Strategy-Proofness, Investment Efficiency, and Marginal Returns: An Equivalence

John William Hatfield  
McCombs School of Business  
University of Texas at Austin

Fuhito Kojima  
Department of Economics  
Stanford University

Scott Duke Kominers  
Society of Fellows  
Harvard University

April 9, 2017

Abstract

Classical insights from mechanism design imply that ex post efficient mechanisms induce agents to make efficient ex ante investment choices only if they are strategy-proof. For mechanisms that fail to be exactly strategy-proof and/or efficient, we derive a correspondence between the degree of failure of strategy-proofness and/or efficiency and the degree of failure to induce efficient investment. Our results extend to settings with uncertainty. Our results imply both that the worker-optimal stable mechanism incentivizes workers to make efficient human capital investments before entering the labor market, and that uniform-price and double auctions induce approximately efficient investment in large markets.

JEL Classification: C78, D44, D47, D82

Keywords: Strategy-proofness, Investment efficiency, Providing marginal rewards, Vickrey-Clarke-Groves mechanisms, Mechanism design

*An extended abstract of this paper appeared in the Proceedings of the 15th ACM Conference on Electronic Commerce. The authors thank Nageeb Ali, Susan Athey, Eduardo Azevedo, Gary S. Becker, James Best, Jörn Boehnke, Lars Böerner, Eric Budish, Gabriel Carroll, Pierre-André Chiappori, Vincent P. Crawford, Alexander Frankel, William Fuchs, Drew Fudenberg, Joshua Gans, Pingyang Gao, Edward L. Glaeser, Ed Green, Guillaume Haeringer, Lars Peter Hansen, Wei He, James J. Heckman, John Eric Humphries, Paul Klemperer, Edward Lazear, David McAdams, Andrew K. McKenzie, Paul R. Milgrom, Kevin M. Murphy, Roger Myerson, Yusuke Narita, Shunya Noda, Mallesh M. Pai, Andrew Postlewaite, Philip Reny, Alvin E. Roth, Tim Roughgarden, Ran Shorrier, Karl Schurter, Ilya Segal, Ron Siegel, Bradly Stadie, Bruno Strulovici, Yeneng Sun, Balázs Szentes, Satoru Takahashi, Alexander Teytelboym, Juuso Toikka, Kentaro Tomoeda, John Vickers, E. Glen Weyl, a referee, and numerous seminar audiences for helpful comments, and especially thank Larry Samuelson for a fantastic discussion. Christina Chiu, Jinjae Park, and Fanqi Shi provided excellent research assistance. Kojima gratefully acknowledges the support of a Sloan Research Fellowship, as well as the National Research Foundation through its Global Research Network Grant (NRF-2013S1A2A2035408). Kominers gratefully acknowledges the hospitality of Microsoft Research New England, as well as the support of National Science Foundation grants CCF-1216095 and SES-1459912, the Harvard Milton Fund, the Ng Fund of the Harvard Center of Mathematical Sciences and Applications, the American Mathematical Society, and the Simons Foundation. Any comments or suggestions are welcome and may be emailed to john.hatfield@utexas.edu, fkojima@stanford.edu, and kominers@fas.harvard.edu.
1 Introduction

Progress in the study of market design has recently enabled economists not only to analyze existing allocation mechanisms but also to design better ones. Auctions have become essential in a variety of markets, including those for spectrum licenses, electricity, search advertising, and U.S. Treasury bonds; moreover, many auction theorists have contributed to the design of auction formats to improve efficiency and revenue (Milgrom, 2004). Meanwhile, matching theory has been instrumental in the design of labor markets. For example, in the labor market for medical residents and hospital residency programs, the National Resident Matching Program (NRMP) matches residents to residency programs via an algorithm designed by Roth and Peranson (1999). Moreover, several school districts in the United States have collaborated with economists to design school choice mechanisms that have desirable properties in terms of incentives, efficiency, and fairness (Balinski and Sönmez, 1999; Abdulkadiroğlu and Sönmez, 2003; Abdulkadiroğlu, Pathak, Roth, and Sönmez, 2005; Abdulkadiroğlu, Pathak, and Roth, 2005, 2009).

Market designers have successfully designed mechanisms that achieve desirable properties such as strategy-proofness (i.e., incentive compatibility) and efficiency at the market-clearing stage. However, agents typically make some decisions before participating in market mechanisms. Prior to spectrum auctions, telecom companies must make investments in technologies that enable them to make use of the spectrum licenses that are to be auctioned off. Prior to bidding for manufacturing contracts, component manufacturers must invest in machinery, and assemblers must invest in process development. Prior to medical residency matching, medical students spend a great deal of time and effort studying, while hospitals use resources to develop residency programs for prospective residents. Investments comprise an important part of market participants’ decisions, and they can have large welfare consequences. Moreover, investment decisions are endogenous, in the sense that the returns to investments are partially determined by the market mechanism; hence, the choice of the market mechanism affects agents’ investment incentives. In other words, a mechanism can affect the participants’
behavior in the investment stage even before the mechanism is actually run.

The goal of this paper is to characterize the mechanisms that provide strong incentives for efficient *ex ante* investment. We consider a general public choice setting that includes, but is not limited to, object allocation settings such as auctions and two-sided matching environments such as labor markets. In our setting, the payoff for an agent may depend not only on the outcome generated by the mechanism but also from pre-mechanism investment by that agent. We say that a mechanism *induces efficient investment* by an agent if the socially efficient investment choices for that agent maximize the agent’s utility. A key insight from the work of Rogerson (1992) is that a mechanism induces efficient investment for an agent if it *provides marginal rewards* to that agent, i.e., the agent exactly internalizes the social gains or losses from a change in his valuation over outcomes. This insight, combined with the celebrated Green-Laffont-Holmström Theorem, leads to a characterization result for *ex post* efficient mechanisms: for such mechanisms, inducing efficient investment and provision of marginal rewards are each equivalent to strategy-proofness. In particular, an *ex post* efficient mechanism induces efficient investment by a particular agent if and only if that mechanism is strategy-proof for that agent.

Thus we see that for *ex post* efficient mechanisms, strategy-proofness—a concept that imposes no direct restrictions on investment behavior—has strong implications for investment incentives. Any mechanism that is *ex post* efficient and strategy-proof automatically induces participants to make efficient investments before participating in the mechanism. For many environments, however, it is impossible to design mechanisms that are both strategy-proof and efficient; nevertheless, in those environments it is often possible to bound the degree to which mechanisms fail to be strategy-proof and/or *ex post* efficient. For instance, in the recent U.S. spectrum reassembly “incentive” auction, a version of the *deferred acceptance auction* introduced by Milgrom and Segal (2015) was used. The deferred acceptance auction is strategy-proof but not fully efficient; nevertheless, in practice, the degree of inefficiency
appears to be small.\footnote{For instance, in many practical settings, the set of permissible outcomes roughly corresponds to those satisfying the constraints of a knapsack problem (Milgrom and Segal, 2015). In such settings, the degree of inefficiency of computationally tractable optimization algorithms can be bounded by the maximal possible value of adding one item to the knapsack.}

To address mechanisms which are not fully strategy-proof or efficient, we consider how the failure of a mechanism to be strategy-proof and/or \textit{ex post} efficient reduces incentives for efficient \textit{ex ante} investment. We show that a mechanism induces efficient investment within $\epsilon$ if and only if it provides marginal rewards within $\epsilon$. Moreover, for a mechanism that is efficient within $\eta$, if that mechanism induces efficient investment within $\epsilon$ then it is strategy-proof within $\epsilon + \eta$. Surprisingly, the converse does not hold: for a mechanism that is efficient within $\eta$, even if that mechanism is strategy-proof within $\epsilon$, it does not necessarily induce efficient investment within $\epsilon + \eta$; rather, it induces efficient investment within $\epsilon + \eta$ times a factor that depends on the set of possible alternatives.

We further extend our results to settings with imperfect information, in which agents may have uncertainty about how investments impact valuations; we also allow for the possibility of residual (post-investment) uncertainty about valuations that is not resolved until after an outcome is chosen. We show that our results continue to hold in the presence of uncertainty. For example, we find a close link between strategy-proofness and investment efficiency in private-values settings where agents can invest in information acquisition before participating in an auction; this generalizes the main result of Bergemann and Välimäki (2002) for private-values auctions.

Efficiency and strategy-proofness are already key goals of market design; our results imply that pursuing these goals also helps to align \textit{ex ante} investment incentives. While the earliest form of this insight follows from the work of Rogerson (1992) and others, our results show that this insight is robust and applies in many contexts. In particular, so long as a mechanism is nearly efficient and nearly strategy-proof, then the private and social returns from \textit{ex ante} investment are nearly aligned.

