Essays in Behavioral Strategy

Matthew Heyser Stephenson

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy under the Executive Committee of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2022
Abstract

Essays in Behavioral Strategy

Matthew Heyser Stephenson

The following essays explore ways in which the environment affects and is affected by organizations. The first essay, “Trust and the Division of Labor” considers that the trust environment of a firm helps determine its structure. Jointly with Stephan Meier and Patryk Perkowski, I show that exogenously imposed culture leads to variation in organizational form. An experiment primes trust using past performance from a pilot study and demonstrate that the level of trust within an organization affects division of labor and consequently organizational productivity. This evidence is consistent with a cross-country link between trust and the division of labor that we observe in data from the European Social Survey. A simple evolutionary game theoretic model is provided to illustrate the results. The second essay, Nobody Likes a Rat”, considers the impact of norms against certain types of behavior (in this case dishonesty) on behavior and organizational composition. Jointly with Ernesto Reuben, I investigate the intrinsic motivation of individuals to report, and thereby sanction, fellow group members who lie for personal gain. We find that when groups can select their members, individuals who report lies are generally shunned, even by groups where lying is absent. This facilitates the formation of dishonest groups where lying is prevalent and reporting is nonexistent. Finally, "NFTs, Volume, and Social Influence" observes how organizations and individuals use environmental cues like rankings and volume for sensemaking in a market with high quality uncertainty. Using observational data scraped from the top 1000 NFT collections I find a significant positive
relationship between volume and price. Then, using plausibly exogenous variation in blockchain-level transaction fees, I fit an instrumental variable model which helps validate the causal interpretation that changes in volume lead to changes in price. I further add an experiment on NFTs to tease out two plausible channels through which volume could affect prices: user attention and normative social influence. The experiment finds strong evidence that being told an NFT is higher volume leads subjects to pay more attention to that NFT, whereas this has no significant effect on a subject’s reported preference for the NFT. A second experimental treatment, in which subjects were told the NFTs were ordered by volume transaction costs, does observe a significant positive affect on reported preferences (as well as attention).
## Table of Contents

Acknowledgments ................................................................. vi

Dedication ........................................................................ vii

Introduction or Preface ............................................................ 1

Chapter 1: Cultures of Trust and the Division of Labor ................... 4
  1.1 Introduction ................................................................. 4
  1.2 Conceptual Considerations .............................................. 8
  1.3 Illustrative Evidence on Trust and Division of Labor across Countries 14
  1.4 Experimental Design .................................................... 20
    1.4.1 Task ................................................................. 20
    1.4.2 Trust Manipulation ............................................... 22
    1.4.3 Procedure .......................................................... 23
  1.5 Experimental Results .................................................... 25
  1.6 Conclusion ................................................................. 29

Chapter 2: Nobody likes a rat: On the willingness to report lies and the consequences thereof 31
  2.1 Introduction ................................................................. 31
  2.2 Experimental design and procedures .................................. 34
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2.1 The whistleblowing game</td>
<td>34</td>
</tr>
<tr>
<td>2.2.2 Theoretical Predictions</td>
<td>36</td>
</tr>
<tr>
<td>2.2.3 Experimental Procedures</td>
<td>38</td>
</tr>
<tr>
<td>2.3 Results</td>
<td>38</td>
</tr>
<tr>
<td>2.3.1 Overstating</td>
<td>39</td>
</tr>
<tr>
<td>2.3.2 Reporting Behavior</td>
<td>43</td>
</tr>
<tr>
<td>2.3.3 Organization member selection</td>
<td>48</td>
</tr>
<tr>
<td>2.3.4 Final payoffs</td>
<td>51</td>
</tr>
<tr>
<td>2.4 Conclusions</td>
<td>53</td>
</tr>
<tr>
<td>Chapter 3: NFTs, Volume, and Social Influence</td>
<td>55</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>55</td>
</tr>
<tr>
<td>3.1.1 Volume as information and influence</td>
<td>57</td>
</tr>
<tr>
<td>3.2 An NFT-focused overview of the Ethereum Blockchain</td>
<td>58</td>
</tr>
<tr>
<td>3.2.1 The Ethereum Blockchain</td>
<td>58</td>
</tr>
<tr>
<td>3.2.2 NFTs as Digital Objects</td>
<td>60</td>
</tr>
<tr>
<td>3.2.3 NFTs as Property</td>
<td>61</td>
</tr>
<tr>
<td>3.2.4 The NFT Market</td>
<td>63</td>
</tr>
<tr>
<td>3.3 The causal effect of volume on price: evidence from NFT Data</td>
<td>64</td>
</tr>
<tr>
<td>3.3.1 Identification</td>
<td>64</td>
</tr>
<tr>
<td>3.4 Experiment on NFTs and Social Influence</td>
<td>66</td>
</tr>
<tr>
<td>3.4.1 Explanation of Experimental Conditions and Methods</td>
<td>67</td>
</tr>
<tr>
<td>3.4.2 Experimental Results</td>
<td>69</td>
</tr>
</tbody>
</table>
List of Figures

1.1 Properties of the Specialization Equilibrium ........................................ 13
1.2 Trust and Division of Labor ................................................................. 17
1.3 * ........................................................................................................ 17
1.4 Clicking Task ...................................................................................... 21
1.5 Switching Screen .................................................................................. 22
1.6 Low Trust ............................................................................................ 24
1.7 High Trust ............................................................................................ 24
1.8 Trust Manipulation ............................................................................... 24
1.9 A CDF Representation of Specialist Behavior by Condition ................. 26
1.10 Division of Labor over Time ............................................................... 27
1.11 Effort across Periods and Treatments ................................................. 28

3.1 A screenshot of opensea.com/rankings (the default ranking is volume.) .... 56
3.2 Changes in the price of computation on Ethereum, 7/20-7/21 .................. 60
3.3 Changes in the supply of computation on Ethereum, 7/20-7/21 ............... 60
3.4 The visual representation of the NFT "Autoglyph 343". The Contract, TokenID pair {0xd4e4078ca3495DE5B1d4dB434BEbc5a986197782, 343} is provably unique on Ethereum and with on-chain metadata containing instructions to reproduce this image. ......................................................... 63
3.5 A screenshot of question 1, featuring Fidenza #560 by Tyler Hobbs .......... 68
List of Tables

1.1 Trust and the Division of Labor in the ESS Dataset ................. 19
1.2 .................. ............................................. 19
1.3 The Causal Impact of Trust on Division of Labor .................. 28
1.4 .................. ............................................. 28
3.1 The Effect of Volume on Price in NFT Markets ....................... 65
3.2 Results from Treatment 1 (Volume) ............................. 70
3.3 Results from Treatment 2 (Credible Volume) ....................... 71
Acknowledgements

Thank you to my committee – Bruce Kogut, Stephan Meier, Dan Wang, RA Farrokhnia, Jorge Guzman – who made this work inestimably better. The PhD office which helped at every step: Elizabeth Elam, Dan Spacher. And the many other teachers who’ve taught me so much Vernon Smith, Jerry Kim, Casey Ichniowski, Damon Phillips, Eric Abrahamson, Andrea Prat, Michael Zargham. My parents Ed and Kat, my sister Ashlee. And of course my incredible (and incredibly patient!) wife Stephanie.
Dedication

To Hewitt and Stephanie. "Ask him now if he is glad he went and 'yes' he will tell you. 'Yes.'"
Preface

Firms are affected by the environments in which they operate. Famously, the existence of a firm as a form of production is thought to be explained by the external factor of “marketing costs”—costs of using the market mechanism which are distinct from the costs of coordinating activity within a firm. Beyond their presumed role in determining the boundary of the firm, these essays consider the ways that environmental costs can affect the structure of firms and of markets themselves.

The first essay, “Trust and the Division of Labor”, finds that a culture of trust leads to the formation of more specialized and efficient teams. In low trust environments, the cost of division of labor is high as agents spend more time duplicating efforts to ensure some minimal outcome. This essay establishes the plausibility and relevance of the relationship between reported trust and the division of labor via observational data from the European Social Survey. A plausible instrument is used to help determine causality, and a causal of trust on division of labor is strengthened further via a laboratory experiment. Lastly, a simple Evolutionary Game Theoretic model is presented which observes that under high trust a Division of Labor is the unique Bayesian Nash equilibrium, whereas as trust declines multiple equilibria are observed which include specialist as well as generality team structures.

If a specialized organizational structure requires trust then low trust environments can effectively impose a cost on those structures. However, it doesn’t mean that organizations cannot specialize. For instance, a firm may forego hiring the most talented prospects in favor of trustworthy prospects so it can achieve specialize. This may explain the prevalence of family firms in low
trust environments, for instance, with these firms restricting their available talent pool to kin but presumably maintaining a level of trust. Such a phenomenon suggests that organizational form is also affected by the ability of organizations to select their members. In my second essay, "Nobody Likes a Rat", I focus on an organizational membership and selection in a context of external regulation. "Nobody Likes a Rat" considers the consequences of "Whistleblowing" behavior for group membership selection. The experimental setting is such that, within a group, individuals have a selfish incentive to be dishonest since dishonest behavior can pay more. But group-members are also able to "whistleblow" by reporting the dishonest behavior of others in their group, with a report resulting in a sanction of the dishonest member. I then create a sort of "job market" in which one group member is randomly removed from a group and another member is randomly chosen to join, conditional on a vote of the remaining members. Before voting on the "applicant" group members are able to observe their past performance (and thus probable dishonesty) as well as previous decisions to report others. I find that members who have reported in the past are "shunned"; they are more likely to be rejected by groups, even if those groups are honest. The experimental settings of "Nobody Likes a Rat" intentionally mirrors a regulatory environment in which potentially lucrative misbehavior can be punished if it is revealed. Of course, misbehavior can also affect the "arm’s length" transactions that take place outside the firm (Arrow 1970). One such example is Wash Trading, which is a "self-financed sale" which manipulates market volume.\(^1\) While Wash Trading doesn’t directly affect price, it “may influence a trader’s sentiment about a given virtual asset” and thus have a secondary order effect. In “NFTs, Volume, and Social Influence”, I test the theory volume influences sentiment along the dimensions of social influence and attention.

Popular NFT charts and analytics seem to prize volume highly—for instance, “trading volume” is the default “NFT Ranking” on OpenSea which is the largest NFT marketplace. With Wash

\(^1\)Wash Trading was made illegal by the Commodity Exchange Act of 1936 and it is defined recently by the CFTC as “entering into... transactions to give the appearance that purchases and sales have been made, without incurring market risk or changing the trader’s market position.” The “essential characteristic” of a wash trade is the “intent”, such that if one “believes they are bargaining faithfully” then they are not wash trading, whatever the result. Thus, a wash trade is a form of dishonesty—unless once is trying “to give the appearance” of a sale, then one has not violated the rule.
Trading in a regulatory grey area, and Volume seemingly prized as an important metric, NFT marketplaces are a motivating case and a useful touchpoint for understanding the behavioral impacts of Volume Manipulation.

"NFTs, Volume, and Social Influence" features an experimental study testing (1) the causal effect of volume information as a channel of social influence and (2) the causal effect of volume information on attention. The experiment is supplemented by observational data which finds that plausibly exogenous negative shocks to NFT volume result in a lower observed NFT price.
Chapter 1: Cultures of Trust and the Division of Labor

1.1 Introduction

Organizations differ in observed performance but they also differ internally; for instance, in some organizations the division of labor is extensive while in other it is not. And internal differences between firms can help account for their performance differences, as when differing management practices within an organization affect firm productivity. While it has been known since Adam Smith’s time that an organization’s internal structure, e.g. the degree of division of labor, may profoundly affect its productivity, the process by which such structure evolves and is sustained is less understood. Part of this internal form may be determined by the production characteristics of its industry, but there exist substantial differences in organizational structures even within an industry [1, 2]. Explaining differences in firms’ organizational structures is an important aspect of explaining productivity differences.\(^1\)

This paper investigates whether some of the difference in organizational structure can be traced back to corporate culture. In particular, we argue and show empirically that changing an organization’s level of trust affects its division of labor and thus that exogenously imposed differences in culture endogenously lead to different organizational forms. Existing literature suggests that corporate culture might play an important part in explaining firms’ performance [12] and that trust is crucial for cooperation within organizations [13, 11]. We want to extend those two strands of literature by arguing that the trust dimension of corporate culture affects performance through organizational structure, in particular the degree of division of labor. Theoretically, a link between trust (culture) and the division of labor (organizational structure) is evident when considering that

\(^1\)The organization of a firm has been argued to depend on many factors, including strategy [3], technology [4], and environment [5, 6]. Additionally, factors such as identity [7, 8] social comparison [9], and reciprocity [10, 11] can affect the organization and function of firms.
the division of labor is limited by coordination costs, one of which is “whether workers trust each other” [14, p. 303].

To provide some intuition on how trust affects the division of labor, consider the example of a group of co-authors specializing in different sections of a paper. This can be efficient because 1) the person doing the literature review and write-up does not need to spend time decoding the proofs, re-focusing, and/or physically going from the field to the computer lab; and 2) that same writer has familiarized herself with relevant literature, major papers, and authors to an extent beyond that of her co-authors. Meanwhile, her co-authors have done the same *mutatis mutandis*. This efficiency has come at the expense of, for instance, the writer’s ability to understand and improve the proofs in the formal model that another co-author has been developing. Division of labor thus imposes an evident cost, and in a case where the value of the product is effectively the minimum effort in each subtask—so that a paper with a thorough literature review but a non-functioning model is not publishable—the cost can be great.

In light of such a cost we think the existence and sustainability of division of labor depend on whether the workers trust one another. In the absence of trust, they can lose efficiency switching between tasks, thereby “de-specializing” the group. An evocative example of this phenomenon set in a firm comes from management guru Stephen Covey, who describes a circumstance in which a division he managed depended on another division to help meet a customer’s needs. Because Covey believed that this other division had a bad reputation, he opted to take the “easy but expensive way out” by doing everything within his division, “creat(ing) our own redundant systems”. He added that “the whole organization was taxed for it in terms of the time and effort we had to put into something that should have been done by somebody else” [17, p. 264]. A more subtle example of this phenomenon comes from team sports: Dean Oliver, analyst for the NBA’s Denver Nuggets, notes that zone defense relies on “trust” that teammates will “cover their own zones” [18]. As in Covey’s example, a coach may have specified high division of labor—with each player

\[\text{\footnotesize{2These efficiencies are adapted from [15], which we will return to in Section 2.}}\]

\[\text{\footnotesize{3That is to say that trust affects task allocation. We focus here on “endogenous” task allocation by measuring actual effort provision in tasks, though delegation-based task allocation is also very important and plausibly affected by trust. For a review of these foundational concepts, see [16]}}\]
defending a specialized zone—but players and employees can alter their tasks to de-specialize the group.

The question of the impact of trust on the division of labor is difficult to study empirically since, first, data on trust levels within a large number of firms and their degrees of division of labor is difficult to obtain. Second and more importantly, a correlation between trust and division of labor could be due to omitted variables, e.g. the management team puts in place structure or institutions that affect both trust levels and division of labor. Reverse causality, i.e. the organizational structure affecting trust levels, makes it almost impossible to interpret a correlation between trust and division of labor as causal. Due to this problem, we provide two types of complementary evidence for the link between trust and division of labor. The main evidence comes from a laboratory experiment that establishes a causal effect between trust and division of labor. We also use cross-country evidence to establish a correlation between trust and division of labor.

