11 are High treatments.
   (c) Sessions 12-14 and A.High treatments.
2. E,E^d: Average effort of the principal conditional on no delegation and delegation.
3. e,e^d: Average effort of the agent if the principal under no delegation and delegation. Elicited by Strategy Method.
4. **Principal: Male Female** Total number of male and female principals in a session.
5. **Agent: Male Female** Total number of male and female agents in a session.
6. **% Delegation** Percentage of periods in which control rights were delegated to the agent.
7. **Single** Average effort exerted in the single player version of the game.

Table 3.12: Summary Statistics for Low Treatments

	P-Formal		A-Formal		Principals		Agents			
Session	E	e	E^d	E^d	Male	Female	Male	Female	% Delegation	Single
1	65	8.2	15	76.9	4	2	2	4	6.7	67.9
2	57.1	25.0	24.3	68.7	3	2	3	2	14.0	58.8
3	56.1	19.1	7.5	68.3	2	3	4	1	28.0	60.8
4	69.9	14.6	17.2	70.6	3	2	2	3	14.0	67.2
Total	62.6	16.3	14.2	71.4	12	9	11	10	15.2	64.1

Table 3.13: Summary Statistics for High Treatments

	P-Formal		A-Formal		Principals		Agents			
Session	E	e	E^d	E^d	Male	Female	Male	Female	% Delegation	Single
5	33.4	21.6	15.3	44.8	0	5	2	3	68.0	46.9
6	38.5	20.9	7.9	62.0	3	2	4	1	28.0	57.2
7	52.4	35.5	13.5	64.0	2	3	2	3	34.0	58.1
8	43.8	42.5	12.3	66.1	4	1	2	3	26.0	73.1
9	45.2	39.6	36.0	57.3	3	2	1	4	10.0	57.1
10	57.6	15.5	24.5	63.3	3	2	4	1	42.0	61.6
11	48.0	14.1	22.7	47.3	3	1	2	2	37.5	58.0
Total	46.0	18.6	17.3	58.1	18	16	17	17	35.0	58.9

Table 3.14: Summary Statistics for A.High Treatments

	P-Formal		A-Formal		Principals		Agents			
Session	E	e	E^d	E^d	Male	Female	Male	Female	% Delegation	Single
12	61.3	11.8	8.3	73.7	2	3	3	2	48.0	59.9
13	54.5	29.5	14.0	63.0	2	3	2	3	20.0	56.9
14	53.9	14.4	41.8	61.8	4	1	2	3	44.0	63.4
Total	56.6	18.6	22.5	62.8	8	7	7	8	37.3	60.0

Table 3.15: Avg Search in Single Player Game

	Principal	**Agent**
Low	68.4	63.5
High	59.2	61.3
A.High	59.2	64.2

Bibliography

Acemoglu, D. & Robinson, J. A. (2006), Persistence of power, elites and institutions, Working Paper 12108, National Bureau of Economic Research.

Aghion, P. & Tirole, J. (1997), 'Formal and real authority in organizations', *Journal of Political Economy* **105**(1), 1–29.

Akinade, O. (1999), 'Illicit traffic in cultural property in Nigeria: Aftermaths and antidotes', *African Study Monographs* **20**(2), 99–107.

Alonso, R., Dessein, W. & Matouschek, N. (2008), 'When does coordination require centralization'.

Anderson & Holt (1997), 'Information cascades in the laboratory', *American Economic Review* **87**(3), 797–817.

Appiah, A. (2006), 'Whose culture is it?', The New York Review of Books, Feb. 9.

Argenziano & Gilboa (2006), History as a coordinating device, Working Paper 8-2006, The Foerder Institute for Economic Research.

Bagli, C. (1993), 'Hot frieze, missing since 1989, found in madison's fortuna', N.Y. Observer, Aug. 16.