Finally, we present several applications. We first consider the labor market matching
setting of Kelso and Crawford (1982). We use our results to show that the worker-optimal stable mechanism is strategy-proof for workers and induces efficient investment by those workers in the presence of continuous transfers. We then use our results to bound the inefficiency in workers’ investments that arise under the worker-optimal stable mechanism when transfers are discrete.\(^2\) We then apply our results in the procurement auction setting of Kothari, Parkes, and Suri (2003, 2005), showing that the computationally tractable, approximately strategy-proof, and approximately efficient Kothari et al. (2003, 2005) mechanism approximately induces efficient investment by sellers. As another application, we consider uniform-price auctions as described by Friedman (1960, 1991). We show that uniform-price auctions induce approximately efficient investment when many bidders are present, building on the well-known intuition that, when many bidders are present, the returns to manipulation in uniform-price auctions are small. Finally, we show that the McAfee (1992) double auction—which is exactly strategy-proof and approximately efficient in large markets—induces approximately efficient investment in large markets.\(^3\)

**Related Literature**

Our analysis builds upon a number of papers in mechanism design that study investment incentives. In a groundbreaking paper, Rogerson (1992) studied investment incentives under various informational assumptions, and offered mechanisms that induce efficient investment while satisfying efficiency and incentive compatibility conditions. Unlike Rogerson (1992), who studies exactly efficient and strategy-proof mechanisms, we consider the relationship

\(^2\)In many settings, transfers are highly discretized to simplify the preference formation and reporting process. For instance, Crawford (2008) proposed that the NRMP use a mechanism he called the Flexible-Salary Match, which allows residency programs to offer several (e.g., three) salary levels and then implements the resident-optimal stable matching.

\(^3\)In Appendix C, we show that the seller-optimal stable mechanism provides marginal rewards for unit-supply sellers in the trading network setting of Hatfield et al. (2013); our results then imply that the seller-optimal stable mechanism is both strategy-proof for unit-supply sellers and induces efficient investment by those sellers. In work in progress (Hatfield et al., 2016), we consider auctions with endogenous entry, using our results for settings with uncertainty to obtain both efficient and revenue-maximizing mechanisms, building on the work of McAfee and McMillan (1987), Engelbrecht-Wiggans (1993), and Jehiel and Lamy (2015).
between deviations from strategy-proofness, efficiency, and inducement of efficient investment. We also extend our results to settings with uncertainty, showing the linkages between the ideas of Rogerson (1992) and Bergemann and Välimäki (2002).

Others have studied incentives for pre-mechanism investment in specific contexts: For example, Arozamena and Cantillon (2004) studied firms’ incentives to make cost-reducing investments before a procurement auction. They showed that using a first-price auction results in inefficient underinvestment, while using a second-price auction induces efficient investment. As the second-price auction is both efficient and strategy-proof, the latter result of Arozamena and Cantillon (2004) can be seen as a special case of the equivalence result for efficient mechanisms. Makowski and Ostroy (1995) considered a model in which agents choose occupations and then engage in trade, where the costs and benefits of a good for an agent may depend on the occupation chosen. Makowski and Ostroy (1995) showed that when each agent fully appropriates his contribution to social welfare, it is an equilibrium for each agent to choose the socially efficient occupation—while their model can not be embedding within our framework (as Makowski and Ostroy (1995) consider an infinite alternative space), the logic underlying their result corresponds to the insight that efficient ex ante investment will be induced by the provision of marginal rewards. Our work is also related to the extensive literature on network formation. Kranton and Minehart (2001) considered a model of buyer-seller markets where unit-demand buyers must make an ex ante investment in order to “link” with a seller, thereby enabling transactions with that seller. Kranton and Minehart (2001) showed that when an ascending bid auction is used to allocate goods, an efficient network of links arises in some equilibria; this result can be seen as a special case of the results we present in Appendix C. As will be seen, results on efficient investment do not necessarily depend upon the specific environments considered in previous work; rather, they hinge upon the strategic and efficiency properties of the mechanisms used.

In contrast to our results, Cole, Mailath, and Postlewaite (2001a,b) found conditions for

---

4 See also the recent work of Felli and Roberts (2002), Chatterjee and Chiu (2007), and Elliott (2014).
efficient *ex ante* investment by *both* workers and firms in a two-sided one-to-one matching market, even absent an efficient mechanism that is strategy-proof for both sides of the market.\(^5\) As we discuss in our analysis of labor market matching, the results of Cole et al. (2001a,b) are not a special case of ours, as Cole et al. (2001a,b) mechanisms rely on the mechanism designer having knowledge of agents’ *ex ante* investment costs—which we do not assume.

There is also a large literature on pre-contractual investment in contract theory; this literature goes back to at least the work of Klein, Crawford, and Alchian (1978), who investigated the hold-up problem.\(^6\) Grossman and Hart (1986) argued that inefficient investments can occur when complete contracts cannot be written (and that this insight may help in understanding the boundaries of the firm); Hart and Moore (1990) extended the Grossman and Hart (1986) analysis to the case with more parties and a richer set of asset structures. Subsequent research sought to understand under which types of contracting environments we should expect investment inefficiency: Aghion et al. (1994) showed that investment efficiency can be recovered with “more complete” contracts in the sense of appropriate provisions regarding renegotiation. Che and Hausch (1999) showed that inefficiency can happen if an agent’s investment can affect other parties. Here we investigate a general mechanism design context, where agents make investments before taking part in a centralized mechanism. By focusing on mechanism design, we can determine how partial failures of mechanism properties are interrelated; for example, we can measure how the *ex post* efficiency of a mechanism determines its ability to provide good *ex ante* incentives for investment.

Finally, our work is related to that of Hatfield, Kojima, and Narita (2016) who, in the school choice setting of Balinski and Sönmez (1999) and Abdulkadiroğlu and Sönmez (2003), considered incentives for schools to improve their programs in order to attract preferred students. While sharing our motivation, the model of Hatfield, Kojima, and Narita (2016) is

---

5 *Ex ante* investment in two-sided labor markets is also discussed by Mailath, Postlewaite, and Samuelson (2013) and Nöldeke and Samuelson (2015).

6 See also the work by Coase (1960) and and Williamson (1979, 1983) for early important contributions on transaction costs and the hold-up problem.
different from ours in several respects. Most importantly, motivated by school choice, the Hatfield, Kojima, and Narita (2016) model does not allow monetary transfers. By contrast, our efficiency concept requires the maximization of social surplus, a concept which is compelling only with the existence of monetary transfers. Moreover, Hatfield, Kojima, and Narita (2016) did not explicitly consider the cost of investment, while the cost of investment is central to our analysis.

Organization of the Paper

The remainder of the paper is organized as follows: Section 2 introduces our framework. Section 3 presents equivalence results for exactly ex post efficient mechanisms. Section 4 introduces how we measure deviations from efficiency, strategy-proofness, provision of marginal rewards, and inducement of efficient investment and studies how partial failures of those properties are interrelated. Section 5 extends our results to incorporate multiple forms of uncertainty. Section 6 presents applications to labor markets, procurement auctions, uniform-price auctions, and double auctions. Section 7 concludes. All proofs are presented in Appendix D.

2 Model

2.1 Framework

There exists a finite set $I$ of agents and a finite set $\Omega$ of alternatives. Each agent $i \in I$ has a valuation function $v^i : \Omega \rightarrow \mathbb{R}$. (As we describe in the next section, agents’ valuation functions are determined endogenously, through ex ante investment.) The set of possible valuation functions for $i$ is denoted $V^i$. Since $\Omega$ is finite and the value for each alternative is simply a real number, the space of valuation functions is isomorphic to a subset of $\mathbb{R}^\Omega$. Throughout, we assume that $V^i$ is compact. The space of all valuation profiles is denoted by $V \equiv \times_{i \in I} V^i$. The (ex post) utility of agent $i \in I$ for an alternative $\omega \in \Omega$ and transfer vector
$t \in \mathbb{R}^I$ is given by $u^i((\omega, t); v^i) = v^i(\omega) - t^i$.\footnote{We use the modifier \textit{ex post} here to indicate that the utility does not include the cost of investment.} We call an alternative–transfer pair $(\omega, t)$ an outcome.

An allocation rule $\mu : V \to \Omega$ is a map from the space of valuation profiles to the set of alternatives. A transfer rule $r : V \to \mathbb{R}^I$ is a map from the space of valuation profiles to the set of transfer vectors. We assume that the transfer from each agent is bounded, that is, $\sup_{v \in V; i \in I} \{|r^i(v)|\} < \infty$.\footnote{Boundedness is a weak restriction, and appears to be reasonable in our mechanism design context. It is satisfied, for example, if the magnitude of a transfer never exceeds the maximum total social welfare.} A mechanism $\mathcal{M} \equiv (\mu, r)$ is an ordered pair consisting of an allocation rule $\mu$ and a transfer rule $r$; we denote $\mathcal{M}(v) \equiv (\mu(v), r(v))$. We focus on direct revelation mechanisms, that is, mechanisms that take agent types (i.e., valuation functions) as input.