In the laboratory experiment, individuals engage in a productive enterprise in groups of three and they are given the ability to alter how they divide labor among specialist and generalist task allocations. Performance is compensated at the group level, an arrangement “far more common than individual performance-based pay” within firms [23, p. 1196]. For simplicity we also assume a minimum-effort production process in which the group performance is the result of the least effort among the three tasks. Individuals can all specialize on their own task and if they do they will be maximally productive as a group. However, individuals are also given the ability to engage in a non-specialized task, which comes at a productivity cost but also confers the advantage of ensuring that non-specialized task is completed to some standard. This game effectively models the real world division of labor situations we describe above, in which members of a team can choose to “de-specialize” at a cost. Before the individuals engage in the real effort task we randomly manipulate their level of trust, allowing us to study whether exogenously changing a culture of trust endogenously creates different levels of division of labor.

Our laboratory findings show a significant effect of exogenously imposed trust on specialist

\[<sup>4</sup>\]For experimental methods used in answering questions in organizational studies and strategy, see [19], [20], [21], [22].
behavior at the intensive and extensive margins. Specifically, we find positive and significant effects of trust on the level of realized specialization within a group, as well as the total performance of the group. Exogenously changing the trust level across our organizations, i.e. our groups of three, leads to the emergence of different forms of working together, i.e. extent of division of labor. We find that even exogenously imposed trust levels can minimize uncertainty about other co-workers. Furthermore, by repeating this experiment with feedback, we observe that these effects intensify over time, with high-trust groups increasing in specialization over time.

Suggestive cross-country evidence, using data from the European Social Survey, supports this finding in that the level of generalized trust within a country corresponds to a higher level of specialization within that country’s industries. The correlation between our proxy of division of labor—a measure of specialized job descriptions within a country’s industries—and generalized trust is robust to economic controls and country and time fixed effects. We also instrument for trust with consistent and significant results. While the evidence cannot be interpreted as causal, it supports the experimental evidence by showing that trust level is correlated to the division of labor in observational data. Our study makes at least three important contributions to the literature:

First, our paper contributes to the debate about the effect of culture on organizational structure. A recent paper by [24] finds that trust is associated with larger firm size and a flattening of hierarchical structures within the firm. We test a distinct but complementary theory: that trust affects the task division within firms, with higher-trust environments endogenously producing more specialized task allocations. The trust environment within a firm is an important feature of a company’s culture [25], and we find that this can help a firm to facilitate a stable division of labor. We show that initial differences in trust level can endogenously lead to different degrees of specialization within an organization. These initial differences are important for three reasons: 1) Firm structure is inertial [26], so the early structure of an organization remains important even as a firm grows in size, 2) A firm’s culture is formed early and is difficult to change [27], and 3) Trust only increases by a small amount as a function of relationship duration, so the existence of relationships among individuals within a firm does not by itself impact trust substantially over time [28].
Second and related to the first contribution, our paper can shed light on a mechanism explaining the relationship between trust and growth. Trust levels within a country have been thought to affect a number of economic variables, most notably growth. A number of studies show a significant correlation between trust and economic growth (e.g., [29] and [30]). However, the precise mechanism through which trust affects growth is still an open question. Our evidence shows that a potential mechanism for the relationship between trust and growth is the organization of firms. If trust affects the division of labor as we observe in our study, and the division of labor affects the “wealth of nations”, this illuminates a plausible mechanism that links trust to growth: organizational structure.

Third, we contribute to the literature on the relationship between trust and economic behavior. There exists significant (mostly experimental) work on the ways in which institutions and organizational forms causally foster the trusting and trustworthy behavior of individuals ([31]; [32]; [33]; [34]). Our contribution lies in studying the converse of this research: addressing the manner in which individual’s trust-levels causally impact institutions and organizational forms. The only other paper that we are aware of that studies the effect of trust on economic behavior is [35]. This paper applies a very similar manipulation of high and low trust to the one we use, but investigates the role of trust and trustworthiness in inducing high and low effort equilibria in a gift exchange game, whereas we study the role of trust in the division of labor.

The paper proceeds as follows: Section 2 discussed concepts used within the paper and includes a model of the effects of trust on specialization. Section 3 presents illustrative evidence from cross-country data on the relationship between generalized trust and division of labor. Section 4 introduces the experimental design and the results of the experiment are exhibited in Section 5. Section 6 provides concluding thoughts and suggestions for further research.

1.2 Conceptual Considerations

The division of labor is thought to be paramount in accounting for the productivity of organizations going back to Adam Smith. However, the process by which the division of labor emerges
has received, to the best of our knowledge, limited prior attention in the scholarly literature.⁵ [39, forthcoming] constitutes a rare recent exception, focusing on factors that precipitate and sustain the division of labor.

We present a simple game theoretic model that illustrates the manner in which trust affects the division of labor. We find that, in the presence of gains from specialization, trust makes a division of labor more likely. Increasing trust also increases the payoff from specialization, as well the risk profile of choosing to specialize vs. generalize.

Gains to Specialization

Let there exist some set of tasks \( y \subseteq \{y_1, \ldots, y_n\} \) divided among \( N \) workers. \( R_i \) is a switching variable denoting the number of tasks a worker engages in, \( R_i \in [1,N] \). Thus if there are two workers who work on all tasks in \( y \), then \( R_i=N=2 \) whereas if they divide the tasks with no overlap \( R_i=1 \).

Effort for a worker \( i \) in task \( k \) is \( e_{ik} \) and their production in task \( k \) is as follows:

\[
y_{ik} = \frac{e_{ik}}{R_i^\alpha}
\]

Where \( \alpha \) is a constant that determines the productivity effects of specialization. Smith’s classical treatment of gains from specialization identified distinct efficiencies it brings about: saving on switching costs and the increase in skill that comes from repeating a task (1887).⁶ [40] refers to this as “learning by doing”. Recognizing these efficiencies, we focus on the “gains from specialization” case in which \( \alpha > 1 \).⁷ In such a case, with fixed effort and two or more workers, individual productive output is decreasing in the number of tasks worked on, reflecting decreased switching costs and/or learning by doing. Under fully divided labor with no task overlap, overall productivity for worker \( i \) is \( \sum_k y_{ik} = e_i \) whereas full generalization gives \( \sum_k y_{ik} = e_i R_i^{1-\alpha} < e_i, \forall R > 1 \). Note

⁵Vernon Smith and coauthors also have several experimental papers that importantly investigate aspects of the division of labor. They explore the role of specialization in the formation of long-distance trade under different institutional environments [36], the discovery and emergence of specialization [37], and the emergence of property rights due to specialization and gains from exchange [38].

⁶Smith identifies a third efficiency—the creation of productive tools—but this is outside the scope of our paper.

⁷Observe that with \( \alpha = 1 \) there would be no productivity losses from switching tasks.
that $\alpha$ is not specialization but rather the productive benefits that accrue to specialization, benefits that vary by the productive technology available in a given industry for instance.

It is useful here to restrict attention to a simple case of two workers and two tasks. We follow [14]’s paper on the division of labor and specify a “minimum effort” production technology. Output $Y$ is thus equal to the minimum production in either of the two tasks: $Y = \min((y_{11}+y_{21}), (y_{12}+y_{22}))$. That is, productivity is equal to either the sum of worker 1 and worker 2’s production in task $y_1$ or the sum of worker 1 and worker 2’s production in task $y_2$, whichever is less. Payoffs are then split equally between the workers, and effort and task-allocation are non-contractible, so that $\pi_i = \frac{Y(y_{ik})}{N}$.

Trust

Having introduced a set of tasks with production benefits from greater division of labor, we will now consider trust. Trust is a notoriously hard concept to define, but within the existing definitions there appears to be a convergence on the role of belief as an important element. [41] summarizes an inter-disciplinary consensus as follows: trust is the “subjective probability with which an agent assesses that another agent or group of agents will perform a particular action, both before he can monitor such action and in a context in which it affects his own action.” [42] propose that, in the productive context, a trusting belief is one that regards the other’s “cost or pleasure of accomplishing the task” (p. 494). Following this, we define trust as a probabilistic belief about another agent’s incentive problem.

With respect to the productive tasks defined above, suppose that there exist two types of workers who differ only in effort provision for these tasks, type $\theta_i = \tilde{\theta}$ sets $e_i=1$ while $\theta_i = \overline{\theta}$ sets $e_i=0$ in all cases. Effort is thus “all or none” and these types, respectively, are workers who find it profitable to exert effort given the potential payoffs or not at all (we will treat this assumption explicitly later). Workers have a trust parameter, a common prior $\theta=(0,1)$ about the probability of a worker being type $\tilde{\theta}$. We can think of this parameter reflecting the “social capital” of their country or

---

[8] The “trust effect” in their model is about perceived incentive problems, which contrasts with a “profitability effect” regarding worker ability. We do not treat the profitability effect explicitly here, and in our experimental treatment heterogenous ability is quite unlikely.
environment. They also know their own type \( \theta_i \) with certainty. The risk-neutral worker then sets a level of specialization based on the level of trust and their type, choosing \( \hat{R}_i(\theta, \theta_i) \) to maximize their expected payoffs in a simultaneous move game between workers.

Suppose that worker 1 is of type \( \bar{\theta} \). To see what values of \( \theta \) produce a specializing (\( R=1 \)) equilibrium, we check for profitable deviations by fixing worker 2’s decision as being full specialization (\( \hat{R}_2=1 \)) with \( y_{22}=\theta \) and \( y_{21}=0 \). This produces payoffs as follows:

\[
\pi_1(\hat{R}_1) = \frac{Y(\hat{R}_1)}{N} = \min \left( (y_{11} | \hat{R}_1), (y_{12} | \hat{R}_1) + \theta \right) N^{-1}
\]

Because worker 1, as type \( \bar{\theta} \), sets \( e=1 \) and worker two sets \( e=1 \) with probability \( \theta \), we can omit the subscripts on \( e \). \( \hat{R} \) is a discrete choice variable that can equal 1 or 2 in the two-person case, so we have:

\[
\pi_1(\hat{R}_1) = \begin{cases} 
\min \left( e, \theta e \right) N^{-1} = \frac{\theta e}{N} & \text{if } \hat{R}_1 = 1 \\
\min \left( \frac{e}{2\sigma}, \frac{\theta e}{2\sigma} + \theta e \right) N^{-1} = \frac{\theta e}{2\sigma N} & \text{if } \hat{R}_1 = 2
\end{cases}
\]

Worker 1 will not find it profitable to deviate from \( \hat{R}=1 \) if \( \theta e \geq \frac{\varepsilon}{2\sigma} \). By symmetry, this is also the case for worker 2 conditional on being \( \bar{\theta} \). Thus we can see then that the existence of equilibrium specialization benefits from increasing both gains from specialization (\( \alpha \)) and trust (\( \theta \)). Treating \( \alpha \) as fixed by a technology, at a sufficient level of trust specializing constitutes a Bayesian Nash Equilibrium.

Under what conditions is generalizing an equilibrium? Now suppose that worker 2 fully generalizes (\( \hat{R}_2 = 2 \), \( y_{21} = y_{22} = \frac{\theta e}{2\sigma} \)) and we thus check worker 1’s incentive to deviate:

\[
\pi_1(\hat{R}_1) = \begin{cases} 
\min \left( e + \frac{\theta e}{2\sigma}, \frac{\theta e}{2\sigma} \right) N^{-1} = \frac{\theta e}{2\sigma N} & \text{if } \hat{R}_1 = 1 \\
\min \left( \frac{e(1+\theta)}{2\sigma}, \frac{e(1+\theta)}{2\sigma} \right) N^{-1} = \frac{e(1+\theta)}{2\sigma N} & \text{if } \hat{R}_1 = 2
\end{cases}
\]

\(^9\text{Because the actions of each worker in the game are independent, a worker of type } \bar{\theta} \text{ is indifferent in their choice of } \hat{R} \text{ since they set } e=0 \text{ in all cases. Thus their } \hat{R} \text{ does not affect payoffs and they never have incentive to deviate.} \)
Generalizing is always an equilibrium since \( e(1+\theta) \geq \theta e \rightarrow e \geq 0 \) and \( e_i \geq 0 \ \forall \ i \). If trust is low \( (\theta e \leq \frac{e}{2}) \) and \( \theta_i = \hat{\theta} \) then generalizing constitutes the unique equilibrium.\(^{10}\) With the level of trust greater than or equal to \( \frac{e}{2\pi} \) we have multiple equilibria, one in which both specialize and one in which both generalize.

Trust and organizational form

At a level of trust such that two organizational forms are feasible \( (\theta e \geq \frac{e}{2\pi}) \), we can say more about the equilibrium properties of these forms. Specializing will payoff dominate generalizing only if \( \theta e > \frac{\theta(1+\theta)}{2\alpha} \rightarrow \theta(2\alpha - 1) > 1 \). The relative payoff of the specialized form of organization is greater with more gains the specialization \( \alpha \) and trust \( \theta \).

To clarify these relationships, Figure 1 displays a graph of the existence and dominance properties of the specialization \( (R=1) \) equilibrium. Any combination of values for trust and gains to specialization that lies above the solid line creates a specializing equilibrium in the game. Specializing is payoff dominant for values above the dashed line, and is risk dominant\(^{11}\) for values above the dotted line. To observe the effect of trust on division of labor, it is useful to pick a value for \( \alpha \), supposing a fixed production technology. With \( \alpha = 4/3 \), an individual’s productivity gains from specialization are \( \approx 25 \) percent for a constant level of effort.\(^{12}\) At this value of \( \alpha \), specialization is only an equilibrium if trust \( \approx 0.4 \) or greater. Specialization is only payoff dominant with trust greater than \( \approx 0.7 \), and generalizing always risk dominates regardless of the level of trust.

\[43\] observes that the belief component of trust can have lasting effects “only as an equilibrium selection device”. In our model, as trust increases, we observe that division of labor becomes a Bayesian Nash Equilibrium and thus a viable organizational form. This is because a worker’s expectations that the other worker will remain on task increases to the point where they no longer find it preferable to safeguard against low-effort by deviating to the off-equilibrium payoff. How-

\(^{10}\) For \( \theta_i = \hat{\theta} \), all choices of \( \hat{R} \) produce Nash Equilibria, since they exert no effort and don’t affect payoffs. This is a less interesting case.

\(^{11}\) It is risk dominant if it provides a higher payoff on the assumption that the opponent completely randomizes. In this case, that means \( \theta > \frac{2}{2\pi} \).

\(^{12}\) Recall from earlier that overall productivity under generalization is \( e_i R_i^{1-\alpha} \) while productivity under specialization is \( e_i \).
ever, workers might still receive higher payoffs in expectation by fully generalizing. But with trust increasing to a high level, a specialized equilibrium can produce the greatest expected payoff. Even where specializing produces a higher payoff there are still often advantages to the generalist equilibrium though; it can be risk dominant.

Errata and extensions

Our assumption about constant effort from types is fairly strong since we are implicitly assuming that an agent’s incentive problem is either satisfied or not satisfied at all regardless of organizational form. We will show that our conclusions are not impacted substantially when relaxing the assumption. Suppose $\exists \hat{\theta}=(0,1)$, with $\theta+\hat{\theta}+(1-\theta-\hat{\theta})=1$. $\hat{\theta}$ is the portion of the population for which the participation constraint is only met under full specialization and not under generalization. If no specializing equilibrium exists then this type is indistinguishable from $\theta$ as they simply set effort to zero in all cases. This type also has no incentive to unilaterally deviate from generalization so that equilibrium still exists. However, where specialization is an equilibrium, including this type does allow specialization to payoff dominate at a lower level of trust, since $(\theta+\hat{\theta})e>\theta e$.

We also have ignored the task-coordination problem inherent in the specializing equilibrium. That is, we have assumed that if both workers specialize then they specialize in different tasks. Note that the generalizing equilibrium has no such task-coordination problem. Assuming that no task coordinating device exists then the conditions on the existence of the specializing equilibrium
become more stringent, $\frac{\theta_{\omega}}{2} \geq \frac{\epsilon_{\omega}}{2\sigma}$, and trust and/or gains from specialization must be higher than in the baseline to ensure the equilibrium. The same is true for the conditions on payoff dominance and risk dominance.