Baker, G., Gibbons, R. & Murphy, K. (1999), 'Informal authority in organizations', *Journal of Law, Economics, and Organization* **15**(1), 56–73.

Banerjee, A. (1992), 'A simple model of herd behavior', *Quarterly Journal of Economics* **107**(3), 797–818.

Bator, P. (1982), 'An essay on the international trade of art', *Stanford Law Review* **34**(2), 275–384.

Beech, H. (2003), 'Spirited away', Time, Oct. 20.

Bikhchandani, Hirshleifer & Welch (1992), 'A thoery of fads, fashion, custom, and cultural change as informational cascades', *Journal of Political Economy* **100**(5), 992–1016.

Boehn, M. (2005), 'New blockbusters: Traveling exhibits.', NPR: Talk of the Nation.

Borodkin, L. (1995), 'The economics of antiquities looting and a propsosed legal alternative.', *Columbia Law Review* **95**(2), 377–417.

Bowman, D., Minehart, D. & Rabin, M. (1999), 'Loss aversion in a consumption-savings model', *Journal of Economic Behavior and Organization* **38**(2), 155–178.

Butcher, K. & Gill, D. (1990), 'Mischievous pastime or historical science', *Antiquity* **64**, 946–50.

Chamley (1999), 'Coordinating regime switches', *The Quarterly Journal of Economics* **114**(3), 869–905.

Charness, G., Corominas-Bosch, M. & Frechette, G. (2007), 'Bargaining and network structure: An experiment.', *Journal of Economic Theory* **136**, 28–65.

Coggins, C. (1972), 'Archeology and the art market.', *Science* **175**(4019), 262–266.

Datta, S. & Satten, G. (2005), 'Rank-sum tests for clustered data', *Journal of the American Statistical Association* **471**(1), 908–915.

Dewatripont, M. (1989), 'Renegotiation and information revelation over time: The case of optimal labor contracts.', *Quarterly Journal of Economics* **104**, 589–619.

Dodonova, A. & Khoroshilov, Y. (2005), 'English auction with loss averse bidders', Working Paper: SSRN: http://ssrn.com/abstract=922534.

Ellison, G. (1993), 'The evolution of conventions', *Econometrica* **61**(5), 1047–1071.

Erdem, S. (1993), 'A new trojan war.', Reuters, Oct. 13.

Falk, A. & Kosfeld, M. (2006), 'The hidden cost of control', *The American Economic Review* **96**(1), 1611–1630.

Fehr, E., Naef, M. & Schmidt, K. (2006), 'Inequality aversion, efficiency, and maximin preferences in simple distribution experiments: Comment.', *The American Economic Review* **96**(5), 1912–1917.

Fernandez-Val, I. (2007), 'Fixed effect estimation of structural parameters and marginal effects in panel probit models', mimeo.

Fischbacher, U. (2007), 'z-tree: Zurich toolbox for ready-made economic experiments', *Experimental Economics* **10**(2), 171–178.

Foster & Young (1990), 'Stochastic evolutionary game dynamics', *Theoretical Population Dynamics* **38**(1), 219–232.

Gale, D. (1992), 'A walrasian theory of markets with adverse selection', *Review of Economic Studies* **59**(2), 229–255.

Gerstenblith, P. (2001), 'The public interest in the restitution of cultural objects (symposium, ownership and protection of heritage: Cultural property rights for the 21st century)', *Connecticut Journal of International Law* **16**, 197–246.

Gneezy, U., Niederle, M. & Rustichini, A. (2003), 'Performance in competitive environments: Gender differences', *Quarterly Journal of Economics* **98**(1), 1049–1074.

Goeree, Palfrey, Rogers & McKelvey (2007), 'Self-correcting information cascades', *Review of Economic Studies* **74**(3), 733–762.

Greenfield, J. (1996), *The Return of Cultural Treasures*, Cambride University Press, Cambridge.

Hart, O. & Tirole, J. (1988), 'Contract renegotiation and coasian dynamics.', *Review of Economic Studies* **55**, 509–540.