We define the (social) welfare of an alternative $\omega$ to be the sum of the agents’ valuations for that alternative, i.e.,

$$V(\omega; v) \equiv \sum_{i \in I} v^i(\omega). \quad (1)$$

We abuse notation slightly by, for a mechanism $\mathcal{M} = (\mu, r)$, writing $v^i(\mathcal{M}(v)) \equiv v^i(\mu(v))$ and $V(\mathcal{M}(v); v) \equiv V(\mu(v); v)$.

Note that welfare is defined exclusively in terms of the surplus created by the mechanism; since mechanisms are not necessarily required to satisfy a balanced budget condition, it may be the case that the sum of the agents’ utilities does not equal total welfare.

\section*{2.2 Investment}

Before participating in the mechanism, each agent makes an investment decision that determines his valuation over alternatives. We model the choice of investment as an explicit choice of the valuation function $v^i$, with the cost of investment determined by a cost function $c^i : V^i \to \mathbb{R}$. The \textit{ex ante utility} of agent $i$ given an outcome–investment pair $((\omega, t), v)$ is
We define the \textit{ex ante social welfare} of an outcome–investment pair \((\omega, t), v\) as \(V(\omega; v) - \sum_{i \in I} c^i(v^i)\).

3 Characterization of Fully Efficient Mechanisms

In this section, we unify and organize classical ideas about the relationship between inducing efficient investment, providing marginal rewards, strategy-proofness, and \textit{ex post} efficiency. As we show in subsequent sections, the linkages we highlight here allow us to understand how partial failures in strategy-proofness and efficiency affect investment incentives.

3.1 Properties of Mechanisms

In this section, we introduce the mechanism properties that are central to our analysis. The first three—\textit{strategy-proofness}, \textit{provision of marginal rewards}, and \textit{inducement of efficient investment}—are connected through Lemma 1 and Theorem 1. We state these properties as agent-specific conditions, so that we may analyze settings (such as two-sided matching) where some properties (like strategy-proofness) hold for some agents but not for others.

Our definition of strategy-proofness is standard: A mechanism is strategy-proof for \(i \in I\) if reporting the true valuation \(v^i\) is a (weakly) dominant strategy for \(i\), i.e., truthful reporting is a best response for \(i\) for any valuation profile \(v^{I\setminus\{i\}}\) reported by other agents.

\textbf{Definition 1.} A mechanism \(M\) is \textit{strategy-proof} for \(i \in I\) if, for all \(v \in V\) and \(\bar{v}^i \in V^i\),

\[ u^i(M(v); v^i) \geq u^i(M(\bar{v}^i; v^{I\setminus\{i\}}); v^i). \]

A mechanism provides marginal rewards to \(i \in I\) if, assuming agents report truthfully, for every valuation profile \(v^{I\setminus\{i\}}\) of other agents, the change in the utility of \(i\) induced by a

\[ u^i((\omega, t); v^i) - c^i(v^i). \]

\footnote{We use the modifier \textit{ex ante} here to indicate that the utility includes the cost of investment taken before the mechanism is run.}
change in the valuation of $i$ is equal to the change in social welfare. That is, a mechanism provides marginal rewards to $i$ if $i$ captures the full marginal surplus from a change in his valuation.

**Definition 2.** A mechanism $\mathcal{M}$ provides marginal rewards to $i \in I$ if, for all $v \in V$ and $\bar{v}^i \in V^i$,

\[
u^i(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); \bar{v}^i) - \nu^i(\mathcal{M}(v); v^i) = V(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); (\bar{v}^i, v^{I\setminus\{i\}})) - V(\mathcal{M}(v); v).
\]

A mechanism induces efficient investment for $i \in I$ if, assuming that agents report truthfully, for every valuation profile $v^{I\setminus\{i\}}$ of other agents, and for every cost function $c^i$, the valuations that maximize the utility of $i$ are exactly those that maximize social welfare.\(^{10}\)

**Definition 3.** A mechanism $\mathcal{M}$ induces efficient investment by $i \in I$ if, for all $v^{I\setminus\{i\}} \in V^{I\setminus\{i\}}$,

\[
\arg \max_{\bar{v}^i \in V^i} \left\{ \nu^i(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); \bar{v}^i) - c^i(\bar{v}^i) \right\} = \arg \max_{\bar{v}^i \in V^i} \left\{ V(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); (\bar{v}^i, v^{I\setminus\{i\}})) - c^i(\bar{v}^i) \right\}
\]

(2)

for any cost function $c^i : V^i \to \mathbb{R}$.\(^{11}\)

For some of our results, we focus on (ex post) efficient mechanisms, i.e., mechanisms that maximize social welfare given the valuation profile.

\(^{10}\)Providing marginal rewards is equivalent to an earlier condition introduced by d’Aspremont and Gérard-Varet (1979) and Rogerson (1992) called (subjective) discretionarity (see further discussion in Section 3.2).

\(^{11}\)In Appendix A.1, we show that the following condition is equivalent to Definition 3:

For all $v^{I\setminus\{i\}} \in V^{I\setminus\{i\}}$ and cost functions $c^i$, for any sequence of valuations $\{\bar{v}^i\}_{t=1}^\infty$, the sequence $\{\nu^i(\mathcal{M}(\bar{v}^i_t, v^{I\setminus\{i\}}); \bar{v}^i_t) - c^i(\bar{v}^i_t)\}_{t=1}^\infty$ approaches $\sup_{\bar{v}^i \in V^i} \left\{ \nu^i(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); \bar{v}^i) - c^i(\bar{v}^i) \right\}$ if and only if the sequence $\{ V(\mathcal{M}(\bar{v}^i_t, v^{I\setminus\{i\}}); (\bar{v}^i_t, v^{I\setminus\{i\}})) - c^i(\bar{v}^i_t)\}_{t=1}^\infty$ approaches $\sup_{\bar{v}^i \in V^i} \left\{ V(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); (\bar{v}^i, v^{I\setminus\{i\}})) - c^i(\bar{v}^i) \right\}$.

Note that this alternate condition has bite even when the sets of maximizers on both sides of (2) are empty. Intuitively, for any mechanism that satisfies the condition of Definition 3, a valuation $\bar{v}^i$ nearly maximizes the utility of agent $i$ if and only if $\bar{v}^i$ nearly maximizes social welfare.
**Definition 4.** A mechanism $M$ is efficient if, for all $v \in V$,

$$V(M(v); v) = \max_{\omega \in \Omega} \{V(\omega; v)\}.$$ 

### 3.2 Results

If a mechanism grants an agent the full social returns to his investment, then that agent is naturally incentivized to make socially efficient investment decisions *ex ante*. Rogerson (1992) made this intuition formal, showing that any *ex post* efficient mechanism that provides marginal rewards (i.e., a Vickrey-Clarke-Groves mechanism) induces efficient investment, as under such mechanisms, “each agent receives the full marginal expected social product from his own investment” (p. 779). A similar insight arises in the work of Bergemann and Välimäki (2002), which shows that Vickrey-Clarke-Groves mechanisms induce efficient information acquisition *ex ante*. In fact, as we highlight here, the link between providing marginal rewards and inducing efficient investment does not depend on *ex post* efficiency; Rogerson’s insight can be extended to show that any mechanism that always rewards an agent with the social value his investment creates *conditional upon the alternative the mechanism chooses* will induce that agent to invest in a socially efficient manner. Moreover, as we also show, any mechanism that does not provide marginal rewards will fail to induce efficient investment.

The preceding observations lead to the following lemma.

**Lemma 1.** Consider a mechanism $M$ and an agent $i \in I$. The following two statements are equivalent:

1. The mechanism $M$ induces efficient investment by $i$.
2. The mechanism $M$ provides marginal rewards to $i$.

We provide a simple example of a second-price auction that demonstrates Lemma 1 in action in Appendix B.1.

12
Lemma 1 highlights that mechanisms that induce efficient investment are exactly those that provide marginal rewards. Meanwhile, within the class of ex post efficient mechanisms, it is well-known that providing marginal rewards is generally equivalent to strategy-proofness (Green and Laffont, 1977; Holmström, 1979; Carbajal and Tourky, 2009); combining this observation with Lemma 1 yields the following result.\(^{12}\)

**Theorem 1.** Consider an efficient mechanism \(\mathcal{M}\) and an agent \(i \in I\), and suppose that \(V^i\) is path-connected. The following three statements are equivalent:

1. The mechanism \(\mathcal{M}\) induces efficient investment by \(i\).
2. The mechanism \(\mathcal{M}\) provides marginal rewards to \(i\).
3. The mechanism \(\mathcal{M}\) is strategy-proof for \(i\).\(^{13}\)

Theorem 1 clarifies the relationship between strategy-proofness and inducing efficient investment. The key linkage—as we have already seen—is the provision of marginal rewards condition, which is equivalent to both inducing efficient investment and (for ex post efficient mechanisms) strategy-proofness. Of course, it is well-known that Vickrey-Clarke-Groves mechanisms are strategy-proof because they provide marginal rewards and are ex post efficient; moreover, as the Green–Laffont (1977) theorem indicates, Vickrey-Clarke-Groves mechanisms are the only strategy-proof mechanisms in the class of ex post efficient mechanisms.\(^{14}\) Thus, we find not only that Vickrey-Clarke-Groves mechanisms are fully efficient (à la Rogerson (1992)), but that that they are in fact the only fully efficient mechanisms.