1.3 Illustrative Evidence on Trust and Division of Labor across Countries

In order to motivate the relevance of our idea and experimental results outside the lab, we present cross-country evidence from a repeated panel survey that sheds light on the relation between trust levels and division of labor. A culture of trust in a country can affect a firm’s structure for at least two reasons: First, the level of trust within an organization can act as a constraint on its form, as shown in the model. Thus a firm might see greater potential production under high levels of specialization but is constrained by the trust environment that makes such an arrangement unstable. Hence we have a within-country hypothesis suggesting that the division of labor will respond to changes in trust. The model suggests that this effect should be observed at a given specialization technology, and thus industry-level observations are relevant. Secondly, the nascent structure of a firm is often team-production [44] and, because firm structure is inertial and resistant to change [26], trust’s effects on division of labor could be expected to persist. Thus we would predict a cross-sectional correlation between trust and the division of labor. While we find support for both hypotheses, controlling for time-invariant factors on the country level and instrumenting for trust, but we consider the evidence illustrative. We therefore couple these findings with our experimental evidence investigating the causal effect of trust on division of labor. Nevertheless, the evidence provided in this section resembles evidence on trust and growth [29, 30] and we see it as a complementary illustration to our experimental evidence.

Data sets: We draw on data from the European Social Survey (ESS) for the available years of 2002-2012. The range provides 217,250 individual observations drawn from 35 countries.13 We

13Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Kosovo, Latvia, Lithuania, Luxembourg, Netherlands, Norway, Poland, Portugal, Romania, Russian Federation, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, and the United Kingdom.
then pair this data with country-level economic measures from the Penn World Tables [45].

**Proxy for division of labor:** We exploit a unique feature of the data in the European Social Survey (ESS) to construct a measure of division of labor. The ESS contains classifications for both an individual’s occupation as well as the industry in which they are employed, using the International Standard Classification of Occupations (ISCO-88) and Nomenclature of Economic Activity (NACE) codes respectively. Within our sample we have 492 unique ISCO-classified occupations and 91 NACE-classified industries. Using this data we use an idea from [46] and construct an index of division of labor, \( d \), for each industry \( j \) within a country from this data as follows (for more details, see Appendix II):

\[
d_j = \frac{\sum_{i=1}^{n} x_{ij}^2}{(\sum_{i=1}^{n} x_{ij})^2}
\]

Where \( n \) is the number of possible occupations and \( x_i \) is the number of individuals in a given occupation \( i \). This form is maximized when \( x \) is constant for all \( i \), suggesting a uniform distribution of individuals among the possible occupations within an industry. If we suppose that each possible occupation within an industry is a “bucket” which can be filled by employees, \( d \) measures how levelly those buckets are filled within a country’s given industry. As described above, our division of labor index \( d \) is computed for each year within a domestic industry. That is, within a country, each industry has a unique division of labor score for a given year of the ESS.

To provide intuition on this measure, consider the forestry industry in Greece and Finland. These countries differ substantially in generalized trust, with Finland having a high average and Greece a low one. We also observe that Greece and Finland have different occupational structures within their forestry industries: a majority (61 percent) of Greek forestry workers classify their occupation as “manufacturing laborer”, a classification which denotes no obvious specialized skill. Compare this with Finland, in which there are no observed “manufacturing laborers” but instead there are “wood processing plant operators”, “wood products machine operators”, “wood treaters”, “woodworking machine setters”, “motorized forestry plant operators”, “lifting truck operators”
and so on. Moreover, these occupations are flatly distributed in the Finnish sample and no clear majority occupation is evident. As we would expect, these observations are reflected in the industry $d$ score, within Finland’s forestry industry being much higher than Greece’s. We maintain that the differentiation of skills and specific tasks form the basis on which a worker chooses to either classify themselves as a general laborer or a machine operator. Thus our measure $d$ captures an element of division of labor within industries.

**Trust measure:** The ESS includes a measure of trust. It is based on the question “Generally speaking would you say that most people can be trusted or that you need to be very careful in dealing with people”. Responders provide an answer on a 10-point scale with the lowest category being “You can’t be too careful” and the highest “Most people can be trusted”. The mean response to the generalized trust question in our sample is 4.83 with a standard deviation (s.d.) of 0.986, and a range of 2.7-6.9. [47] note that the “trust in strangers” survey question that we use is most relevant to beliefs within a laboratory, rather than preferences. [48] find that, when the “target of trust” is a stranger, surveyed measures of trust do correlate with trusting behavior in a controlled environment. Because our model and experiment treats an interaction among strangers in a productive context, these surveyed measures of trust (and production activities) are relevant.

**Results:** We evaluate the relationship between the division of labor and trust using the measures described above. Figure 2 displays the results for the aggregated $d$ scores plotted against generalized trust at the country level with a linear fit line projected on the data. Trust and $d$ are positively correlated ($p<.01$) and indicating that an increase of one s.d. in measured trust results in a .45 s.d. increase in the observed $d$-score.

In Table 1 we explore the robustness of the association between trust and division of labor by

---

14 This could, In principle, be explained by the hypothesis that trust increases the thoroughness of the ESS survey and results in more specific answers. However a spearman correlation finds no evidence for a link between interview time and trust ($p=0.769$).

15 So a worker charged with the task of treating the wood, driving the truck, and setting the machine for operation should be more apt to describe their occupation in general terms (“laborer”) than a worker whose entire job consists of treating wood.

16 These statistics are reported at the country level.

17 Results in this section are from an OLS regression unless otherwise specified.
Figure 1.2: Trust and Division of Labor

Figure 1.3: *

Note: Data taken from European Social Survey and measures averaged across countries from 2002-2012. Plot omits graphical outliers Cyprus (CY), Israel (IS), and Kosovo (XK). These countries are not omitted from the statistical analysis.
making use of the panel structure of our data and fitting several models with a variety of controls.\textsuperscript{18} The specification controls for growth as well as population, trade openness, and capital investment which are variables commonly used to explain growth. We also include controls plausibly related to division of labor: the number of industries in a country, number of individuals self-employed, and number of family firms. We also instrument for trust using the "feeling of safety after dark" (surveyed on a scale from 0-5), a measure that should only affect the division of labor through its effect on trust.

Column 1 reports results on a model which considers the smallest unit of analysis given our data: the yearly domestic industry. Note that trust is still measured at the year-country level, and so the estimates reflect changes at the domestic industry level as a response to changes in the country-wide trust environment. Column 2 reports results at the country level, pooling industries within countries and years, and this exhibits a similar coefficient on trust. In columns 3 and 4 report the same specifications but using controls. In both specifications we find a similarly significant and positive relationship between trust and the division of labor with roughly similar effect sizes. In columns 5 and 6 we add time and group level fixed effects and, while the industry-level specification remains significant, the country-level specification is only significant at the 10 percent level. Thus the result is largely robust to controlling for time-invariant effects. Column 7 displays the results of a two stage least squares regression, using "feeling of safety" as in instrument for trust (F<0.01). This instrument helps alleviate concerns about endogeniety, and the coefficient is significant and similar in size to the other results. Finally, column 8 reports the first stage of the IV regression.

In sum, we observe that trust is positively and significantly associated with the division of labor within the ESS dataset. This finding is not simply due to time-invariant country and country-industry factors, and a plausible instrument helps to rule out questions of reverse causality. Prominent research on trust and organizations treats the level of trust within a region as being primarily determined by historical particulars and thus "largely exogenous" with respect to the organization.

\textsuperscript{18}Appendix III exhibits a table with between-country models, including a Tobit regression.
Table 1.1: Trust and the Division of Labor in the ESS Dataset

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trust</td>
<td>0.0782***</td>
<td>0.0683**</td>
<td>0.108***</td>
<td>0.104***</td>
<td>0.179*</td>
<td>0.0713**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0125)</td>
<td>(0.0310)</td>
<td>(0.0162)</td>
<td>(0.0143)</td>
<td>(0.0983)</td>
<td>(0.0339)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self Employment</td>
<td>-0.116***</td>
<td>-0.290***</td>
<td>-0.0145</td>
<td>-0.0511</td>
<td>-0.113***</td>
<td>0.0904***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0313)</td>
<td>(0.0739)</td>
<td>(0.0241)</td>
<td>(0.165)</td>
<td>(0.0204)</td>
<td>(0.0217)</td>
<td></td>
<td></td>
</tr>
<tr>
<td># Family Firms</td>
<td>0.0263</td>
<td>-0.450</td>
<td>0.0855</td>
<td>-0.217</td>
<td>0.0214</td>
<td>-0.0992</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.118)</td>
<td>(0.443)</td>
<td>(0.0792)</td>
<td>(0.637)</td>
<td>(0.0721)</td>
<td>(0.0784)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trade Openness</td>
<td>0.0368**</td>
<td>0.0229</td>
<td>0.00537</td>
<td>0.106**</td>
<td>0.0317**</td>
<td>-0.0553***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0148)</td>
<td>(0.0178)</td>
<td>(0.0536)</td>
<td>(0.0485)</td>
<td>(0.0123)</td>
<td>(0.0129)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capital Investment</td>
<td>-0.0274</td>
<td>0.00120</td>
<td>0.0223</td>
<td>0.0861***</td>
<td>-0.0210</td>
<td>0.130***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0182)</td>
<td>(0.0244)</td>
<td>(0.0308)</td>
<td>(0.0274)</td>
<td>(0.0167)</td>
<td>(0.0169)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP per capita</td>
<td>0.00661**</td>
<td>0.00232</td>
<td>0.0759**</td>
<td>-0.151***</td>
<td>0.00778***</td>
<td>0.0213***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00333)</td>
<td>(0.00385)</td>
<td>(0.0378)</td>
<td>(0.0360)</td>
<td>(0.00258)</td>
<td>(0.00265)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population</td>
<td>0.0206***</td>
<td>0.0116**</td>
<td>0.0891</td>
<td>0.0705</td>
<td>0.0186***</td>
<td>-0.0340***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00488)</td>
<td>(0.00508)</td>
<td>(0.162)</td>
<td>(0.151)</td>
<td>(0.00410)</td>
<td>(0.00407)</td>
<td></td>
<td></td>
</tr>
<tr>
<td># Industries</td>
<td>0.00931</td>
<td>0.00761</td>
<td>0.00228</td>
<td>-0.00556</td>
<td>0.0113*</td>
<td>0.0659***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00610)</td>
<td>(0.00997)</td>
<td>(0.00705)</td>
<td>(0.00645)</td>
<td>(0.00622)</td>
<td>(0.00611)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feeling of Safety</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.228***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.00840)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.585***</td>
<td>0.724***</td>
<td>0.197</td>
<td>0.497***</td>
<td>0.530</td>
<td>0.947</td>
<td>0.247**</td>
<td>0.355***</td>
</tr>
<tr>
<td></td>
<td>(0.0199)</td>
<td>(0.0515)</td>
<td>(0.134)</td>
<td>(0.176)</td>
<td>(1.378)</td>
<td>(1.302)</td>
<td>(0.116)</td>
<td>(0.122)</td>
</tr>
<tr>
<td>Fixed effects</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Observations</td>
<td>9,123</td>
<td>150</td>
<td>6,411</td>
<td>110</td>
<td>6,411</td>
<td>110</td>
<td>6,406</td>
<td>6,406</td>
</tr>
<tr>
<td>N</td>
<td>9123</td>
<td>150</td>
<td>6411</td>
<td>110</td>
<td>6411</td>
<td>110</td>
<td>6406</td>
<td>6406</td>
</tr>
</tbody>
</table>

Note: Data from the European Social Survey. Results from an OLS multiple regression, robust errors in parentheses, clustered at the domestic industry (1,3,5,7) and country (2,4,6) levels. All independent variables in logs. Column 8 reports the first stage of two-stage least squares regression instrumenting Trust with "Feeling of Safety". Fixed effects are time and group-level: domestic industry (5) and country (6). *** p<0.01, ** p<0.05, * p<0.1
of firms (Bloom et al. 2012a). Trust has been found to depend on weather conditions (Durante 2010), past literacy rates and institutions [49] and “crucial events in city-states during the medieval period and earlier” [50, cited in Bloom et al. 2012] Nevertheless, we are sensitive to the possibility of endogeneity given our sample and we prefer to present the results as correlational. Nevertheless we believe that the findings are suggestive and point to external validity for our experiment demonstrating a causal effect of trust on division of labor.

1.4 Experimental Design

To test whether the level of trust in an organization causally affects division of labor, we need (1) a task that measures the extent of specialization and (2) a way to randomly assign individuals/teams into high or low trust environments. We will describe the operationalization of our experiment in turn. Full instructions were available for participants to read and were read aloud prior to the experiment. These instructions are presented in Appendix I.

1.4.1 Task

We sought to devise an experimental game that could approximate the feeling of an actual work environment, leading us to use a real-effort task in continuous time. Our design attempts to capture the process of specialization in our model, treating distinct complementary tasks under “minimal effort” production. However, because the division of labor can be driven by heterogenous skills among a group, as well as beliefs about such skills, it was important to use a design in which individuals have similar levels of potential productivity while allowing for different levels of actual productivity within the game. That is to say, we did not want individuals switching to a different task simply because they thought they would be more skilled at it. To avoid these pitfalls, our design uses the same simple task for all subjects. The interdependence of the tasks comes from the payoff structure and we model specialization by imposing costs of switching between tasks (so \( \alpha > 1 \)).

The task is to clear blocks in a 20x20 grid (see Figure 3 for a screenshot) via simple clicking,
with a 9 second enforced delay between clicks. We expect that all students should have a similar ability in clicking. However, they may have a different utilities of effort as is considered in our model.

Holding the actual tasks constant, we then specialize the participants by dividing the field up into thirds, with each individual only able to work in one-third of the field at a time. Each individual has a specialized unique third of the field, and the game begins with them viewing this portion of the field. To engage in work a non-specialist portion of the field an individual must click “switch” and pay a switching cost equivalent to 18 seconds of effort. Upon clicking the switch button (observable in the bottom right corner of Figure 3), subjects are presented with a switching screen (see Figure 4 for a screenshot). In this screen they can observe the progress of made in the other subfields and choose to switch and then work in them.\textsuperscript{19}

Subjects in our experiment were also permitted to browse the internet during the experiment, acting as the outside option. And the imposed 9 second delay between clicks served to make the option more enticing. A browser window displaying Google.com was open on their desktop

\textsuperscript{19}Subjects must pay the switching cost before observing this screen.
at the time of the experiment and, informally, we observed about half of the participants using
the internet. In each of 2 experimental rounds, participants had 13 minutes time to work on their
subfield and/or switch to another subfield and work there. Thus if every individual fully specialized
the maximum production within a subfield would be 13 minutes times 60 seconds divided by 9
seconds waiting time per click, or 86. An individual in group j’s payoff is determined by

\[ \$5 + \frac{20 \times \min(\sum_i y_{ij1}) + 20 \times \min(\sum_i y_{ij2})}{3 \times 86} \]

This payoff includes a $5 show-up fee plus $20 times the minimum production y within a
subfield i in both round 1 and round 2. These round-specific payoffs are then split evenly among
the group (there are 3 people per group and thus they are divided by 3) as well as normalized by
the maximum possible production (divided by 86). Thus the maximum payoff within this game is
just over $18 for each individual.

After Period 1, i.e. the first 13 minutes, subjects were given information on the performance
of their group. Subjects observed their group production for the round—\( \min(\sum_i y_{ij1}) \)—and were
reminded of the maximum number of clicks (86) for context. They then repeated the task again in
period 2.