Hoff, K. & Pandey, P. (2005), 'Opportunity is not everything: How belief systems and mistrust shape responses to economic incentives', *Economics of Transition* **13**(3), 445–472.

Holt, Langan, L. & Villamil, A. P. (1986), 'Market power in oral double auctions', *Economic Inquiry* **24**, 107.

Kahneman, D. & Tversky, A. (1991), 'Loss aversion in riskless choice: A reference-dependent model.', *The Quarterly Journal of Economics* **106**(4), 1039–1061.

Kandori, M. (1992), 'Social norms and community enforcement', *Review of Economic Studies* **59**(1), 61–80.

Koszegi, B. & Rabin, M. (2006), 'A model of reference-dependent preferences.', *Quarterly Journal of Economics* **121**(4), 1133–65.

Koszegi, B. & Rabin, M. (2007), 'Reference-dependent risk attitudes.', *American Economic Review* **97**(4), 1047–1073.

Kremer, M. & Jayachandran, S. (2006), 'Odious debt.', *American Economic Review* **96**(1), 82–92.

Merryman, J. (1986), 'Two ways to of thinking about cultural property.', *American Journal of International Law* **80**(4), 831–853.

Niederle, M. & Vesterlund, L. (2007), 'Do women shy away from competition? do men compete too much?', *Quarterly Journal of Economics* **122**(3), 1067–1101.

North, D. (1990), *Institutions, institutional change and economic performance*, Cambridge University Press, New York, new York.

Pogge, T. (2001), 'Achieving democracy.', *Ethics and International Affairs* **15**(1), 3–23.

Prott, L. & O'Keefe, P. (1990), *Law and Cultural Heritage.*, Professional Books Limitied, England.

Rantakari, H. (2008), 'Governing adaptation', mimeo.

Ross, D. (1995), 'Disturbing history: Protecting mali's cultural heritage.', *African Arts* **28**(4), 1–10.

Roth, A., Prasnikar, V., Okuno-Fujiwara, M. & Zamir, S. (1991), 'Bargaining and market behavior in jerusalem, ljubljana, pittsburgh, and tokyo: An experimental study.', *American Economic Review* **81**(3), 1068–1995.

Rothschild, M. & Stiglitz, J. (1976), 'Equilibrium in competitive insurance markets: an essay on the economics of imperfect information', *Quarterly Journal of Economics* **90**(2), 629–650.

Shurchkov, O. (2008), 'Performance in competitive environments: Are women really dierent?', mimeo.

Shyloon, F. (2000), 'The recovery of cultural objects by african states through the unesco and unidroit conventions and the role of arbitration.', *Uniform Law Review* pp. 219–241.

Smith, L. & Sorensen, P. (2000), 'Pathological outcomes of observational learning', *Econometrica* **68**(2), 371–398.

Tirole, J. (1996), 'A thory of collective reputation (with applications to the persistence of corruption and firm quality)', *The Review of Economic Studies* **63**(1), 1–22.

Van Huyck, Battalio, R. (1997), 'On the origin of convention: Evidence from coordination games', *The Economic Journal* **107**(442), 576–596.

Van Huyck, Battalio, S. (2001), 'Optimization incentives and coordination failure in laboratory stag hunt games', *Econometrica* **69**(3), 749–764.

Walker, J. & Williams, A. (1988), 'Market behavior in bid, offer, and double auctions: A reexaminiation', *Journal of Economi Behavior & Organization* **9**, 301–314.

Welch (1992), 'Sequential sales, learning and cascades', *The Journal of Finance* **47**(2), 695–732.

Wendel, P. T. (2007), 'Protecting newly discovered antiquities: Thinking outside the "fee simple" box.', *Fordham Law Review* **76**, 1016–1062.

Young, P. (1993), 'The evolution of conventions', *Econometrica* **61**(1), 57–84.