Theorem 1 organizes a number of classical ideas about Vickrey-Clarke-Groves mechanisms and the relationships between strategy-proofness, providing marginal rewards, and inducing

\(^{12}\)As we clarify below, the observations codified in Theorem 1 are not novel in and of themselves. Rather, our intention is to formally and clearly state here the relationships between inducing efficient investment, providing marginal rewards, and strategy-proofness that we generalize in later sections.

\(^{13}\)The assumption that \(V^i\) is path-connected is only employed in showing that strategy-proofness implies the other two conditions.

\(^{14}\)If we were to assume that \(V^i\) is smoothly connected, then we could apply the Holmström (1979) version of the Green–Laffont theorem directly; Suijs (1996) and Heydenreich et al. (2009) generalize this result by relaxing the smooth-connectedness assumption on the space of valuations so as to cover the path-connected case.
efficient investment. Theorem 1 serves in part as a template for the more general results we prove in the sequel, as the linkage through providing marginal rewards exposes how to generalize the insights of Theorem 1 to mechanisms beyond Vickrey-Clarke-Groves. Thus, while Theorem 1 has many antecedents, we derive it here. The approach we develop generalizes to allow us to understand how much a failure of conditions such as strategy-proofness reduces incentives for efficient investment.

Theorem 1 is “sharp,” in the sense that all of the hypotheses are required for the full equivalence to hold. In particular, the requirement that the mechanism always chooses the efficient alternative from the set of alternatives is necessary in general; in Appendix B.2, we show that a posted-price mechanism is strategy-proof, but not efficient, and, in fact, does not provide marginal rewards. Similarly, in Appendix B.3, we consider a setting with a valuation space that is not path-connected and show that it admits an efficient and strategy-proof mechanism that does not provide marginal rewards.

Note that, even though efficiency is required by Theorem 1, in fact what is actually required is efficiency within a fixed space of alternatives. Thus, our formulation of Theorem 1 immediately allows us understand how constraining the outcome space may affect investment incentives. We say that a mechanism is constrained ex post efficient if it always chooses the socially efficient alternative from some fixed restriction of the alternative space $\bar{\Omega} \subseteq \Omega$. For instance, in many settings, there are legal diversity constraints in procurement (Fryer and Loury, 2013; Pai and Vohra, 2014), such as set-asides that require that a certain percentage of contracts must be awarded to small businesses or minority-owned firms. In our context, set-asides can be implemented by considering a reduced alternative space $\bar{\Omega}$; thus, Theorem 1 implies that a strategy-proof, constrained ex post efficient mechanism will induce efficient investment.

Using Theorem 1, we can analyze equilibrium outcomes of the game induced by a mechanism $\mathcal{M}$. Given cost functions $\{c^i\}_{i \in I}$, the game induced by $\mathcal{M}$ is a game between the agents in $I$ where each agent $i$ chooses a valuation $v^i \in V^i$ and receives payoff $u^i(\mathcal{M}(v); v^i)$ —
$c^i(v^i)$; agents’ cost functions are assumed to be common knowledge. Theorem 1 implies that, when each agent other than $i$ is choosing a socially optimal investment, it is privately optimal for $i$ to also choose a socially optimal level of investment; this yields a “fully efficient” Nash equilibrium.

**Corollary 1.** For any efficient and strategy-proof mechanism $\mathcal{M}$, there exists a Nash equilibrium $v$ of the game induced by $\mathcal{M}$ such that $v$ maximizes ex ante social welfare.

Unfortunately, it is not the case that every Nash equilibrium induced by an efficient and strategy-proof mechanism $\mathcal{M}$ maximizes ex ante social welfare.\(^{15}\) As we illustrate via example in Appendix B.4, inefficient investment outcomes can arise in an equilibrium because of miscoordination: Investment by one agent may preclude investment by the other agent, even if the latter agent can generate greater value—and the fact that the latter agent does not invest then makes it optimal for the former agent to invest.\(^{16}\) Nevertheless, Corollary 1 ensures that an efficient Nash equilibrium does exist and so, if players can coordinate prior to investment, efficiency should be achieved.

### 4 Bounding Departures from Full Efficiency

#### 4.1 Generalized Properties of Mechanisms

We now introduce analogues of the exact properties introduced in Section 3.1 that allow for bounded departures from strategy-proofness, providing marginal rewards, inducing efficient investment, and ex post efficiency.

To measure deviations from exact strategy-proofness, we say a mechanism is strategy-proof

---

\(^{15}\)Indeed, as Tomoeda (2015) has recently shown, it is not possible to implement only the fully efficient outcome in Nash equilibrium; however, Tomoeda (2015) shows that adding an ex post investment stage can sometimes make the fully efficient outcome the unique equilibrium outcome.

\(^{16}\)Makowski and Ostroy (1995) showed that miscoordination equilibria exist in their model of occupational choice as well. Quint and Einav (2005) (in the setting of firm entry into an oligopolistic product market) and McAdams (2014) (in the setting of a single-good second-price auction) considered the possibility that miscoordination may be mitigated by sequencing agents’ investment choices.
for $i$ within $\epsilon$ if reporting the truth results in an outcome for $i$ that provides a utility within $\epsilon$ of the highest utility $i$ can attain under any report.

**Definition 5.** A mechanism $M$ is *strategy-proof within $\epsilon$* for $i \in I$ if, for all $v \in V$ and $\tilde{v}^i \in V^i$,

$$u^i(M(v); v^i) + \epsilon \geq u^i(M(\tilde{v}^i, v^{I \setminus \{i\}}); v^i).$$

Similarly, a mechanism provides marginal rewards within $\epsilon$ if, assuming agents report truthfully, the change in the utility of $i$ induced by a change in the valuation of $i$ is within $\epsilon$ of the change in social welfare.

**Definition 6.** A mechanism $M$ provides marginal rewards within $\epsilon$ to $i \in I$ if, for all $v \in V$ and $\tilde{v}^i \in V^i$,

$$\left| (V(M(v); v) - V(M(\tilde{v}^i, v^{I \setminus \{i\}}); (\tilde{v}^i, v^{I \setminus \{i\}}))) - (u^i(M(v); v^i) - u^i(M(\tilde{v}^i, v^{I \setminus \{i\}}); \tilde{v}^i)) \right| \leq \epsilon.$$

A mechanism induces efficient investment within $\epsilon$ by $i$ if, assuming agents report truthfully, every utility-maximizing choice of valuation by $i$ maximizes social welfare within $\epsilon$.

**Definition 7.** A mechanism $M$ induces efficient investment within $\epsilon$ by $i \in I$, if, for all $v^{I \setminus \{i\}} \in V^{I \setminus \{i\}}$, if

$$\hat{v}^i \in \arg\max_{\tilde{v}^i \in V^i} \{u^i(M(\tilde{v}^i, v^{I \setminus \{i\}}); \tilde{v}^i) - c^i(\tilde{v}^i)\},$$

then we have

$$V(M(\hat{v}^i, v^{I \setminus \{i\}}); (\hat{v}^i, v^{I \setminus \{i\}})) - c^i(\hat{v}^i) + \epsilon \geq \sup_{\tilde{v}^i \in V^i} \left\{V(M(\tilde{v}^i, v^{I \setminus \{i\}}); (\tilde{v}^i, v^{I \setminus \{i\}})) - c^i(\tilde{v}^i)\right\}$$

for all cost functions $c^i : V^i \to \mathbb{R}$.\(^{17}\)

\(^{17}\)Note that this definition considers *exact* optimization on an agent’s part while requiring only maximization of social welfare *within* $\epsilon$; that is, Definition 7 requires that optimal investment choices for $i$ are socially efficient within $\epsilon$, rather than requiring socially optimal investment choices be optimal within $\epsilon$ for $i$. We chose
Finally, our definition of bounded efficiency loss is standard: A mechanism is efficient within $\eta$ if it selects an outcome that achieves social welfare within $\eta$ of the maximum with respect to the reported valuations.