1.4.2 Trust Manipulation

In order to manipulate different trust environments, we use a method that is similar to the one
proposed by [35]. We prime the trust of individuals by presenting them with actual examples of
past performance in this task, taken from a pilot study. We present this information neutrally,
describing it as “an actual example of a subject’s performance in a game played previously”.\textsuperscript{20} While the text is the same for all subjects, the examples are different. The primes are included below in Figure 5: Panel A for Low Trust treatment and Panel B for High Trust treatment. Subjects were told that “The Red Line shows the total number of clicks possible” and “The Blue Bars show the actual (cumulative) number of clicks.”\textsuperscript{21}

Because subjects read the instructions and learned about the task beforehand, we expect that this prime affects their beliefs about others’ perception of the task. \cite{42} refer to this as the “trust effect” which is the belief about another’s “cost or pleasure of accomplishing the task” (p.494). They contrast this with a “profitability effect” which comes from receiving information about another’s ability. The latter is an unlikely channel in our experimental set-up, amounting to the belief that some students are incapable of clicking a mouse every nine seconds. To further ensure that we were manipulating trust, we asked subjects in a post-experimental survey to recall how trustworthy they believed their team was (on a scale of 1-7) “as they began the task” and “at the end of the task.”\textsuperscript{22} Among the subjects who responded we find a significant difference in how trustworthy they thought their team was at the end of the experiment, as well as how likely they were to experience a decline in trust. Subjects in the “Low Trust” condition are significantly more likely ($p < 0.05$)\textsuperscript{23} to report a lower “trustworthy” rating after the experiment than those in the “High Trust” condition. Subjects in “Low Trust” were also significantly more likely to exhibit an overall decrease in trust (“trust after” minus “trust before”), and the mean change in trust was negative in low trust groups (-0.285).

1.4.3 Procedure

We conducted the experiment at the Columbia Experimental Laboratory for the Social Sciences (CELSS), using 63 Columbia University undergraduates who were recruited via ORSEE \cite{51}.\textsuperscript{20} It was made clear to the participants beforehand that the experiment would involve no deception.\textsuperscript{21} The data was taken from a pilot experiment using the same game, but in which there was a longer delay imposed between clicks and thus the maximum was lower than 86.\textsuperscript{22} This survey was voluntary and there was some observed attrition (10 of the 63 subjects did not complete the survey) which reduces our number of observations somewhat.\textsuperscript{23} All reported $p$ values are from OLS regressions clustered at the group level unless otherwise noted.
Figure 1.6: Low Trust

Figure 1.7: High Trust

Figure 1.8: Trust Manipulation
There were 3 sessions lasting approximately 45 minutes apiece. A show up fee of $5 combined with the incentivized earnings produced an average payoff of $15 per subject. Participants were unaware of the nature of the experiment before entering the laboratory.

1.5 Experimental Results

In the following, we treat the number of task switches as examples of generalist behavior and we measure the effect of trust on generalist behavior on the intensive and extensive margins. We define a task switch as occurring when one subject switches to another’s field and then proceeds to click one or more times on this non-specialized field, switching back to the specialized task is costless and not counted as a task switch. Monitoring switches—in which one switches to another field but then (costlessly) switches back without working—were possible in this game but somewhat uncommon. Of the 131 acts of switching we observe, only 14 (10.6 percent) were monitoring switches. We test whether there is a significant effect of trust on a pure monitoring switch, and the results are insignificant (p=0.86) with a coefficient of .01. While the inclusion of these monitoring switches does not substantively impact our results (see Appendix V), we exclude them for our main analysis and focus on task switching, in which subjects actually work on non-specialized tasks.

Figure 6 presents the number of switches in a cumulative distribution function for both treatments. The figure shows clearly the difference between the treatments in specialized behavior. In the High Trust treatment, 66 percent of the participants fully specialize, i.e. they never switched to another field. In stark contrast, only 28 percent of participants in the Low Trust treatment never switched, i.e. fully specialized on their field. The figure shows not only that there are differences on the extensive margin but also that the number of switches are higher in the Low Trust treatment (on average, subjects in this treatment switched 0.92 times) compared to the High Trust treatment (0.5 switches).

While Figure 6 pools our observations across both periods, Figure 7 shows the trend in specialist behavior across periods. While we observe positive trends in specialization across both
conditions, the increase in specialist behavior is only significantly different in the High Trust treatment. Using a logit regression on specialization and clustering at the group level, the high trust groups demonstrably improve from Period 1 to Period 2 \((p < .05)\), while the change for the low trust groups is not significant \((p = .26)\). The differences between the two treatments in the time trend is, however, not significant on any conventional level. But the point estimate indicates that the difference in division of labor in the two environments intensifies over time.

Table 2 shows the results of the Figure 6 and Figure 7 in a regression framework in which the dependent variable in model (1) is a dummy variable that is 1 if a participants fully specialized, i.e. never switching, and 0 otherwise. In model (2), the dependent variable is a count measure on the number of switches. Standard errors are clustered on the group level.

The table confirms the results from the figures: the High Trust treatment increases the probability of fully specializing behavior by 34.3 percent \((p < 0.05)\). The High Trust treatment also decreases the number of switches (Model 2). Participants in the High Trust treatment switch 1.036 times less than participants in the Low Trust treatment. The coefficients on “Round” show that for participants in the Low Trust treatment, the incidence of fully specializing increases somewhat but this increase is not statistically significant \((p = 0.25)\). The interaction “High Trust x Round” in-
Figure 1.10: Division of Labor over Time

dicates that for the High Trust treatment specialized behavior increase further \( (p = 0.327) \). While the difference in the time trend between the treatments is not statistically significant, the increase in specialized behavior is statistically significant for the High Trust treatment \( (p < 0.05) \).

The results so far show that creating a high trust environment endogenously leads to more specialization and less switching tasks. One possible explanation for increased specialized behavior in the high trust groups is that the high trust manipulation might have induced more effort, perhaps because the prime functions as an anchor. However, our prime does not specify that the clicks were made in a specialist field, and thus it is not trivial that specializing would be a straightforward result from a “high effort” anchor. Figure 8 plots effort levels in the two treatment for the two periods together. There are no significant differences in effort overall \( (p=.13) \). Splitting the sample by round, round 1 effort is higher in the High Trust treatment compared to the Low Trust treatment \( (p < 0.05) \), and there is no difference in effort in round 2. Thus, even with indistinguishable effort in round 2, the Low Trust groups exhibit significantly less specialization.

In sum, the experimental evidence shows that exogenously changed trust levels affect how specialized members of a group work together. In a high trust environment, we observe more division of labor, i.e. individuals work on their specialized task and do not switch as often to the
Table 1.3: The Causal Impact of Trust on Division of Labor

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Specialize Dummy</th>
<th># of Switches</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td></td>
</tr>
<tr>
<td>High Trust (=1)</td>
<td>1.372**</td>
<td>-1.036**</td>
</tr>
<tr>
<td></td>
<td>(0.623)</td>
<td>(0.473)</td>
</tr>
<tr>
<td></td>
<td>[0.343]</td>
<td></td>
</tr>
<tr>
<td>Round</td>
<td>0.496</td>
<td>-0.667**</td>
</tr>
<tr>
<td></td>
<td>(0.434)</td>
<td>(0.293)</td>
</tr>
<tr>
<td></td>
<td>[0.124]</td>
<td></td>
</tr>
<tr>
<td>High Trust × Round</td>
<td>0.633</td>
<td>0.273</td>
</tr>
<tr>
<td></td>
<td>(0.646)</td>
<td>(0.331)</td>
</tr>
<tr>
<td></td>
<td>[0.158]</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-1.190**</td>
<td>1.733***</td>
</tr>
<tr>
<td></td>
<td>(0.472)</td>
<td>(0.381)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.190</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>126</td>
<td>126</td>
</tr>
</tbody>
</table>

Table 1.4: *

*Note: Results from a logit (1) and an OLS (2) regression, robust errors clustered at the group level and reported in parentheses. Marginal effects reported in brackets.

***p < 0.01, **p < 0.05, *p < 0.1

Figure 1.11: Effort across Periods and Treatments
task of others. As a result, they also earn more. Payoffs are $15.10 in High Trust groups versus $14.56 in Low Trust groups (p<0.01). Trust makes groups more specialized and more productive, with level of specialization seem to increase over time. The experimental approach allows us to provide internally valid effects of our manipulation (trust) on our outcome measure (division of labor), which complements the results from the cross-country analysis (section 2).

1.6 Conclusion

This paper investigates whether differences in corporate culture can explain the emergence of different organizational structure, especially different degrees of division of labor. We begin our study with a model of the effect of trust on organization form, and then provide evidence on the cross-country correlation between trust and the division of labor. We then experimentally evaluate a productive group that can attain different levels of productivity depending on its level of division of labor. This provides causal evidence on the question of trust and division of labor by showing in an experiment that exogenously ‘shocking’ an organization with high or low trust levels leads to the emergence of different degrees of division of labor. This evidence suggests that one aspect of corporate culture—trust—can have an effect on corporate performance through organizational structure. Demonstrating that trust affects organizational structure is important both because it sheds light on the nature of firms—an age-old puzzle in economics and strategy—and also because it represents a plausible mechanism which could drive the relationship between trust and economic growth.

While the experimental evidence provides high internal validity, the cross-country evidence is high on external validity. Taken together, these two pieces of empirical evidence provide a solid foundation to study further the effect of trust on division of labor.

While demonstrating the causal impact of trust on specialization, this paper raises several important questions and points to areas where further research is needed. Of major interest is how easy it is to change an existing culture of trust, since levels of trust can have long-lasting effects on organizations. To that end it would be useful to document how start-ups confront this problem by
establishing strong cultures. Additional field evidence in this vein would be of significant value. There is also the matter of trust as only one component of corporate culture, but of course culture extends far beyond this and it thought to determine orientations toward innovation, fairness, experimentation and more [52]. Finally, it could be very useful to analyze the effect of trust on the span of control, including task division as well as task allocation. The relationships among organizational form and these other aspects of culture should prove to be a fruitful avenue of exploration.
Chapter 2: Nobody likes a rat: On the willingness to report lies and the consequences thereof

2.1 Introduction

Reporting on another’s deceptive behavior is an act that arouses diverse and conflicting opinions. Children are sometimes scolded for being “tattle-tales” and “snitch” is a derogatory term used to describe those who report on others. Yet this act can also be deemed praiseworthy, as in the case of whistleblowers or crime informants. One stereotype portrays a whistleblower as an ungrateful coworker eyeing a lucrative book deal, while another brings to mind a self-sacrificing hero. In this paper, we study the intrinsic motives of people to report on others’ lies and evaluate the consequences of this act when it can be rewarded by others.

A growing body of research in economics focuses on deception and the inclination of some people to tell the truth despite it being in their material interest to lie [53, 54]. For instance, [57] demonstrate that individuals are averse to lying (to varying degrees), but they can be tempted to lie when doing so is profitable enough. We extend the work on lying aversion by studying whether the predilection to tell the truth also extends to the willingness to uphold a truth-telling norm by punishing and disassociating oneself from people who lie. If this is the case, the cost of deceiving others could be even higher than with lying aversion alone, which has important implications for, among other things, the transmission of information in asymmetric information [58, 59] and the design of institutions that monitor the actions of economic agents (see discussion in [60]).

We run a laboratory experiment in which subjects play a repeated “whistleblowing” game. In each repetition of the game, subjects receive a random number that corresponds to their “real”

\[1\] In this study, we do not distinguish between a preference for truth-telling and an aversion to lying. This distinction is important only if it is possible to not tell the truth and not lie at the same time, e.g., by remaining silent (see [55]). We also do not differentiate between an aversion to lying and an aversion to being observed lying [56].
earnings. Subsequently, they are given the opportunity to lie by overstating their earnings in order to receive a higher payoff. Importantly, subjects are divided into groups within which they can observe each other’s real and stated earnings. Upon observing whether lying occurred, subjects have the opportunity to report others in their group. Subjects who are reported for overstating their earnings are sanctioned by a factor proportional to the size of their lie. However, subjects who report others do not receive any monetary benefits from their action. This game captures basic elements of many situations of interest. In particular, it explores situations where lying is individually profitable but, if discovered by a central authority, it is heavily sanctioned. Critically, the central authority does not monitor behavior—perhaps because it is prohibitively expensive to do so—and instead relies on individuals within the organization to report lying.\textsuperscript{2} We are interested in whether enough people disapprove of and are willing to sanction lying (by reporting it) such that overstatements do not occur.

Some evidence suggests that people are willing to sanction those who tell them lies \cite{61, 62, 63}, even after controlling for the economic damage the lie might have imposed \cite{64}. This evidence is consistent with the enforcement of a truth-telling norm. However, besides being in the pecuniary interest of the person who lies, in these studies a lie is to the detriment of the person being lied to. Therefore, lying also conveys an intention to hurt the person who is subsequently administering the punishment.\textsuperscript{3} In this study, we test the willingness to punish liars even when lies do not affect the pecuniary interest of, and are not directed at, the person administering the punishment. If the punishment of lies occurs in this setting, it indicates that individuals consider lying as behavior that deserves to be sanctioned.\textsuperscript{4}

In addition to studying motivations to report lies, we also evaluate the consequences of reporting. One could reasonably expect that sanctioning behavior that is generally disapproved of would be welcomed. However, some empirical evidence indicates that this is not always the case. For

\begin{itemize}
  \item \textsuperscript{2}A common factor for most whistleblowers is a low cost of obtaining information regarding the malfeasance and the ease with which it can be reported to the appropriate authority \cite{60}
  \item \textsuperscript{3}For evidence that unkind intentions trigger punishment see \cite{65}
  \item \textsuperscript{4}\cite{66} show that some individuals are not willing to lie even when doing so is in the pecuniary interest of the person doing the lying and the person that is being lied to, which is further evidence that lying per se is the act that is disapproved of.
\end{itemize}
example, [60] report that the career prospects of employees who report corporate malfeasance are so dismal that it is actually surprising that people blow the whistle at all. Indeed, the existence of strong community norms against reporting others has been documented by both investigative journalists and academics and is well epitomized by the phrase “snitches get stitches” [67, 68]. These reports point to fear of ostracism and punishment by their peers as one of the main reasons why people do not report others’ wrongdoings [69].

To introduce these types of peer effects in the experiment, we give subjects a say in who gets to join their group. Specifically, every three repetitions of the game, some subjects are randomly removed from their group and have to rejoin a group to continue to earn money. In order to join a group, however, displaced subjects must be accepted by a unanimous vote of the group’s current members. To cast their vote, current group members are informed (i) whether displaced subjects reported others for lying, (ii) whether they were sanctioned for lying, and (iii) their mean stated earnings. This design allows us to determine whether subjects welcome people who report lies or whether they prefer to avoid them. Moreover, we can observe whether subjects who do not lie are the only ones who vote in favor of people who report lying or whether reporting lies is well regarded more generally. Similarly, we can observe whether subjects who regularly lie are the only ones who vote against those who report lying or whether reporting lies is generally poorly regarded. Finally, to determine the importance of these effects in the overall amount of lying and reporting, we run another treatment without voting where displaced subjects rejoin groups at random. These comparisons can give us insights regarding the formation and impact of anti-reporting groups.

Field research that explores the causes and motivations for reporting lies faces a complex task. For one, lies are difficult to observe, and even where lying is evident, extrinsic incentives may exist that can influence the reporting of lies beyond just an aversion to dishonesty. An important example is whistleblowing. In their review, [70] report that whistleblowing is affected by factors that are plausibly unrelated to reporting dishonest behavior, such as qui tam statutes that financially reward whistleblowers. Furthermore, as [71] points out, numerous selection effects make the evaluation of whistleblowing difficult. For instance, if unproductive employees are more likely
to be whistleblowers then we would overestimate the impact of whistleblowing on employability if the job prospects of non-whistleblowers with average productivity are used as the counterfactual. Due to the difficulties posed by field research, an experiment provides the ideal setting to isolate the reasons and consequences for lying and reporting lies. In our experiment, there are no financial incentives for reporting lies and there is no externality associated with lying. Thus, we can be assured that reporting someone for lying reflects intrinsic motives such as upholding a norm of truth telling. Moreover, we avoid the aforementioned selection effects because all subjects are equally productive, they randomly leave groups, and they do not have a say in which groups they want to rejoin. In fact, by comparing the treatment with random reassignment to the treatment where subjects get to choose who joins their group, we can identify the importance of this type of selection on the prevalence of lying.