**Definition 8.** A mechanism $\mathcal{M}$ is (ex post) efficient within $\eta$ if, for all $v \in V$,

$$V(\mathcal{M}(v); v) + \eta \geq \max_{\omega \in \Omega} \{V(\omega; v)\}.$$

### 4.2 Results

We begin by showing that a mechanism induces efficient investment within $\epsilon$ if and only if it provides marginal rewards within $\epsilon$.

**Lemma 2.** Consider a mechanism $\mathcal{M}$ and an agent $i \in I$. For any $\epsilon \geq 0$, the following two statements are equivalent:

1. The mechanism $\mathcal{M}$ induces efficient investment within $\epsilon$ by $i$.

2. The mechanism $\mathcal{M}$ provides marginal rewards within $\epsilon$ to $i$.

We show next that if a mechanism provides marginal rewards within $\epsilon$ to $i$ and is efficient within $\eta$, then the mechanism is strategy-proof within $\epsilon + \eta$.

**Theorem 2.** Consider a mechanism $\mathcal{M}$ and an agent $i \in I$. For any $\epsilon \geq 0$ and $\eta \geq 0$, if the mechanism $\mathcal{M}$ provides marginal rewards within $\epsilon$ to $i$ and is efficient within $\eta$, then $\mathcal{M}$ is strategy-proof within $\epsilon + \eta$ for $i$.

Given Theorem 1, we might conjecture a converse of Theorem 2, i.e., if the mechanism $\mathcal{M}$ is strategy-proof within $\epsilon$ for $i$ and efficient within $\eta$ (and $V^i$ is path-connected), then $\mathcal{M}$

---

Definition 7 as, from the mechanism designer’s point of view, it may be difficult to force the agent to only approximately optimize, while approximate social welfare maximization may still be desirable. Nevertheless, as a robustness check, we show in Appendix A.2 that an alternative definition requiring exact social welfare maximization and individual optimization within $\epsilon$ is equivalent to Definition 7. In particular, this equivalence implies that the inducement of efficient investment given in Definition 3 is a special case of Definition 7 for $\epsilon = 0$. We thank Philip Reny for suggesting this extension.
provides marginal rewards within $\epsilon + \eta$ to $i$ (and hence induces efficient investment within $\epsilon + \eta$ by $i$). The following example, however, shows that this conjecture is false, even for exactly efficient mechanisms.

**Example 1.** Suppose there are two agents, $i$ and $j$, and there are $N$ units of a homogeneous good to be allocated. The set of alternatives is then given by $\Omega \cong \{0, \ldots , N\}$ where $\omega = n$ represents the allocation of $n$ units to agent $i$ and $N - n$ units to agent $j$.\(^{18}\)

The valuation space $V^i$ of agent $i$ is parameterized by $a^i$, where $a^i \in [0, N + 1]$: The valuation function $v^i$ is given by

$$v^i(\omega; a^i) = \sum_{m=1}^{\omega} (a^i - m + 1);$$

that is, the first unit of the good is worth $a^i$ to $i$, the second unit is worth $a^i - 1$ to $i$, and so forth. This gives rise to the utility function

$$u^i((\omega, t); v^i(\cdot; a^i)) = \sum_{m=1}^{\omega} (a^i - m + 1) - t^i.$$

Agent $j$, by contrast, simply values each unit of the good at $1 + \eta$, with $\eta \geq 0$ a (small) constant; hence, the utility function of $j$ is fixed at

$$u^j((\omega, t); v^j(\cdot; a^j)) = (N - \omega)(1 + \eta) - t^j.$$

Consider the mechanism $\mathcal{M} = (\mu, r)$ where

$$\mu(v) = \left\lfloor a^i \right\rfloor$$

$$r^i(v) = \left\lfloor a^i \right\rfloor (1 - \epsilon)$$

$$r^j(v) = 0,$$

\(^{18}\)The setting of Example 1 assumes a lack of free disposal, but the argument does not in any substantive way depend on this assumption.
with $\epsilon \geq 0$ a (small) constant.

This mechanism $\mathcal{M}$ is strategy-proof within $\epsilon$ for each agent. To see this, first note that $\mathcal{M}$ is clearly strategy-proof for $j$, as $V^j$ is a singleton set. Thus, we consider agent $i$ with valuation $v^i(\cdot; a^i)$ such that $n \leq a^i < n + 1$ where $n \in \{0, \ldots, N + 1\}$. If agent $i$ reports his true valuation, he obtains $\lfloor a^i \rfloor$ units of the good and utility $\sum_{m=1}^{\lfloor a^i \rfloor} (a^i - m + \epsilon)$. If agent $i$ instead reports a valuation $\bar{v}^i(\cdot; \bar{a}^i)$ such that $n \leq \bar{a}^i < n + 1$, then $i$ obtains the same number of units while paying the same transfer. If agent $i$ instead reports a valuation $\bar{v}^i(\cdot; \bar{a}^i)$ such that $\bar{a}^i < n$, then $i$ obtains $\lfloor a^i \rfloor - \lfloor \bar{a}^i \rfloor$ fewer units of the good, while paying $((\lfloor a^i \rfloor - \lfloor \bar{a}^i \rfloor)(1 - \epsilon))$ less, making him weakly worse off. Finally, if agent $i$ instead reports a valuation $\bar{v}^i(\cdot; \bar{a}^i)$ such that $\bar{a}^i \geq n + 1$, then $i$ obtains $\lfloor \bar{a}^i \rfloor - \lfloor a^i \rfloor$ more units, while paying $((\lfloor \bar{a}^i \rfloor - \lfloor a^i \rfloor)(1 - \epsilon))$ more; this can increase his utility by at most $\epsilon$.

Furthermore, $\mathcal{M}$ is also efficient within $\eta$ when $\eta$ is small, i.e., $\eta < 1$. To see this, note that at most one unit of the good is misallocated between $i$ and $j$ for any valuation $v \in V$ (when agents report truthfully).

However, the mechanism $\mathcal{M}$ does not provide marginal rewards within $\epsilon + \eta$ (unless $\epsilon = \eta = 0$). To see this, let $a^i = 0$ and consider $\bar{a}^i \in [0, N + 1]$. The increase in the utility of agent $i$ associated with changing his valuation parameter from $a^i = 0$ to $\bar{a}^i$ is $\sum_{m=1}^{\lfloor \bar{a}^i \rfloor} (\bar{a}^i - m + \epsilon)$, while the change in social surplus is $\sum_{m=1}^{\lfloor \bar{a}^i \rfloor} (\bar{a}^i - m - \eta)$; hence, the difference between the increase in the utility of $i$ and the increase in social surplus is given by $\lfloor \bar{a}^i \rfloor (\epsilon + \eta)$, which is greater than $\epsilon + \eta$ if $\bar{a}^i \geq 2$ (unless $\epsilon = \eta = 0$).

Our next result shows that the loss in investment efficiency for $i$ arising from weakening both strategy-proofness for $i$ and ex post efficiency depends on the number of equivalence classes of alternatives faced by agent $i$. Formally, we define

$$\Omega^i \equiv \{ \Psi \subseteq \Omega : \forall \psi \in \Psi, \forall \omega \in \Omega, [\forall v^i \in V^i, v^i(\psi) = v^i(\omega)] \iff \omega \in \Psi \}$$

to be the set of equivalence classes of alternatives for $i$; that is, each element of $\Omega^i$ is the class
of all alternatives that are equally desirable to agent $i$ for every possible valuation function $v^i \in V^i$.

**Theorem 3.** Consider an agent $i \in I$ and suppose that $V^i$ is path-connected. For any $\epsilon \geq 0$ and $\eta \geq 0$, if the mechanism $\mathcal{M}$ is strategy-proof within $\epsilon$ for $i$ and efficient within $\eta$, then it provides marginal rewards to $i$ (and induces efficient investment by $i$) within $|\Omega^i|(\epsilon + \eta)$.\(^\text{19}\)

For many problems, the number of alternatives can be very large, while the number of equivalence classes of alternatives for a given agent $i$ is small. For instance, in a simple bilateral exchange problem with $N$ unit-demand buyers and $N$ unit-supply sellers of a homogenous good, there are \(\frac{(2N)!}{(N!)^2}\) alternatives—a very large number even if $N$ is reasonably small. However, from the perspective of a given agent $i$, there are really only two alternatives: $i$ either engages in trade or does not. Thus, in the case of $\epsilon$-strategy-proof and $\eta$-efficient bilateral exchange, $|\Omega^i| = 2$ for each agent $i$, so the bound in Theorem 3 is $|\Omega^i|(\epsilon + \eta) = 2(\epsilon + \eta)$.

Theorems 2 and 3 together imply an approximate form of the Green–Laffont–Holmström theorem: a mechanism is nearly efficient and nearly strategy-proof if and only if its transfer rule is close to that of a Groves mechanism (see Appendix E).

## 5 Information Acquisition and Uncertain Returns

### 5.1 Framework

We now allow for two different types of uncertainty. The first type of uncertainty, *investment uncertainty*, represents the agent’s uncertainty over the impact of his choice of investment; this uncertainty is resolved prior to the running of the mechanism. The second type of uncertainty, *valuation uncertainty*, represents the agent’s residual uncertainty over his valuation that remains after his investment; this uncertainty is not resolved until after the running of the mechanism.

\(^{19}\)Example 1 can be extended to show that this bound is tight.
We represent valuation uncertainty by allowing an agent \( i \) to have a distribution \( v^i \) over valuations in \( V^i \); we call such a distribution \( v^i \) a *valuation distribution* for \( i \). Let the space of valuation distributions for agent \( i \) be given by \( \mathcal{V}^i \equiv \Delta(V^i) \); more specifically, let \( \mathcal{V}^i \) be the set of Borel probability measures over \( V^i \), and endow it with the standard weak* topology.\(^{20}\) Hence, for an alternative \( \omega \), the *expected value* for agent \( i \) with a valuation distribution \( v^i \) is given by

\[
v^i(\omega) \equiv \int v^i(\omega) \, dv^i(v^i).
\]

We denote \( \mathcal{V} \equiv \times_{i \in I} \mathcal{V}^i \); note that this construction implies that agents’ valuations are drawn independently from each other. In this context, for an alternative \( \omega \in \Omega \) and transfer vector \( t \in \mathbb{R}^I \), and given a valuation distribution \( v^i \in \mathcal{V}^i \), the *expected utility* of agent \( i \) is defined by

\[
u^i((\omega, t); v^i) \equiv v^i(\omega) - t^i = \left( \int v^i(\omega) \, dv^i(v^i) \right) - t^i.
\]

We define the *expected welfare* of an alternative \( \omega \) as the sum of the agents’ expected valuations for that alternative, i.e.,

\[
\mathcal{V}(\omega; v) \equiv \sum_{i \in I} v^i(\omega) = \sum_{i \in I} \left( \int v^i(\omega) \, dv^i(v^i) \right). \tag{3}
\]

We represent the other type of uncertainty, investment uncertainty, by allowing agent \( i \) to select an *investment choice* \( \psi^i \), which is a distribution over valuation distributions. The space of investment choices for agent \( i \) is \( \mathcal{V}^i \equiv \Delta(\mathcal{V}^i) \). We let \( \mathcal{V} \equiv \times_{i \in I} \mathcal{V}^i \); note that this construction implies that agents’ valuation distributions are drawn independently from each other.

The cost of an investment choice is determined by a *cost function* \( c^i : \mathcal{V}^i \rightarrow \mathbb{R} \). Hence, the *ex ante expected utility of agent* \( i \) for an alternative \( \omega \in \Omega \) and transfer vector \( t \in \mathbb{R}^I \),

\(^{20}\)The Lévy-Prokhorov metric is known to induce this topology over distributions.
given an investment choice \( v^i \), is given by

\[
\mathbb{E}_{v^i}[u^i((\omega, t); v^i)] \equiv \int u^i((\omega, t); v^i) \, dv^i(v^i),
\]

where here and throughout the expectation is taken with respect to an investment choice.

An allocation rule maps valuation distributions to outcomes, i.e., it is a map \( \mu : \mathcal{V} \rightarrow \Omega \). A transfer rule is a map \( r : \mathcal{V} \rightarrow \mathbb{R}^I \). We assume that the transfer from each agent is bounded, that is, \( \sup_{v \in \mathcal{V}, i \in I} \{ |r^i(v)| \} < \infty \). A mechanism \( \mathcal{M} \equiv (\mu, r) \) is an ordered pair of an allocation rule \( \mu \) and a transfer rule \( r \); we denote \( \mathcal{M}(v) \equiv (\mu(v), r(v)) \); note that while mechanisms take as input valuation distributions, in practice a mechanism could depend on just the expected value of each alternative for each agent.\(^{21}\) We abuse notation slightly by, for a mechanism \( \mathcal{M} = (\mu, r) \), writing \( v^i(\mathcal{M}(v)) \equiv v^i(\mu(v)) \) and \( V(\mathcal{M}(v); v) \equiv V(\mu(v); v) \).

### 5.2 An Example

We now consider a simple example to illustrate the uncertainty structure in our framework.

**Example 2.** Consider a second-price auction with two bidders and one item, where \( I = \{i, j\} \) and \( \Omega \cong I.\(^{22}\)\)

We wish to consider a setting where the agents may engage in *ex ante* information acquisition. In particular, agent \( i \) is uncertain before the investment stage about the value of the item: it is worth 10 or 20 to him with equal probability. Agent \( i \) can choose to spend 1 to learn the true value of the item before making his report to the mechanism, or he can choose to spend 0, in which case he does not learn anything further about the value of the item and is left with his prior beliefs. Agent \( j \), meanwhile, has no uncertainty about his valuation—the item is worth 14 to him for certain.

\(^{21}\)Note that if a mechanism maximizes expected welfare, then (generically) its associated allocation rule can only depend on the expected value of each alternative for each agent.

\(^{22}\)Here, a second-price auction assigns the item to the agent with the highest expected value for the item, and that agent pays a transfer equal to the other agent’s expected value for the item.
Formally, then, the valuation space for agent $i$ is given by $V^i = \{a\mathbb{1}_{\omega=i} : a \in \{10, 20\}\}$. We let $v^i = 10\mathbb{1}_{\omega=i}$, under which $i$ has a value of 10 for receiving the item, and $\bar{v}^i = 20\mathbb{1}_{\omega=i}$, under which $i$ has a value of 20 for receiving the item.

The space of valuation distributions is given by $\mathcal{V}^i = \Delta(V^i)$. However, we will focus our attention on three particular valuation distributions:  

1. $v^i$, which has the probability density function $\delta_{v^i}$, and thus puts a probability weight of 1 on the value of the item to $i$ being 10;  

2. $\bar{v}^i$, which has the probability density function $\delta_{\bar{v}^i}$, and thus puts a probability weight of 1 on the value of the item to $i$ being 20;  

3. $\chi^i$, which has the probability density function $\frac{1}{2}\delta_{v^i} + \frac{1}{2}\delta_{\bar{v}^i}$, and thus puts a probability weight of $\frac{1}{2}$ on the value of the item to $i$ being 10 and puts a probability weight of $\frac{1}{2}$ on the value of the item to $i$ being 20.  

The space of investment choices for $i$ is given by $V^i = \Delta(\Delta(V^i))$. However, we focus our attention on two investment choices:  

1. $v^i$, which has the probability density function $\frac{1}{2}\delta_{v^i} + \frac{1}{2}\delta_{\bar{v}^i}$, and thus puts a probability weight of $\frac{1}{2}$ on $v^i$ and a probability weight of $\frac{1}{2}$ on $\bar{v}^i$;  

2. $\chi^i$, which has the probability density function $\delta_{\chi^i}$, and thus puts a probability weight of 1 on the distribution $\chi^i$.

Here, then, the investment choice $v^i$ represents $i$ learning whether the true value of the item to him is 10 (i.e., $i$’s valuation distribution is $v^i$) or 20 (i.e., $i$’s valuation distribution is $\bar{v}^i$). The investment choice $\chi^i$ represents $i$ not learning the true value of the item (i.e., $i$’s valuation distribution is $\chi^i$). We let the cost function of agent $i$ over investment choices be  

\footnote{Here, we use the notation $\delta_z$ to denote the Dirac delta function centered on $z$.}
given by

\[
c^i(y^i) = \begin{cases} 
1 & y^i = v^i \\
0 & y^i = x^i \\
2017 & \text{otherwise.}
\end{cases}
\]

The valuation space for agent \( j \) is given by \( V^j = \{v^j\} \) where \( v^j = 14\mathbb{1}_{(\omega=j)} \), i.e., the value for agent \( j \) for the alternative \( i \) is always 0, and the value for agent \( j \) for the alternative \( j \) is always 14. Hence, the space of valuation distributions \( \nu^j = \Delta(V^j) \equiv \{v^j\} \) is a singleton set, where \( v^j = \delta_{v^j} \). It follows that the space of investment choices \( \nu^j = \Delta(\nu^j) \equiv \{\delta_{v^j}\} \) is also a singleton set; that is, there is a single investment choice for \( j \) that corresponds to putting full weight on the only possible valuation distribution \( v^j \). We let the cost function of agent \( j \) be given by \( c^j(\delta_{v^j}) = 0 \).

The optimal strategy for \( i \) during the investment phase is to choose \( v^i \). If the item is worth 10, then agent \( i \) will not receive the item, resulting in a total payoff of \(-1\), as he incurs a cost of 1 from choosing \( v^i \). If the item is worth 20, then agent \( i \) will receive the item, resulting in a payoff of 5, as, in the second-price auction \( i \) will pay a transfer of 14 for the item and incur a cost of 1 from choosing \( v^i \). Hence, the expected payoff for \( i \) from choosing \( v^i \) is \( \frac{-1+5}{2} = 2 \). By contrast, if \( i \) chooses \( x^i \) during the investment phase, he will always receive the item while paying a transfer of 14, and so obtains an expected payoff of 1.