Our main findings are that, in addition to restraining themselves from lying, enough subjects report those who lie that in the average group lying is largely unprofitable. However, we also find that when group members have a say in who joins their group, subjects who report lies are generally shunned, making this behavior very costly. In fact, the subjects who have better prospects of rejoining (some) groups are those who lie. Accordingly, reporting lies is less common in this treatment and lying is more frequent, particularly in groups that develop a rule of no reporting lies.

2.2 Experimental design and procedures

2.2.1 The whistleblowing game

Here we describe the game used in the experiment in more detail. For simplicity, we describe the game already with the parameters used in the experiment. Consider a “society” composed of $i \in \{1, 2, \ldots, 12\}$ and $g \in \{1, 2, 3\}$ organizations. Each organization $g$ is staffed by $n_g \in \{2, 3\}$ individuals. The game is played repeatedly for nine periods and each period is divided into two stages. In the first stage, each individual first observes their “true” earnings $t_i$, which are independently drawn from a uniform distribution with support $[0, t_{g\text{max}}]$ where $t_{g\text{max}}$ are the maximum earnings in $i$’s organization $g$. The value of $t_{g\text{max}}$ increasing with the size of the organization.
Specifically, we set $t_{g}^{max} = 300$ points for organizations of $n_g = 3$ individuals and $t_{g}^{max} = 225$ points for organizations of $n_g = 2$ individuals. After observing $t_i$, each individual simultaneously decides on the earnings she wishes to state $s_i$. Individuals are free to state any feasible earnings, i.e., $s_i = [0, t_{g}^{max}]$. Barring any sanctions in the second stage, an individual’s payoff is equal to her stated earnings and not her true earnings. In the second stage, all individuals observe both the true and stated earnings of others in their organization. Thereafter, they simultaneously decide whether they report their organization to a central authority. If at least one individual reports her organization, the organization is inspected and all individuals who overstated their earnings, i.e., chose $s_i > t_i$ are sanctioned by three times the overstated amount.\(^5\) Hence, the payoff of individual $i$ of organization $g$ in a period is given by:

$$\pi_i = \begin{cases} s_i - 3(s_i - t_i) & \text{if } g \text{ is inspected and } s_i > t_i \\ s_i & \text{otherwise} \end{cases}$$

At the end of the second stage, individuals are informed of the payoff and actions of all individuals in their organization. Hence, as in other studies of lying behavior, in our game individuals have a monetary incentive to lie by overstating their true earnings, but in our case lying is profitable only if individuals believe that others will not report their actions.

Now we describe how organizational membership is determined. We refer to individuals who are part of an organization as being active and to those who are not as being inactive. At the beginning of the game, all individuals in the society are randomly assigned to one of the three organizations and are therefore active individuals. However, every third period (i.e., after periods 3 and 6 are played), one individual in each organization of $n_g = 3$ is selected at random to be separated from it and hence become an inactive individual.\(^6\) Moreover, everyone in the society gets to observe the following information of each inactive individual: (i) their mean stated earnings

\(^5\)To avoid confusion in the experiment, subjects were given the opportunity to report only if one or more individuals overstated their earnings.

\(^6\)Organizations with $n_g = 2$ individuals do not lose members. This was done to avoid the added complexity of disappearing organizations.
over the last three periods, (ii) whether they reported their organization in the last three periods, and (iii) whether they were sanctioned for overstating their earnings in the last three periods. Before play resumes, inactive individuals have the opportunity to rejoin an organization. Individuals who remain inactive do not receive or state earnings and obtain a payoff of $\pi_i = 0$ points per period. In the experiment, we implement two treatments, each using a different procedure to convert inactive individuals into active individuals. In the Random treatment, all inactive individuals are randomly reassigned to organizations.\footnote{Note that in Random nobody remains inactive and individuals are simply being randomly reassigned in a similar way as in experiments with a strangers matching protocol. Unlike these experiments, here only some individuals are reassigned and there are information asymmetries between individuals who join a group and those who are already in it (unless an individual is reassigned to the same group she used to belong to).}

In the Selection treatment, inactive individuals must be accepted into organizations using the unanimity rule. Specifically, active individuals first indicate whether they accept or veto each inactive individual. Thereafter, inactive individuals are randomly assigned to an organization where they were unanimously accepted. If no such organization exists, the individual remains inactive for the next three periods (until the next vote or until the game ends). Meanwhile organizations with $n_g = 2$ individuals continue to play but with reduced maximum earnings of $t_{max}^g = 225$ points.

2.2.2 Theoretical Predictions

We now briefly discuss the theoretical predictions of the game. We start with the assumption that all individuals are risk neutral and own-earnings maximizers. In this case, in both treatments and during all periods, all individuals always state the maximum earnings and they never report others because they are indifferent to the outcomes of other players and reporting carries the opportunity cost of lying.\footnote{If an organization is inspected then all members who overstated, including the reporter, get sanctioned.}

Finally, in the Selection treatment, all inactive individuals are accepted because organizations with $n_g = 3$ individuals allow for higher stated earnings than organizations of $n_g = 2$ individuals and vetoing someone carries the risk of ending up with a smaller organization.\footnote{More precisely, strategies where active individuals coordinate their voting such that all organizations end up with three players are also equilibria of the game. However, accepting all inactive individuals weakly dominates these voting strategies and requires no coordination, making it the safer option.}
We next discuss how the above predictions change if some individuals have an aversion to lying. Namely, they incur disutility if they tell a lie. By assumption, models of lying aversion predict less overstating in the whistleblowing game as some individuals state their earnings truthfully. None of these models, however, consider motivations for individuals to report others for overstating. It is not our aim to develop such a model here. Instead, we hope to contribute to an understanding of reporting behavior so that such models can be developed in the future. For now, we simply provide an informal discussion of what we could expect if we assume that, in addition to honest individuals who incur disutility if they lie, there are also indignant individuals who incur a utility loss when they or others in their organization lie, but whose utility loss is smaller if liars are sanctioned. If the parameters were such that indignant individuals are willing to report overstatements, lying behavior would be further reduced or even eliminated. Introducing indignant individuals also allows for differences in behavior between the Random and Selection treatments. In particular, active individuals who are neither honest nor indignant might find it optimal to veto inactive individuals who reported others. In fact, we chose the parameters of the game so that two individuals who are willing to lie are better off overstating maximally in an organization of \( n_g = 2 \) compared to telling the truth in an organization of \( n_g = 3 \) (in the former they guarantee themselves 225 points and in the latter their expected payoff equals 150 points). In a similar way indignant individuals might find it optimal to veto inactive individuals who (are likely to have) overstated. This introduces an incentive for individuals to sort themselves into honest and dishonest organizations much like the anti-snitching groups described in the introduction. Depending on the frequency of each type, the risk of being left without an organization makes one’s behavior relatively more costly. For instance, if indignant individuals are relatively rare then they would face a high probability of rejection by organizations in Selection treatment, which might be enough to make reporting less common or nonexistent. In this regard, it is interesting to note that indignant individuals might face rejection also by honest individuals. The intuition behind this is that even individuals who have a high cost of lying understand that they might be tempted to lie if the gain from doing so is truly large. Consequently, honest individuals might prefer not to have indignant
We would like to point out that in addition to indignation at observing lies, other motivations for reporting overstatements exist. In particular, reporting could be motivated by lying averse individuals who are also envious (i.e., they dislike disadvantageous inequality, [72, 73]) or possess competitive preferences (i.e., they like advantageous earnings disparities [74, 75]). In our statistical analysis, we test whether these motivations are present by looking for a positive correlation between reporting others and the difference in the reporter’s and liar’s stated earnings.

2.2.3 Experimental Procedures

The computerized experiment was conducted using the typical procedures of anonymity, neutrally worded instructions, and monetary incentives. In total, 192 students participated in the one-hour experiment. Each session of the experiment consisted of 24 subjects who were randomly assigned to one of two societies. We employed a between-subjects treatment design so that in each session subjects in one society played the Random treatment and those in the other played the Selection treatment. This gives us eight independent observations (societies) per treatment. Points were converted to US dollars at a rate of 150 points = $1 and subjects received a show-up fee of $15. Average earnings equaled $22.84.

Upon arrival, each subject drew a card to be randomly assigned to a seat in the laboratory. Once everyone was seated, subjects read the instructions for the experiment (a copy of the instructions is available in the online appendix) and answered a few questions to ensure their understanding of the game. When all subjects had correctly answered the questions, the computerized experiment (programmed in z-Tree, [76]) started. After the game ended, subjects were confidentially paid their earnings in cash.

2.3 Results

In this section, we present the experimental results. We start by analyzing the subjects’ over-stating behavior (3.1). Thereafter, we examine their reporting behavior and the ensuing sanctions
(3.2). Lastly, we discuss the subjects’ voting behavior in the Selection treatment (3.3) and how it affects the profitability of overstating earnings (3.4). Note that, throughout this section, we use regression analysis to test the statistical significance of our findings. All regressions use robust standard errors clustered on societies and report p-values values based on two-tailed tests. Moreover, unless it is otherwise noted, we include subject random effects. This method allows us to utilize fully the panel structure of our data.

\[ \frac{(s_i - t_i)}{t_{\text{max}} - t_i} \]

The distribution of this variable is depicted in the left panel of Figure 1. In both treatments, the modal behavior is to be honest: subjects choose \( s_i = t_i \) 62 percent of the time in Random and 42 percent of the time in Selection. However, there are plenty of cases where subjects state higher earnings than the ones they received. In fact, when subjects overstate their earnings they usually choose the maximum earnings they could possibly state: subjects choose

![Distributions of overstatement behavior](image)

**Figure 1 - Distributions of overstatement behavior**

*Note:* The left panel shows the amount overstated as a fraction of the maximum that could have been overstated. The right panel shows the percentage of periods in which subjects overstate.

### 2.3.1 Overstating

To describe the degree to which subjects overstate their earnings, we look at the number of points a subject overstates in a given period as a fraction of the points she could have overstated, i.e., \( \frac{(s_i - t_i)}{t_{\text{max}} - t_i} \). The distribution of this variable is depicted in the left panel of Figure 1. In both treatments, the modal behavior is to be honest: subjects choose \( s_i = t_i \) 62 percent of the time in Random and 42 percent of the time in Selection. However, there are plenty of cases where subjects state higher earnings than the ones they received. In fact, when subjects overstate their earnings they usually choose the maximum earnings they could possibly state: subjects choose
15 percent of the time in Random and 39 percent of the time in Selection (see Figure 1). Thus, conditional on overstating, subjects lie as much as they can 61 percent of the time across the two treatments. In the right panel of Figure 1, we depict the distribution of the percentage of periods each subject overstates. In both treatments, we can see that the most common behavior is to overstate earnings in some but not all periods. In Random, 20 percent of subjects never overstate and 5 percent always overstate, which leaves 75 percent who sometimes overstate. In Selection, 19 percent never overstate, 20 percent always overstate, and 61 percent sometimes overstate. Hence, overstatement behavior across our whole game is consistent with studies that report that people are willing to lie but do not do so maximally (e.g., [77]).

We run a series of regressions with a treatment dummy variable as the independent variable to test whether overstating behavior differs between Random and Selection. The dependent variable in the first regression is the amount overstated as a fraction of the highest feasible overstatement \( \frac{(s_i - t_i)}{t_{\text{max}} - t_i} \). We use Tobit estimates and censor the variable at 1 and 0. The dependent variable in the second regression is simply a dummy variable indicating whether a subject overstates, i.e., chooses \( s_i > t_i \). Since the dependent variable is binary, we use logit estimates. We find that both the amount overstated and the frequency of overstatements are significantly higher in Selection \( (p = 0.010 \text{ and } p = 0.046, \text{ respectively}). \)

Interestingly, the difference between treatments occurs from the beginning. This is illustrated in Figure 2, which plots the means of these two variables over time. As we can see, the amount overstated and the frequency of overstatements is higher in Selection already in the first three periods.\(^{10}\) Hence, even before the selection of inactive subjects takes place, subjects expect that the selection process allows them to overstate more often.

\(^{10}\)Using the same regressions to test for treatment differences but restricting them to the first three periods gives \( p = 0.008 \) for the amount overstated and \( p = 0.046 \) for the frequency of overstatements.
Next, we take a closer look at how subjects adjust their overstating behavior. To do so, we run two logit regressions per treatment. As the dependent variable we use a dummy variable indicating whether subject $i$ overstates her earnings in period $z$. In all cases, we use subject fixed effects, period fixed effects, and robust standard errors clustered on societies.

In the first two regressions we concentrate on the effect of $i$’s previous actions and on factors that directly affect $i$’s monetary payoff. Specifically, we use the following independent variables:

(i) $i$’s temptation to overstate in period $z$, which we define as $i$’s monetary gain of reporting the highest possible earnings, $t^m - t_i$;
(ii) a dummy variable equal to one if $i$ overstated her earnings in period $z = 1$;
(iii) a dummy variable equal to one if $i$ was sanctioned for overstating her earnings in period $z = 1$ (because another subject chose to report); and
(iv) a dummy variable equal to one if $i$ reported another subject in period $z = 1$. In the subsequent two regressions, we also include events observed by $i$ that have no direct effect on her payoff. Namely, we add in (v) a dummy variable equal to one if $i$ observed at least one other subject overstate their earnings in period $z = 1$, and in (vi) a dummy variable equal to one if $i$ was not sanctioned but observed another subject being

![Figure 2 - Overstating behavior over time](image)

Note: The left panel shows the mean amount overstated as a fraction of the maximum that could have been overstated. The right panel shows the mean fraction of subjects that overstate. Error bars correspond to 95 percent confidence intervals.
sanctioned in period \( z - 1 \) because a subject other than \( i \) chose to report. The estimated odds ratios and respective standard errors are presented in Table 1.\(^{11}\)

**Table 1 – Determinants of overstating**

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>Specification I</th>
<th>Specification II</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Random o.r. s.e.</td>
<td>Selection o.r. s.e.</td>
</tr>
<tr>
<td>Temptation to overstate</td>
<td>1.93*** (0.22)</td>
<td>2.29*** (0.34)</td>
</tr>
<tr>
<td>Overstated in ( z - 1 )</td>
<td>4.68*** (0.99)</td>
<td>2.66*** (0.96)</td>
</tr>
<tr>
<td>Sanctioned in ( z - 1 )</td>
<td>0.06*** (0.03)</td>
<td>0.37** (0.15)</td>
</tr>
<tr>
<td>Reported in ( z - 1 )</td>
<td>0.35*** (0.08)</td>
<td>0.48** (0.14)</td>
</tr>
<tr>
<td>Others overstated in ( z - 1 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others sanctioned in ( z - 1 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Period fixed effects</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Subject fixed effects</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Pseudo ( R^2 )</td>
<td>0.21</td>
<td>0.16</td>
</tr>
<tr>
<td># of obs./subj./societies</td>
<td>648/72/8</td>
<td>462/59/8</td>
</tr>
</tbody>
</table>

*Note: Logit regressions with a dummy variable indicating whether subjects overstate their earnings as the dependent variable. Clustered standard errors allow for correlation within societies. Asterisks indicate significance at the 1 percent (***) 5 percent (**), and 10 percent (*) level.*

Consistent with models where lying behavior is sensitive to the benefits of lying (e.g., [53]), Table 1 shows that subjects are significantly more likely to overstate their earnings as the gain from doing so increases. Specifically, an increase of one standard deviation in the temptation to overstate roughly doubles the odds of over-stating. As is common in laboratory experiments, subjects tend to repeat actions that resulted in a high payoff and avoid actions that gave them a low payoff. Specifically, subjects who overstated and were not sanctioned have increased odds of over-stating whereas subjects who overstated and were sanctioned have decreased odds of over-stating.