Moreover, it is socially optimal for \( i \) to make the investment choice \( v^i \). The expected social welfare when \( i \) chooses \( v^i \) is

\[
16 = \frac{1}{2} \cdot 20 + \frac{1}{2} \cdot 14 - 1.
\]

By contrast, the expected social welfare when \( i \) chooses \( x^i \) is

\[
15 = \frac{1}{2} \cdot 20 + \frac{1}{2} \cdot 10,
\]
as in that case $i$ always receives the item.

### 5.3 Inducing Efficient Investment under Uncertainty

In settings with uncertainty, a mechanism induces efficient investment by $i$ within $\epsilon$ if (assuming agents report truthfully) every expected utility-maximizing investment choice by $i$ maximizes expected social welfare within $\epsilon$.

**Definition 9.** A mechanism $\mathcal{M}$ *induces efficient investment within* $\epsilon$ *by* $i \in I$, if, for all $v^{I \setminus \{i\}} \in \mathcal{V}^{I \setminus \{i\}}$, if

$$\hat{\psi}^i \in \arg \max_{\tilde{\psi}^i \in \mathcal{V}^i} \left\{ \mathbb{E}(\hat{\psi}^i, v^{I \setminus \{i\}}) \left[ u^i(\mathcal{M}(v^i, v^{I \setminus \{i\}}); v^i) \right] - c^i(\tilde{\psi}^i) \right\},$$

then we have

$$\left( \mathbb{E}(\hat{\psi}^i, v^{I \setminus \{i\}}) \left[ V(\mathcal{M}(v^i, v^{I \setminus \{i\}}); (v^i, v^{I \setminus \{i\}})) \right] \right) - c^i(\hat{\psi}^i) + \epsilon \geq \sup_{\tilde{\psi}^i \in \mathcal{V}^i} \left\{ \left( \mathbb{E}(\hat{\psi}^i, v^{I \setminus \{i\}}) \left[ V(\mathcal{M}(v^i, v^{I \setminus \{i\}}); (v^i, v^{I \setminus \{i\}})) \right] \right) - c^i(\tilde{\psi}^i) \right\}$$

for all cost functions $c^i$.

Note that in Definition 9 the expectation is taken over $v^{I \setminus \{i\}}$, the investment choices of other agents; effectively, we are requiring that $i$’s privately optimal investment choice is socially optimal within $\epsilon$ given the investment choices of others. This assumption is natural, as, in our framework, all investment choices are made prior to the resolution of investment uncertainty; thus, $i$, when selecting his investment choice, can not be aware of the valuation distributions realized following others’ investment choices.

However, we state our strategy-proofness conditions in terms of $i$’s choice of reported valuation distribution, as opposed to $i$’s investment choice, as valuation distributions have been realized at the time of reports. Similarly, we state our provision of marginal rewards and efficiency conditions in terms of valuation distributions. Thus, in Sections 5.4 and 5.5, we
consider extensions of our generalizations of strategy-proofness, provision of marginal rewards, and efficiency (Definitions 5, 6, and 8) that treat agents as having complete information about others’ realized valuation distributions. In Section 5.4, we require all three conditions to hold “pointwise,” i.e., for all possible profiles of valuation distributions of other agents. However, such a requirement may be unrealizable in many market design settings. Thus, in Section 5.5 we consider a relaxation of all three conditions where we only require that the relevant conditions hold in expectation for any given investment choice profile of the other agents. These latter conditions allow us to bound the ex ante efficiency loss of many commonly-used mechanisms such as uniform-price auctions and double auctions (see Sections 6.3 and 6.4).

5.4 Bounding Departures from Full Efficiency

5.4.1 Properties of Mechanisms

Our definition of strategy-proofness within $\epsilon$ in this context naturally generalizes the definition from the context without valuation uncertainty: A mechanism is strategy-proof within $\epsilon$ for $i$ if reporting his true valuation distribution yields $i$ an expected utility within $\epsilon$ of the highest expected utility he can attain from any report.

**Definition 10.** A mechanism $\mathcal{M}$ is strategy-proof within $\epsilon$ for $i \in I$ if, for all $v \in V$ and $\bar{v}^i \in v^i$,

$$u^i(\mathcal{M}(v); v^i) + \epsilon \geq u^i(\mathcal{M}(\bar{v}^i; v^{I \setminus \{i\}}); \bar{v}^i).$$

Similarly, a mechanism provides marginal rewards within $\epsilon$ in this context if, assuming agents report truthfully, the change in the expected utility of $i$ induced by a change in the valuation distribution of $i$ is within $\epsilon$ of the change in expected social welfare.

**Definition 11.** A mechanism $\mathcal{M}$ provides marginal rewards within $\epsilon$ to $i \in I$ if, for all $v \in V$ and $\bar{v}^i \in v^i$,

$$\left| (V(\mathcal{M}(v); v) - V(\mathcal{M}(\bar{v}^i; v^{I \setminus \{i\}}); (\bar{v}^i; v^{I \setminus \{i\}}))) - (u^i(\mathcal{M}(v); v^i) - u^i(\mathcal{M}(\bar{v}^i; v^{I \setminus \{i\}}); \bar{v}^i)) \right| \leq \epsilon.$$
Finally, we generalize our weakened efficiency concept to incorporate valuation uncertainty.

**Definition 12.** A mechanism $\mathcal{M}$ is efficient within $\eta$ if, for all $\nu \in \mathcal{V}$,

$$V(\mathcal{M}(\nu); \nu) + \eta \geq \max_{\omega \in \Omega} \{V(\omega; \nu)\}.$$

### 5.4.2 Results

We now give analogues of the results of Section 4.2.

**Lemma 3.** Consider a mechanism $\mathcal{M}$ and an agent $i \in I$ in the setting of Section 5.1. For any $\epsilon \geq 0$, the following two statements are equivalent:

1. The mechanism $\mathcal{M}$ induces efficient investment within $\epsilon$ by $i$.

2. The mechanism $\mathcal{M}$ provides marginal rewards within $\epsilon$ to $i$.

**Theorem 4.** Consider a mechanism $\mathcal{M}$ and an agent $i \in I$ in the setting of Section 5.1. For any $\epsilon \geq 0$ and $\eta \geq 0$, if the mechanism $\mathcal{M}$ provides marginal rewards within $\epsilon$ to $i$ and is efficient within $\eta$, then the mechanism $\mathcal{M}$ is strategy-proof within $\epsilon + \eta$ for $i$.

For our analogue of Theorem 3 we modify our equivalence class definition. Formally, here and throughout the remainder of Section 5, we take

$$\Omega^i \equiv \left\{ \Psi \subseteq \Omega : \forall \psi \in \Psi, \forall \omega \in \Omega, \left[ \forall \varphi^i \in \mathcal{V}^i, \varphi^i(\psi) = \varphi^i(\omega) \right] \iff \omega \in \Psi \right\}$$

to be the set of equivalence classes of alternatives for $i$.

**Theorem 5.** Consider a mechanism $\mathcal{M}$ and an agent $i \in I$ in the setting of Section 5.1. For any $\epsilon \geq 0$ and $\eta \geq 0$, if the mechanism $\mathcal{M}$ is strategy-proof within $\epsilon$ for $i$ and efficient within $\eta$, then it provides marginal rewards (and induces efficient investment) within $|\Omega^i| \epsilon + \eta$ to $i$.

---

24Unlike Theorems 1 and 3, we do not need to explicitly assume that $\mathcal{V}^i$ is path-connected, as its path-connectedness follows immediately from the assumption that $\mathcal{V}^i = \Delta(V^i)$. 

27
5.4.3 Discussion

Combining Lemma 3 and Theorems 4 and 5 in the “exactly” strategy-proof and efficient case ($\epsilon = 0$ and $\eta = 0$), we find that Vickrey-Clarke-Groves mechanisms induce efficient investment by all participants. This observation generalizes Theorem 1 of Bergemann and Välimäki (2002), who showed this result in the case in which investment represents acquisition of information about ex post valuations.\(^{25,26}\)

Moreover, our results show that strategy-proofness is not only sufficient but also necessary for ex post efficient mechanisms to induce ex ante efficient investment and that this relationship still roughly holds for mechanisms that do not satisfy exact efficiency or incentive properties.

Using Lemma 3 and Theorem 5, we can analyze equilibrium outcomes of the game induced by efficient and strategy-proof mechanisms, just as we did in the case without uncertainty (Corollary 1). Given cost functions $\{c^i\}_{i \in I}$, the game induced by $\mathcal{M}$ is a game between the agents in $I$ where each agent $i$ selects an investment choice $v^i \in V^i$ and receives the payoff

$$E_{v^i, v^I \setminus \{i\}}[u^i(\mathcal{M}(v^i, v^I \setminus \{i\}); v^i)] - c^i(v^i);$$

agents’ cost functions are assumed to be common knowledge.