With the second specification, we see that subjects are learning from observing the actions of others. That is, observing other subjects overstate increases the odds of over-stating and observing other subjects be sanctioned for over-stating decreases the odds of over-stating.\(^{12}\) Finally, note

\(^{11}\)To facilitate the interpretation of the coefficient of (i), we normalize this variable to have a mean of zero and a standard deviation of one. Moreover, since we use subject fixed effects and there are subjects who never or always overstate, the coefficients are estimated using fewer than 96 subjects and 864 observations.

\(^{12}\)We also observe that subjects are less likely to overstate if they reported someone in the previous period. It is not clear why this might be the case, but it could be due to subjects expecting retaliation from subjects who got sanctioned.
that by and large the dynamics of overstating are similar in both treatments. Hence, the effect of selecting who enters the organization does not seem to affect how subjects react to each other’s actions. Hence, the effect of selecting who enters the organization does not seem to affect how subjects react to each other’s actions.

We summarize these findings as our first result.

**Result 1 (Overstating earnings):** Subjects regularly overstate their earnings, albeit, most do not do so in every period. The possibility of selecting who enters their organization facilitates overstating even before selection takes place. Overstating is more likely if the gain from lying is large and subjects observe or experience unpunished lying in the past.

2.3.2 Reporting Behavior

In both treatments, the overall rate at which subjects report others is not high, which results in few instances in which subjects are sanctioned. In Random, subjects report others 17 percent of the time, resulting in 19 percent of choices being sanctioned, while in Selection they report others 11 percent of the time resulting in 13 percent of choices being sanctioned. Logit regressions confirm that reporting others for overstating and the ensuing sanctions are significantly more common in Random compared to Selection ($p < 0.032$). However, it is more informative to look at reporting behavior conditional on there being a reason to report.
Figure 3 plots the fraction of subjects who report others given that at least one other subject overstated (left panel) and the fraction of subjects who are sanctioned conditional on having had overstated (right panel). Subjects report others when they see them overstate 32 percent of the time in Random and 17 percent of the time in Selection (significantly different with a logit regression, $p = 0.028$). We also see an important treatment difference in the fraction of subjects who never report: 28 percent in Random and 53 percent, a majority, in Selection. The fact that the reporting rate varies between treatments suggests that the willingness to report is sensitive to the incentives introduced by the selection process. We see an even larger difference between treatments in the fraction of overstatements that are sanctioned: a sizable 56 percent of overstatements are sanctioned in Random but only 24 percent in Selection (significantly different with a logit regression, $p = 0.006$).
To evaluate the determinants of reporting we run two logit regressions per treatment. The dependent variable is a dummy variable indicating whether subject $i$ reports other subjects for overstating their earnings in period $z$. In all cases, we use subject fixed effects, period fixed effects, and robust standard errors clustered on societies. In the first two regressions we concentrate on the effect of $i$’s actions and others’ actions in period $z$. We use the following independent variables: (i) the amount by which others overstate their earnings in period $z$, $\sum_{j\neq i}(s_j - t_j)$, which measures the extent to which others lie; (ii) the difference in stated earnings between $i$ and others in period $z$, $\sum_{j\neq i}(s_j - s_i)$, which captures incentives to report that are unrelated to lying but might be important such as inequity aversion and competitive preferences; (iii) a dummy variable equal to one if $i$ overstated her earnings in period $z$; (iv) an interaction term between (i) and (iii); and (v) an interaction term between (ii) and (iii). In the other two regressions, we also control for events in the previous period. To this end, we use (vi) a dummy variable equal to one if $i$ was sanctioned for overstating her earnings in period $z - 1$; (vii) a dummy variable equal to one if $i$ reported others in

<table>
<thead>
<tr>
<th>Table 2 - Determinants of reporting overstatements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Independent variables</td>
</tr>
<tr>
<td>Deviation from true earnings</td>
</tr>
<tr>
<td>Difference in earnings</td>
</tr>
<tr>
<td>Overstated</td>
</tr>
<tr>
<td>Overstated × dev. true earnings</td>
</tr>
<tr>
<td>Overstated × diff. earnings</td>
</tr>
<tr>
<td>Sanctioned in $z - 1$</td>
</tr>
<tr>
<td>Reported in $z - 1$</td>
</tr>
<tr>
<td>Others reported in $z - 1$</td>
</tr>
<tr>
<td>Period fixed effects</td>
</tr>
<tr>
<td>Subject fixed effects</td>
</tr>
<tr>
<td>Pseudo $R^2$</td>
</tr>
<tr>
<td># of obs./subj./societies</td>
</tr>
</tbody>
</table>

Note: Logit regressions with a dummy variable indicating whether subjects report others as the dependent variable. Clustered standard errors allow for correlation within societies. Asterisks indicate significance at the 1 percent (**), 5 percent (**), and 10 percent (*) level.
period $z \geq 1$; and (viii) a dummy variable equal to one if $i$ observed at least one other subject being sanctioned in period $z \geq 1$. The estimated odds ratios and standard errors are presented in Table 2.\textsuperscript{13} Sanction them for doing so. A second powerful determinant of the decision to report is having had overstated. Subjects who overstated are significantly less likely to report others, which is not very surprising as by reporting they also sanction themselves. By contrast, the effect of differences in earnings is significant only in Selection and is markedly smaller compared to the previous two. Lastly, note that the interaction term for having overstated times the deviation from true earnings is not significant in the Selection treatment, which shows there is some willingness to report others even when doing so implies incurring a sanction. We should note, however, that overstating and then reporting is a rare occurrence, so these latter results are based on few observations (11 out of 585 in Random and 10 out of 318 in Selection).

### Figure 4 – Vetoing depending on the subject’s behavior

*Note:* From left to right, the panels show the fraction of active subjects who veto an inactive subject depending on whether the inactive subject: reported others, was sanctioned for overstating, or stated earnings above 233 points. Within panels, voting behavior is separated depending on how often the active subject overstated in the previous three periods. Error bars correspond to 95 percent confidence intervals.

\textsuperscript{13}As before, we normalize the continuous variables (i) and (ii) to have a mean of zero and a standard deviation of one. Moreover, since we use subject fixed effects and some subjects never or always report, the coefficients are estimated using less than 96 subjects and 864 observations.
It is worth mentioning that the coefficients of the two most impactful variables (i.e., the deviation from true earnings and having had reported) are not significantly different across treatments. This shows that the lower reporting rate in Selection is not due to honest subjects being less willing to report, but instead it is due to fewer instances where honest subjects see others overstate. In the next subsection, we analyze how the selection process facilitates this effect by separating subjects who report from those who overstate.

The following result summarizes the findings of this subsection.

**Result 2** (Reporting overstatements): *In the absence of selection, enough subjects report others so that subjects who overstate are sanctioned relatively often. If it is possible for organizations to select their members, reporting is much less frequent and overstatements are sanctioned less often. If they are not lying themselves, subjects are more likely to report others the more egregious the
lying in their organization is.

2.3.3 Organization member selection

In this section, we restrict our analysis to the Selection treatment in order to assess how people decide whom to accept into their organization as well as how the composition of organizations changes over time. Figure 4 presents the fraction of active subjects who veto an inactive subject depending on their behavior in the last three periods. Across panels, it displays the effect of the inactive subjects’ behavior that is available to active subjects when they cast their vote. From left to right, the figure shows whether inactive subjects reported others, whether they were sanctioned for overstating, and whether their stated earnings were suspiciously high (earnings that occur with less than 5 percent probability by chance alone; 35 percent of inactive candidates fall in this category). Within panels, the figure plots separately the voting behavior of active subjects depending on how often they overstated in the previous three periods.

First, let us consider whether active subjects accept or reject inactive subjects who reported others for lying. As we can see, inactive subjects who never reported others fare much better than those who did (the former are vetoed by 48 percent of active subjects while the latter are vetoed by 70 percent). This pattern is expected for active subjects who always or sometimes overstated. What is more interesting is that the same pattern appears for active subjects who never overstated. However, as we pointed out in Section 2.2., this behavior is consistent with a model in which subjects who are generally averse to their own (but not others’) lies understand that they might be tempted to lie in some future instances, and hence they prefer not to have individuals who report others in their organization. Next, we turn to inactive subjects who are known for, or can be suspected of, lying. We see that overstating actually facilitates rejoining organizations. Inactive subjects who display suspiciously high earnings or who were sanctioned for overstating are vetoed by a smaller fraction of active subjects (30 and 48 percent, respectively) than subjects with lower earnings or without sanctions (67 and 61 percent, respectively). Moreover, as before, this pattern is similar for active subjects who never, sometimes, and always lied.
We test whether these differences are statistically significant with two logit regressions. In both cases, the dependent variable is a dummy variable equal to one if active subject $i$ vetoes inactive subject $j$. In both cases, we use subject fixed effects, period fixed effects, and robust standard errors clustered on societies. In the first regression, the independent variables are: (i) a dummy variable equal to one if $i$ was sanctioned for overstating; (ii) $i$’s mean stated earnings; and (iii) a dummy variable equal to one if $i$ reported others for overstating. In the second regression, each of these variables is interacted with three dummy variables: one equal to one if $i$ never overstated, one equal to one if $i$ sometimes overstated, and one equal to one if $i$ always overstated. The results are displayed in Table 3.

We can see that the odds of vetoing significantly increase when the inactive subject is known to have reported others and the effect is significant irrespective of the overstating behavior of the active subjects. We also see a significant effect for the inactive subjects’ stated earnings: the higher the stated earnings, and hence the probability of having had overstated, the higher the chance of acceptance. Finally, once we control for other variables we do not see that being sanctioned significantly influences vetoing.

We next evaluate the impact of the selection process on organizational composition. To this end, we look at the fraction of times subjects overstate in each organization. In the first three periods, before selection takes place, in both Random and Selection the vast majority of organizations display a combination of truthful and untruthful statements. The fraction of fully honest organizations where no subject overstates in any period is 9 percent in Random and 6 percent in Selection. Similarly, the fraction of fully dishonest organizations where all subjects overstate in all periods is 3 percent in Random and 9 percent in Selection. After the third period, however, we see a noticeable change in the distribution of organizations in Selection. This is illustrated in Figure 5. The figure plots the fraction of organizations depending on the fraction of times subjects overstate in periods 4 to 9. As we can see, while in Random there are no fully dishonest organizations, in Selection the modal organization is in fact the fully dishonest one. By contrast, the fraction of organizations that are fully honest is still not that far apart: 9 percent in Random and 15 percent in
Figure 5 presents suggestive evidence that the selection process in the experiment is sufficient to create a substantial number of fully dishonest organizations. However, it is conceivable that the difference between the two distributions is simply the result of individual subjects having a higher propensity to overstate in Selection, and it is not due to selection and peer effects. To assess the importance of these effects, we simulate how the distributions in Figure 5 would appear if the subjects’ propensity to overstate is independent of the organization they are in (plotted as dashed lines).\textsuperscript{14} Compared to the simulated distributions, there is a larger fraction of organizations that are completely honest in both treatments. This suggests that observing each other’s statements, which is the case in both treatments, is enough to produce some conformity in the direction of full honesty. In addition, in Selection we see an even larger difference for the fraction of organizations that are fully dishonest. More specifically, while in Selection a fifth of the organizations are fully dishonest.

\textsuperscript{14}Specifically, we simulate the following scenario 50,000 times: first, we calculate the overstatement rate of each subject (in periods 4 to 9); second, we randomly reassign subjects to organizations (within each treatment); and third, we compute the fraction of overstatements that would occur in each organization if the reassignment does not change the subjects’ overstatement rate.
dishonest, the probability of observing a fully dishonest organization by randomly regrouping individual overstatement rates is only 4 percent. This is compelling evidence that granting subjects the ability to reject others from joining their organization is enough to create organizations with strong norms against reporting others.

We summarize the findings from this subsection in the following result.

**Result 3** (Acceptance into organizations): *Inactive subjects who reported others face more resistance when rejoining organizations as they are rejected both by subjects who overstate and those who do not. By contrast, inactive subjects with high stated earnings, which suggest that they overstated, are generally welcomed. This selection process leads to the formation of a sizable number of highly dishonest organizations with strong norms against reporting others.*

2.3.4 Final payoffs

The subjects’ final payoff is very similar in both treatments. On average, subjects make 129 points per period in Random and 133 points in Selection (not significantly different with a GLS regression, \( p = 0.791 \)). Hence, the additional points subjects make in Selection due to more overstating and lower sanctions are canceled out by a smaller \( t_{\text{max}}^g \) in some instances and the low earnings of inactive subjects who do not manage to rejoin an organization (18 percent of subjects remain inactive after the third period). What is more interesting is to evaluate the effect of overstating and reporting on the subjects’ final payoff.
To see whether overstating pays off in each treatment we calculate each subject’s payoff over the nine periods and the fraction of times each subject overstated and reported others. For each treatment, the left panel in Figure 6 plots the relationship between final payoffs and the fraction of overstatements (divided into terciles). In Random, we see that the tercile of subjects who overstate the most have lower payoffs than subjects who overstate less. By contrast, in Selection the converse is true. Subjects who overstate the most have higher payoffs than subjects who overstate less. On the right panel of Figure 6, we plot the relationship between final payoffs and the fraction of times subjects reported others (also divided into terciles). In this case, we see that reporting others has no perceivable effect on payoffs in Random but it has a strong negative effect in Selection. To test whether these relations are statistically significant, we regress the subjects’ payoff on the fraction of times each subject overstated or the fraction of times each subject reported others. We run the GLS regressions using society fixed effects. In Random we obtain a significantly negative coefficient for the fraction of overstatements ($\beta = 67$ points, $p = 0.012$) and a positive but not significant coefficient for the fraction of reports ($\beta = 28$ points, $p = 0.519$). By contrast, in Selection the coefficient for fraction of overstatements is significantly positive ($\beta = 59$ points, $p = 0.010$) and

\[\text{Figure 6 - Final payoff depending on the amount of overstating and reporting} \]

*Note: Error bars correspond to one standard error.*
the one for the fraction of reports is negative and weakly significant ($\beta = 76$ points, $p = 0.081$).

The findings from this subsection give us our last result.

**Result 4 (Final payoffs):** In the absence of selection, dishonest subjects have lower payoffs than honest subjects. However, if it is possible for organizations to select their members, it is profitable to overstate one’s earnings because dishonest subjects obtain higher payoffs than honest subjects do. In this situation, the lowest payoff is obtained by subjects who report others for overstating.

### 2.4 Conclusions

In this paper, we study whether individuals are willing to sanction other people if they lie for personal gain by reporting them to a central authority. We find that, in randomly assigned groups, enough people are willing to report lying to make lying unprofitable. We also investigate how the frequency of lying and reporting lies is affected when groups can select their members. Our results indicate that this type of selection is enough to increase the frequency of lying and decrease the amount of reporting by facilitating the formation of a significant number of dishonest groups where lying is prevalent and reporting is nonexistent.

The fact that some individuals report others for lying even when lies are not directed toward those doing the reporting and cause no obvious harm to third parties suggests that lying per se is considered by some as normatively undesirable behavior that ought to be punished. This fact calls for the modeling of lying as more than an individual cost and more as a social norm (e.g., [78]). In such a model, whether individuals lie (or report others) would depend on not only the monetary benefit (or cost) of their action but also on the expectations of others regarding their behavior ([79, 80]). This approach to lying aversion helps explain why lying is accepted in some contexts, such as when playing poker, but not in others.

As in most experimental studies, our study abstracts away from many elements present in the field in order to cleanly identify specific effects or motivations. Hence, some caution ought to be exerted when thinking on the external validity of our findings. The presence of individuals who are intrinsically motivated to report dishonest behavior is encouraging and provides one answer to
the question posed in the whistleblowing literature of why some individuals are willing to blow the whistle even though the consequences of doing so are dire. However, this does not mean that extrinsic motivations are not important or even the main factor behind many whistleblowing cases. Indeed, the whistleblowing game can be easily modified to evaluate the relative impact and interaction of intrinsic and extrinsic incentives to report on others’ dishonesty. This type of research can potentially inform policies such as the granting of immunity to whistleblowers or rewarding them through qui tam statutes.

Our study also suggests that relying on reports from individuals within organizations to discourage dishonest behavior is a lost cause if others can identify the reporters and subsequently avoid them. In fact, selection effects might be a lot stronger in the field as selection in our experiment is limited to rejecting individuals from joining organizations. An interesting avenue of research would be to observe the effect of other possible selection mechanisms. For instance, slight modifications to our game can be used to evaluate the effect of reporting others on the likelihood of being unjustly “fired”. Alternatively, if individuals are given the option to leave then honest individuals might prefer to avoid dishonesty by leaving as opposed to reporting the organization.

Finally, our work also highlights that reporting or sanctioning others for being dishonest is not behavior that is always well received or sought after even by individuals who act honestly. This is an important finding because it implies that reporting dishonest actions is very costly as reporters can be ostracized even if most organizations are in fact honest. This helps explain the dismal career prospects of employees who blow the whistle [60] and calls for caution when it comes to policies or actions that reveal the identity of whistleblowers.
Chapter 3: NFTs, Volume, and Social Influence

3.1 Introduction

In 1992 Reiner Knizia, the "Mozart" of strategic board games, created a game based on the art market. This game, called *Modern Art*, sees players compete to own "valuable art" cards by auctioning and buying them off one another. The winner is the player who owns the most valuable art cards at the end of the game, with a card’s value being "exclusively determined by how many times the dealers traded it."[81] That is, in *Modern Art*, Transaction Volume determines the value of art.

The board game *Modern Art’s* assignation of value is artificial; a product of the game rules. But in the real, $17 billion world of Non-Fungible Tokens (NFTs) there is a remarkably similar premium placed on volume. For instance, consider the website on which the great majority of the $17 billion of trading activity took place: Opensea.io. Opensea displays "NFT rankings" (Figure 1) and, by default, "Top NFTs" are those with the highest transaction volume. The next most popular NFT Marketplaces, LooksRare and Rarible, follow suit as well, prominently displaying NFT rankings which default to volume.¹ Popular NFT analytics sites like CryptoSlam.io and Nonfungible.com each feature a ranked list of NFTs which defaults to volume as well.

Why has volume emerged as the focal metric? NFTs—unique, digital blockchain objects—don’t have any obvious methods by which value can be determined. Digital Property may feel "real" to some parties and be valued along idiosyncratic and subjective dimensions [82] but there are few ways of assigning an "objective" value since, for instance, owning an NFT doesn’t entitle one to a stream of dividends as a stock might. So, from this view, Volume might be as good as any other measure. Furthermore, rankings themselves are observed to create status hierarchies [83],

¹More precisely, it sorts NFT collections by the volume of the collection. This is detailed elsewhere in the paper in the section "The NFT Market".

55
impacting reputations and allowing those that come out on top to command higher prices.[84] Status hierarchies can create self-fulfilling prophecies wherein the hierarchy of goods affects the perception of the goods, then feeding back to strengthen the hierarchies themselves, creating a self-perpetuating cycle. Salganik and Watts [85] observe such an effect in a cultural market for music driven by social influence. Thus it’s possible that the attention given to NFT volume is purely artificial, perhaps driven by the decision of the marketplace, and sustained through path dependence a self-fulfilling feedback loops. But rankings offered by multiple sites suggest volume is a social test, and thus we may expect that volume as a criteria is "the outcome institutional processes"[83, 86]. One possibility is that NFT marketplaces prefer high-volume NFTs because they provide more revenue. Given the percentage-fee-per-transaction structure of the major NFT marketplaces this is certainly plausible. But if volume was not informative to consumers and traders we might expect the major analytics sites to offer alternative rankings that were more desired. Thus the puzzle remains: is there something important about volume?

Figure 3.1: A screenshot of opensea.com/rankings (the default ranking is volume.)
3.1.1 Volume as information and influence

In traditional asset pricing theory, current prices reflect all available information, including volume. And even models that allow for heterogenous information have a difficult time accounting for extent of volume observed in, for example, the US stock market. It’s thus assumed that "the bulk" of observed trading volume must be effectively noise "with no consequences for prices". [87]

Despite the theoretical irrelevance of volume to prices, empirical research on the stock market has identified a "volume return premium" in which a positive volume shock to an asset seems to predict excess returns for that asset. [88] One explanation for this effect is attentional, drawing on a model from Merton [89] in which stock visibility helps guide investor attention. By channeling investor attention, visibility increases the investor base for a stock, lessening the firm’s cost of capital and, finally, increasing firm’s market value and price. Insofar as high volume creates visibility for stocks and drives attention, as observed in Shive [90], this could provide a mechanism by which volume could affect price.

Attention is one mechanism through which volume could affect the price of NFT artworks. But NFT Artworks, as art, are cultural objects traded in cultural markets. Such markets, where quality is uncertain, may be affected by social influence.[91] Classic work on social influence suggests two distinct channels: informational influence as we learn new information from others, and normative influence as we feel pressure to accept the opinions of others as valid and appropriate [92]. The latter would be the case if an observing high volume for an NFT caused an individual to like that NFT more (or less.) [85]

In the work that follows, I consider three primary questions to be investigated:

1. Does volume causally affect price in NFT markets? And if so...

2. Is attention a channel through which volume affects price?

3. Is normative social influence a channel through which volume affects price?
To explore the first question, I use observational data on NFT markets to observe the affect of volume on price. I exploit a unique aspect of the Ethereum Blockchain—the fluctuation in transaction fees across the network—as a source of exogenous variation in volume in order to demonstrate a causal effect. To speak to the mechanisms of attention and social influence, I make use of a laboratory experiment in which subjects are told to rate NFT artworks that are presented to them from highest volume to lowest volume. This is then compared against a control condition where different subjects are told the same NFTs (in the same order) are ordered by artist name, then asked to rate them.

A second treatment is added to the experiment in order to establish whether "credible" volume—volume ranked by the cost of transactions—has a distinct effect from "superficial" volume. If volume affects price, then we may expect sellers to drive volume up before selling in an attempt to drive up the price. And indeed this phenomenon of "volume manipulation", in which volume is inflated by flipping artworks back and forth via self-financed sales, is widely observed in NFT markets. Because of this, analytics sites like CoinGecko will often blacklist or ignore volume information from exchanges with minimal transaction fees as the volume data isn’t credible. And so, whereas my first experimental treatment tells subjects that works are ranked simply by the number or transfers, the second condition informs subjects that works are ranked by the costs paid for the transactions. This helps better understand the assurance that transaction fees seem to provide in the marketplace.

3.2 An NFT-focused overview of the Ethereum Blockchain

3.2.1 The Ethereum Blockchain

A blockchain, originally proposed by Nakamoto [93], is a distributed database that cryptographically links and protects data records according to a specified protocol. Most popular NFTs are on the Ethereum Blockchain, and so for convenience most of the specifics of the following discussion will regard Ethereum. Ethereum’s blockchain is best described as a distributed state machine, a quasi-Turing-complete computer that updates its (virtual) machine state by accepting a
limited set of state transition instructions in a given period of time.[94] On Ethereum, distributed computers which run Ethereum software ("nodes") replicate the current state of the Ethereum machine and remain synced with one another by accepting and executing all valid state transitions. What determines the validity of a set of state transitions (a "block") is (1) whether it meets all the protocol rules, for instance all transactions having valid signatures, and (2) that the block has a "certificate of legitimacy", which certifies the completion of a mathematical puzzle.²

The entities who compete to earn Certificates of Legitimacy, thereby gaining the ability to propose a block which Ethereum nodes will accept, are called miners. Because block size is limited, a winning miner must make a choice about what to include in their valid block. Though in principle miners may add any valid transaction they like to the block, miners typically choose to include transactions which pay them the highest fees. And since space is limited, users are willing to offer these fees in order to increase their probability of being included in the winning miner’s next block. The fee required to be added to the next block is referred to as the "gas price", and it denotes the price per unit of computation of block inclusion (as measured with the smallest unit of Ethereum’s native currency.) See Figure 2 for a chart of gas prices from July 2020 to July 2021.

So why wouldn’t a miner add all fee-paying transactions to a block? The reason is that there is a computational limitation on the size of blocks, referred to as Block Gas Limit. This limit is imposed for security reasons, since higher amounts of computation would slow down the network as nodes took longer to verify the correctness of computation and sync with the network. It might effectively mean that slower nodes with less memory couldn’t participate at all, limiting decentralization. In principle the block gas limit is dynamic over time, as each "certificate of legitimacy" entitles its bearer to a one-time change in the block gas limit by a maximum of .05%. In practice, the block gas limit is fairly static and unresponsive to short term changes in demand–see Figure 3 for the change in block gas limit and compare it with the substantial price fluctuations of gas in Figure 2. The short run price fluctuations appear thus to be almost completely reflective of changes

²As of May, 2022 Ethereum miners continue to compete using this "Proof of Work" system, wherein energy resources are spent on computation in order to find an input which meets certain conditions. A bit like brute-force-guessing a simple password that changes every block, with the first correct guess unlocking the ability to add a single block.
3.2.2 NFTs as Digital Objects

An NFT is a provably unique digital object on a blockchain. NFTs are linked to different types of media, including videos, text, animated GIFs, and audio, but currently the most popular NFTs are images. NFTs on Ethereum largely conform to one of two token-standards, ERC-721 or ERC-1155, each of which can ensure an NFT is provably unique on Ethereum. The database-level uniqueness of an NFT depends on the standard to which it accords: ERC-721 NFTs are unique in the sense that each token within a given contract address must be unique. More specifically, a Token Identifier (typically an unsigned integer) which corresponds to a particular contract address (a 42 character hexadecimal that identifies some executable code) cannot be repeated. In ERC-1155 an NFT is a token ID that is assigned the integer "1", which means there is only one of them. These NFTs can be owned by Ethereum accounts, such as those belonging to individuals, so that NFTs can be easily bought, sold, and transferred.

Beyond the uniqueness of the TokenID and Contract Address, NFTs usually specify their correspondence to a piece of media, or some set of aesthetic attributes. For instance an NFT can
correspond to a picture stored as a jpeg file, but due to the high storage cost on Ethereum, the NFT will rarely contain the data to reproduce the actual image. Instead it will typically "point" to the image file in some way, such as having some of the NFT’s data contain a hash of the jpeg it corresponds to.\(^3\) While most NFTs reference or point to a file that exists off the blockchain, some NFTs are created to be completely "on-chain", such that the aesthetic attributes of the NFT can be fully recovered data on the Ethereum Blockchain itself without the need to refer to an external source. An example of this is found in the project Autoglyphs, in which the instructions for creating the aesthetic component of the NFT is found in the contract and the token id itself.

3.2.3 NFTs as Property

The NFT’s existing on a blockchain helps establish "answers to such questions as who owns, previously owned, and created the NFT, as well as which of the many copies is the original."[95] In light of this, a useful perspective on what NFTs are and why they might be valuable is provided by Property Rights Theory. Property Rights Theory directs our attention to the concept of ownership, and the way that ownership is enforced. Ownership is the possession of some set of rights over a resource, most particularly (1) the right to exclude others from the resource ("exclusivity") and (2) the right sell or transfer ownership of the resource ("alienability")[96]. These rights may be enforced by a third party, such as legal enforcement by police officers acting as agents of a court, but they may also emerge as a result of non-legal cooperative action.

Ownership over NFTs is not legally enforced, but in practice, through the cooperative action of Ethereum nodes\(^4\), anyone who possesses an NFT has the rights of exclusivity and alienability over it. That is, an NFT is owned by one address and cannot be claimed by other addresses without some action from the owning address, and once that account has transferred an NFT, it is gone unless it gets sent back.

Owning an NFT on Ethereum gives rights of exclusivity and alienability to the account which

---

\(^3\)This hash can be part of the address for a data store which uses content addressing, such as IPFS, where the address to view the file uses the hash itself. Usually this will be in the token’s metadata, specifically the TokenURI.

\(^4\)Recall that validating "nodes" (computers running Ethereum software) ultimately determine what is and is not part of the Ethereum Blockchain.
owns it. But, infamously, owners do not get property rights over their NFT’s corresponding aesthetic qualities.[97] See figure 1 for an example of an Autoglyph NFT and its corresponding visual identity and note that displaying the picture in Figure 1 did not require the author to own the Autoglyph 343 NFT. While this is often a source of confusion for people first encountering NFTS, this aspect of ownership has been long understood in property rights theory. In Barzel (1997), it's recognized that some attributes of a privately owned object are in the public domain. Adapting one of Barzel’s examples, suppose that you own a sandwich and I take a bite without permission. In that case I have violated your property rights. But if I look at the sandwich, or even photograph the sandwich and hang the photo on my wall, I likely have not violated your property rights. Thus, from Barzel’s perspective, the visual attributes of a displayed sandwich are in the public domain. NFT owners seem content to freely display the aesthetic qualities of their NFT, allowing anyone to view them, download them, etc. for free.

If not the aesthetic qualities of the NFT, what, then is owned? The unique object on a blockchain is owned, but note that to is easily possible to create a new, provably unique, NFT with the same aesthetic qualities as an existing NFT. That is, anyone is free to "mint" an NFT with the visual representation of Autoglyph 343. And while this has sometimes been done to scam buyers who believe they are getting the "real" NFT, this has not been viable way to create a valuable NFT. The "original" NFT, the one that holds value, is the one that corresponds to the NFT creator’s contract.

It is the blockchain’s ability to help answer "who owns, previously owned, and created the NFT, as well as which of the many copies is the original."[95] that makes an NFT valuable. This is a sort of authentication, particularly the provision of "indexical authenticity", so-named because the object indexes an authentic creation event (e.g. an indexically authentic Picasso is one that he actually created.)[98] The history offered on a blockchain can authenticate NFTs, providing information about whether the NFTs "are what they appear to be or are claimed to be, and therefore worth the price that is asked for them"[99]. That one NFT is authentic and another is not means they are differentiated, potentially serving as the basis of value creation.
3.2.4 The NFT Market

Though NFTs have existed on Ethereum since at least 2017\(^5\) they have exploded in popularity in 2021. According to a report from NonFungible.com, the value of NFT transactions grew to "more than $17 billion in 2021" a 21,350% increase over the $82.5 million in transactions seen observed in 2020. Despite the boom in their popularity, NFT research is "still limited, and focuses mostly on technical aspects."[95]

Most NFT sales occur on the website Opensea, selling as fixed price listings or auctions. Popular NFTs often belong to a "collection", which is a series of similar-but-not-identical NFTs by the same artist.\(^6\) NFTs within these collections sometimes have identifiable dimensions along which they will differ, referred to as "traits", and an NFT’s rarity can be calculated according to these traits. Traits are typically assigned during the "minting" process, which is when a user interacts with the NFT’s smart contract and creates an NFT which they then own. Thus during a mint, a

---

\(^5\)The first NFT token standard on Ethereum, ERC-721, was proposed in 2017. But several notable NFT projects precede that standard including for instance Cryptopunks, Digital Zones, and Moon Cats.

\(^6\)For historical reasons, the number of NFTs in a collection tends to be around 10,000.
specific NFT is created and transferred to the user. Because the minting process typically assigns traits randomly, specific aspects of an NFT’s appearance and rarity is not known to the minter until the NFT is created. Once the limit of available NFTs to be minted (usually 10,000 or so) has been reached, the only way to acquire an NFT from that collection is by purchasing it from an exchange, the most popular of which is Opensea.

3.3 The causal effect of volume on price: evidence from NFT Data

This research focuses on a unique aspect of NFT markets, which is the substantial attention to collection volume. The first fact to be established is whether volume causally affects price in NFT markets. To evaluate this, I scraped daily transaction and price data for the 963 most popular NFT collections over the past year. This was compiled into a panel of 108,778 collection x day observations. The data was then enhanced by daily gas prices taken from etherscan.

Transaction-level data was analyzed for cases in which a particular token was traded more than three times in one day. This technique is used by NFT data analysts to weed out fictitious transactions made for the "purpose of artificially inflating the trade volume of a given NFT". Values for a collection’s daily volume and average price were calculated both including and excluding these likely fictitious transactions.

Both volume and gas prices were observed to be non-stationary using a Dickey-Fuller test, making them suitable for regression analysis. The average number of daily transactions for a given collection is 71, .007% of the total number of daily Ethereum transactions.

3.3.1 Identification

The relationship between price and volume is fraught identification issues; for instance, since price plausibly affects volume a standard regression cannot rule out reverse causality[100]. We thus exploit a unique feature of the Ethereum Blockchain in order to instrument volume and determine its effect on price. Ethereum’s transaction fee ("gas price") changes dynamically according to

---

7The reasoning is that two transfers of an NFT could be the sign of a mistake or a sale one regretted and then bought back. But a third daily transfer is difficult to justify outside of volume manipulation.
demand for computation across the Ethereum Network. Supply also plays a role, but total supply and the rate at which supply can increase is hard-capped for security reasons.

We consider NFT traders to be price takers\(^8\) with respect to gas fees, disincentivized to trade when gas fees are especially high. Accordingly, a shock that increases gas prices will cause traders to reduce their likelihood to transact, thereby reducing volume. However, the value of an NFT is not likely to be directly tied to short-run Ethereum gas fees in any meaningful way. Thus gas fees make for a reasonable instrumental variable which affects price through the channel of transaction volume.

Table 3.1: The Effect of Volume on Price in NFT Markets

<table>
<thead>
<tr>
<th></th>
<th>(1) OLS</th>
<th>(2) IV</th>
<th>(3) IVSS</th>
<th>(4) OLS</th>
<th>(5) IV</th>
<th>(6) IVSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume (USD)</td>
<td>0.00201**</td>
<td>0.103**</td>
<td>0.154</td>
<td>0.00202**</td>
<td>0.0222**</td>
<td>0.0374***</td>
</tr>
<tr>
<td></td>
<td>(0.000845)</td>
<td>(0.0461)</td>
<td>(0.137)</td>
<td>(0.000846)</td>
<td>(0.00929)</td>
<td>(0.0105)</td>
</tr>
<tr>
<td>Age</td>
<td>10.88*</td>
<td>11.69*</td>
<td>47.62***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(6.126)</td>
<td>(5.967)</td>
<td>(9.245)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cragg-Donald</td>
<td>YES</td>
<td>YES</td>
<td>31.22</td>
<td>49.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contract FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td></td>
</tr>
</tbody>
</table>

Robust standard errors clustered at contract level in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 2.1 provides results from the data analysis. A simple OLS regression (column 1) observes a significant positive correlation between volume and price and in column 2, instrumenting gas prices for volume, a positive causal effect is evident (Cragg-Donald F = 11.61). We will return to column 3, which implements a shift-share design, below. Proceeding onto the columns 4 and 5, I add a control for the NFT project’s age, since other analyses of NFT prices have found project age to be strongly predictive. [101] Adding age to the OLS and IV regressions produces similarly positive and significant results.

\(^8\)While it is true that an NFT’s daily trading volume is itself a factor in demand for computation, this is likely to be a negligible effect. Recall that each NFT collection was responsible for an average of .007% of the daily transactions on the network, and, moreover, this is only a marginal effect since these NFT transactions are largely taking the place of other competing transactions.
Columns 3 and 6 implement a shift-share design to account for plausibly differential impacts of gas price shocks. The well-studied shift share design of Autor et al. [102] uses a shock to import competition thought to have differential effects based on the industrial composition of US regions. While there is no clear analogue for region industrial composition in the NFT data, there is the notion that gas prices would affect NFT projects more severely if they had many transactions. Consider that there are some NFT projects, such as Tom Sach’s Rocket Factory, in which the goal is to obtain several matching NFTs and combine them together.[103] Such a project requires multiple NFT transactions of its users, whereas another NFT project may not. To proxy this effect, we treat the total number of NFT transactions as a "share" in a weighted IV regression. Column three finds that the coefficient on this regression is positive but insignificant and the first-stage is weak (Cragg Donald F < 10). However, adding the control for age makes the regression positive and significant, with a non-weak instrument. This may be because the shift-share design overweights older projects, which have many transactions simply by virtue of existing for a long time. Controlling for age eliminates this effect.

3.4 Experiment on NFTs and Social Influence

Having established that volume has a causal effect on price in NFT markets, we now move to explore two possible mechanisms. Lacking access to the motivations of NFTs traders, we conduct a laboratory experiment to test the impact of information about volume on (1) stated NFT preferences and (2) observed NFT attention.

For the laboratory experiment, 428 english fluent subjects were recruited via the website prolific. Subjects were asked to rate a series of 54 NFT artworks from different collections, all created on the platform Artblocks. The 54 Artblocks NFTs are displayed to subjects in order of total volume, with the highest volume NFTs shown first.

To explain the conditions, subjects were told that Artworks "are sometimes transferred from owner to another. For example, when an artwork sells it is transferred to the buyer. There is a cost
to transfer an artwork, and that cost is sometimes very low and other times very high. Subjects were then told one of the following three things, depending on the condition they were (randomly) assigned to:

- **Baseline Condition**: "The following artworks are presented in order according to the artist's name."

- **Volume Condition**: "The following artworks are presented in order according the number of times that art from that collection has been transferred from one owner to another."

- **Costly Volume**: "The following artworks are presented in order according to how much money has been paid to transfer artworks in that collection from one owner to another."

After being shown the information for their condition, each subject was asked a comprehension question before beginning the ratings and again after the ratings, to ensure that they understood the order of artworks they were shown. Around a quarter of the respondents failed the comprehension questions, reducing the effective sample size to 320. Subjects were then presented with the each of the 54 NFTs and asked to them on a scale of 0-100, with 0 indicating "I like this artwork not at all" and 100 indicating "I like this artwork very much" (scale from [104]). The time subjects took to assess each artwork was also recorded.

### 3.4.1 Explanation of Experimental Conditions and Methods

Recall that all subjects, regardless of condition, were presented the NFT in the same order. The order was such that the NFT with the highest total volume came first, with volume descending monotonically until the 54th NFT, with the lowest total volume, was displayed last. This was done to hold constant any ordering effects that might arise endogenously from aspects of the NFTs themselves (e.g. whether subjects might like a large picture after being shown a small one.) Thus the only change across conditions is what subjects were told about the ordering.

---

9This is phrased so neutrally because in the NFT market sometimes transfers are not sales.
Figure 3.5: A screenshot of question 1, featuring Fidenza #560 by Tyler Hobbs
The control condition of "artist name" was chosen as a neutral ordering so that subjects in all conditions were told that the artworks were presented "in order" of something. No further information about how author’s names were ordered was provided. The "credible volume" condition is added because low cost volume may be less credible than high cost volume.

Artblocks was chosen in part because it is selective of NFT art and it favors abstract, "on-chain" NFT art\(^{10}\). The abstract nature of Artblocks NFTs reduces potential confounds that may exist among other NFT projects, where different NFTs could potentially have identical traits (e.g. NFT project Cryptopunks and copycat NFT project Fast Food Punks look very similar). The fact that Artblocks curates NFT artists according to strict criteria helps ensure some consistency and minimal quality for these NFTs. All 140 Artblocks NFT Collections were sorted according to volume, then I assessed collections by-hand looking for those which had no visual component and/or made use of non-pictorial media (e.g. a game or a video). This was done to make each NFT easily comparable, differing only visually. Of these 140 collections, 54 were judged to be primarily or entirely visual.

3.4.2 Experimental Results

The study was designed to test whether information about the volume of transactions of a digital artwork affects people’s preferences and/or attention regarding that artwork. I begin by standardizing the response variables (reported preferences and observed attention) using individual z scores. Attention is operationalized as the amount of time subjects spent before providing their rating, and ratings are their reported ratings on a scale of 1-100.

For volume and credible volume I use the order of the questions within each condition, from 1-54. Because the ordering was consistent across conditions, each number also corresponds to a unique artwork—1 denotes the first presented, and highest volume, artwork across all conditions, and 54 denotes the last presented, and lowest volume, artwork. For clarity, the variables "Volume" \(^{10}\)On-chain art is NFT art that can be recreated using only information stored on the blockchain. This contrasts with content-addressed NFTs which may include a link to the media and a hash of that media, but with the media needed to be stored off-chain.
and "Credible Volume" are inverted. This means that the highest volume NFT in treatment 1 has a value of "-1" for Volume, and the lowest ranked "-54". This makes interpreting the coefficients more straightforward (e.g. a positive coefficient on Volume when regressed on "Ratings" would mean that as perceived volume increases, reported ratings increase.)

Table 3.2: Results from Treatment 1 (Volume)

<table>
<thead>
<tr>
<th>Ind. Var.</th>
<th>(1) Ratings</th>
<th>(2) Ratings</th>
<th>(3) Attention</th>
<th>(4) Attention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume</td>
<td>0.000262</td>
<td>0.0000628</td>
<td>0.00371***</td>
<td>0.00374***</td>
</tr>
<tr>
<td></td>
<td>(0.50)</td>
<td>(0.12)</td>
<td>(7.12)</td>
<td>(7.07)</td>
</tr>
<tr>
<td>Attention</td>
<td>0.0815***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(6.87)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attn×Vol</td>
<td>-0.0000541</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.10)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ratings</td>
<td></td>
<td>0.0826***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(6.87)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rate×Vol</td>
<td></td>
<td>0.0000682</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.00000599</td>
<td>-0.00000987</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(-0.01)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>0.000168</td>
<td>-0.00165</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(-0.09)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.00379</td>
<td>0.000386</td>
<td>0.0536***</td>
<td>0.0579</td>
</tr>
<tr>
<td></td>
<td>(0.32)</td>
<td>(0.01)</td>
<td>(4.49)</td>
<td>(1.39)</td>
</tr>
<tr>
<td>Observations</td>
<td>11369</td>
<td>10992</td>
<td>11394</td>
<td>10992</td>
</tr>
</tbody>
</table>

* t statistics in parentheses
Standard errors are robust
* p < 0.05, ** p < 0.01, *** p < 0.001

In table 2, you can see regression coefficients for the effects of volume on ratings (columns 1 and 2) and attention (columns 3 and 4). Focusing first on Ratings, there is no significant effect of volume on ratings, with or without controls. Moving on to Attention, the regression coefficients are positive and significant for the basic OLS, as well as adding controls for demographic effects,
ratings, and the interaction between volume and attention. Informing subjects that NFTs are high volume causes a positive and significant increase in the attention they pay.

**Result 1:** Individuals to pay more attention to digital artworks that have been transacted more often.

Table 3.3: Results from Treatment 2 (Credible Volume)

<table>
<thead>
<tr>
<th>Ind. Var.</th>
<th>(1) Ratings</th>
<th>(2) Ratings</th>
<th>(3) Attention</th>
<th>(4) Attention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Credible Volume</td>
<td>0.00153**</td>
<td>0.00140**</td>
<td>0.00305***</td>
<td>0.00291***</td>
</tr>
<tr>
<td></td>
<td>(2.91)</td>
<td>(2.60)</td>
<td>(5.83)</td>
<td>(5.43)</td>
</tr>
<tr>
<td>Attention</td>
<td>0.0766***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(6.42)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attn×Vol</td>
<td>-0.00121*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-2.16)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ratings</td>
<td></td>
<td></td>
<td>0.0756***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(6.32)</td>
<td></td>
</tr>
<tr>
<td>Rate×Vol</td>
<td></td>
<td></td>
<td>-0.00124*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(-2.22)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.000225</td>
<td>-0.000467</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.29)</td>
<td>(-0.60)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>-0.000106</td>
<td>-0.00131</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.01)</td>
<td>(-0.07)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.0220</td>
<td>0.0296</td>
<td>0.0438***</td>
<td>0.0637</td>
</tr>
<tr>
<td></td>
<td>(1.83)</td>
<td>(0.74)</td>
<td>(3.66)</td>
<td>(1.58)</td>
</tr>
<tr>
<td>Observations</td>
<td>11261</td>
<td>10884</td>
<td>11286</td>
<td>10884</td>
</tr>
</tbody>
</table>

* t statistics in parentheses
  Standard errors are robust
  * p < 0.05, ** p < 0.01, *** p < 0.001

Table 3 reports results from a regression of **credible volume** on ratings (columns 1 and 2) and attention (columns 3 and 4). Here there is a significant and positive effect of volume on ratings, robust to controls. The effect of credible volume on attention is positive and significant as well.

**Result 3:** Individuals give higher ratings to digital artworks that have more costly volume.
Result 4: Individuals to pay more attention to digital artworks that have more costly volume (more accumulated transaction costs.)

The interaction terms Rate×Vol and Attn×Vol, are negative and significant. These effects can be interpreted as higher volume being associated with lower ratings when attention is high and higher volume being associated with lower attention when ratings are high. That means, for example, that there is a "delay" before a subject gives a low rating to an NFT with high credible volume. This delay could suggest cognitive dissonance when ones ratings disagree with the crowd, which would strengthen the interpretation of normative social influence. That is, someone who is rating a "high costly volume" artwork poorly may be inclined to take longer since they are aware that their decision is counternormative. And on the other end of things, a high rating being given to a high credible volume artwork takes less time. So subjects perhaps feel more secure and spend less time when their decision seems to be sanctioned by the group.

(Unhypothesized) Result 4: Individuals take more time when they give low (high) ratings to artworks with high (low) costly volume.

3.5 Conclusion

Volume is focal in NFT markets, with most if not all prominent NFT rankings sites utilizing volume as a default ranking. Rankings are thought to impact the reputations of organizations which, in turn, allow them to command a price premium.[84] I first examined whether higher volume led to higher prices, plausibly a reputational effect. Using exogenous shocks to the cost of transacting, I established that volume has a causal effect on prices in NFT markets.

Secondly I sought to better understand the mechanisms through which volume might affect price. Given the prominence in rankings, it seems that volume functions as an effective social test [86, 83]. Drawing on literature from behavioral finance, I first considered the hypothesis that volume has a first order affect on attention. This was verified in the experimental results.
I also considered that volume could be a channel through which normative social influence could affect behavior. In particular, do subjects rate art more highly if it has higher volume? It seems that mere volume is not enough to have a significant effect, but "credible" volume—volume sorted by transaction fees—does exhibit an effect. Lastly, I find that subjects pay additional attention before giving a bad rating to a "popular" NFT (and vice versa.) This delay could suggest cognitive dissonance when one’s ratings disagree with the crowd’s, which would strengthen the interpretation of normative social influence.
References


[49] G. Tabellini, “Presidential address: Institutions and culture,”


[51] B. Greiner, “An online recruitment system for economic experiments,”


[53] T. Ellingsen and M. Johannesson, “Promises, threats and fairness,”


[56] L. Hao and D. Houser, “Perceptions, intentions, and cheating,”