**Corollary 2.** Consider a mechanism $\mathcal{M}$ in the setting of Section 5.1. If $\mathcal{M}$ is efficient and strategy-proof, then there exists a Nash equilibrium $v$ of the game induced by $\mathcal{M}$ such that $v$ maximizes ex ante social welfare.\(^{27}\)

\(^{25}\)To map the model of Bergemann and Välimäki (2002) into our framework, we take
- the set $\Omega$ of alternatives to be the set of Bergemann–Välimäki alternatives, $X$;
- the set $V^i$ of valuations for agent $i$ to be the set of Bergemann–Välimäki states for $i$, $\Omega_i$;
- the valuation distributions $v^i$ to be the Bergemann–Välimäki signal realizations, $p_i$; and
- the investment choices $v^i$ to be the Bergemann–Välimäki distributions of signals, $F^{\alpha_i}$.

With this identification, the investment efficiency criterion (Definition 2) of Bergemann and Välimäki (2002) corresponds to our inducement of efficient investment condition, Definition 9, with $\epsilon = 0$.

\(^{26}\)Bergemann and Välimäki (2002) also suggest that Vickrey-Clarke-Groves mechanisms induce efficient investment beyond settings in which agents invest only in information, although without formal arguments.

\(^{27}\)By efficient, we mean that $\mathcal{M}$ is efficient within 0, and by strategy-proof, we mean that $\mathcal{M}$ is strategy-proof within 0 for each $i \in I$.  

28
Finally, we note that the main result of Crémer and Riordan (1985) shows that, in our setting, one can always construct a mechanism that is efficient, budget-balanced, and strategy-proof for all but one agent $i$, and for which truth-telling is optimal (in the Bayesian sense) for $i$. This mechanism is constructed by providing all but one agent transfers that are equal to those provided by a Vickrey-Clarke-Groves mechanism, while the remaining agent’s transfer includes a supplementary transfer (that relies on his and others’ reports) that balances the budget. Lemma 3 and Theorem 5 then imply that a Crémer and Riordan (1985) mechanism would, in our setting, induce efficient investment by all but one agent.

5.5 Bounding Departures from Full Efficiency in Expectation

Many commonly used mechanisms, such as the uniform-price auction and standard double auction mechanisms, do not satisfy our “pointwise” properties as defined in Section 5.4. For instance, in the uniform-price auction, there exists a configuration of valuation distributions $v^I$ such that there may be large gains for $i$ from misstating his valuation distribution: For example, suppose that there are $k$ identical objects available and that the expected values for other agents (at the reporting stage) are such that $k - 1$ of them place a high expected value on obtaining one object (and no value on subsequent objects) while other agents place a low expected value on obtaining an object. In this realization, if $i$ has a high (expected) value for the object, $i$ can significantly reduce the price he pays by reporting a valuation distribution which under-represents his (expected) value for the object.

Despite the preceding observation, for most realizations of valuation distributions $v^{I \setminus \{i\}}$, agent $i$ can not gain much by lying regardless of his valuation; it follows that the expected gain for $i$ from lying is small in expectation. Consequently, the uniform-price auction is nearly strategy-proof in expectation, in a sense that we formalize. In this section, we bound the distortion in investment incentives away from socially optimal investment for mechanisms that are strategy-proof within $\epsilon$ in expectation and efficient within $\eta$ in expectation. In Section 6.3, we apply the results from this section to uniform-price auctions; in Section 6.4, we apply the
results from this section to the McAfee (1992) double auction.

5.5.1 Properties of Mechanisms

In considering the properties of mechanisms in expectation, we relax our earlier assumption that $V^i = \Delta(\nu^i)$ and instead only require that $V^i \subseteq \Delta(\nu^i)$.\(^{28}\) If $V^i = \Delta(\nu^i)$ (in particular, if $\delta_{v^i} \in V^i$ for all $\nu^i \in \nu$), then the conditions we discuss in this section reduce to the conditions of Section 5.4.

First, we introduce our concept of strategy-proofness within $\epsilon$ in expectation for $i$. This condition requires that (for every possible choice of investment by other agents) $i$’s gain from misreporting optimally is at most $\epsilon$ in expectation, i.e., averaging across all the possible realized valuation distributions for other agents.

**Definition 13.** A mechanism $\mathcal{M}$ is strategy-proof within $\epsilon$ in expectation for $i \in I$ if, for all $\nu^{I \setminus \{i\}} \in \nu^{I \setminus \{i\}}$,

$$
\mathbb{E}_{\nu^{I \setminus \{i\}}} \left[ \sup_{\nu^i, \bar{\nu}^i \in \nu^i} \left\{ u^i(\mathcal{M}(\bar{\nu}^i, \nu^{I \setminus \{i\}}); \nu^i) - u^i(\mathcal{M}(\nu^i, \nu^{I \setminus \{i\}}); \nu^i) \right\} \right] \leq \epsilon.
$$

Intuitively, this condition requires that, in expectation, for any possible investment choices $\nu^{I \setminus \{i\}} \in \nu^{I \setminus \{i\}}$ by the other agents, the maximum $i$ can possibly gain from lying is $\epsilon$. We note two key features of this condition:

1. The supremum is taken with respect to both the actual valuation distribution $\nu^i$ of $i$ as well as the reported valuation distribution $\bar{\nu}^i$ of $i$.

2. We allow $i$’s actual and reported valuation distributions to depend on the reported valuation distributions $\nu^{I \setminus \{i\}}$ of the other agents, as opposed to just depending on $\nu^{I \setminus \{i\}}$.

Intuitively, a mechanism is strategy-proof within $\epsilon$ in expectation for $i$ if, for most realizations of valuation distributions for agents other than $i$, agent $i$’s gain from misrepresenting...
his valuation distribution is bounded regardless of his actual valuation distribution in such a way that i’s expected gain from misrepresentation is at most \( \epsilon \). As \( \epsilon \) goes to 0, Definition 13 becomes a form of strategy-proofness in expectation: for almost all realizations of other agents’ valuation distributions, there is no gain from misreporting. Consequently, if a mechanism \( \mathcal{M} \) is strategy-proof within \( \epsilon \) in expectation and the \( \epsilon \) bound can be taken to 0 as the number of agents grows, then the mechanism is *strategy-proof in the large* in the sense of Azevedo and Budish (2015) (under certain technical conditions).\(^{29}\)

Our strategy-proofness in expectation concept builds on the classical idea that under a strategy-proof mechanism, truthful reporting is a dominant strategy for each agent. In Definitions 5 and 10, we generalized this by requiring that truthful reporting is an \( \epsilon \)-best response for any reports by other agents. Here, we further generalize this concept by only requiring that truthful reporting is an \( \epsilon \)-best response in expectation, i.e., that the \( \epsilon \)-best response property holds in expectation across resolutions of investment uncertainty.

In parallel to our definition of strategy-proofness within \( \epsilon \) in expectation, we now state our definition of providing marginal rewards within \( \epsilon \) in expectation.

**Definition 14.** A mechanism \( \mathcal{M} \) provides marginal rewards within \( \epsilon \) in expectation to \( i \in I \)

\[ \sup_{v^i,\bar{v}^i \in \mathcal{V}^i} \left\{ \mathbb{E}_{v_{-\{i\}}} \left[ u^i(\mathcal{M}(\bar{v}^i, v^i_{-\{i\}}); v^i) - u^i(\mathcal{M}(v^i, v^i_{-\{i\}}); v^i) \right] \right\} \leq \epsilon. \]

Strategy-proofness in the large is weaker than requiring that a mechanism is strategy-proof in expectation within \( \epsilon \) (where, for each \( j \in I \), \( \mathcal{V}^j \) is the set of all investment choices with full support over \( v^j \)) where the \( \epsilon \) bound can be taken to 0 as the market grows large, as strategy-proofness in the large only allows \( i \) to best respond to the distribution of valuation distributions for the other agents as opposed to the realized valuation distributions. Moreover, strategy-proofness in the large is only defined for symmetric distributions \( v \), while strategy-proofness in expectation with \( \epsilon \) allows for different agents to select different investment choices. That said, we are unaware of any real-world mechanisms for settings with transferrable utility that are strategy-proof in the large but are not strategy-proof within \( \epsilon \) in expectation where the \( \epsilon \) bound can be taken to 0 as the number of agents grows large.
if, for all \( v^{I\setminus\{i\}} \in V^{I\setminus\{i\}} \),

\[
\mathbb{E}_{v^{I\setminus\{i\}}} \left[ \sup_{v^i, \bar{v}^i \in V^i} \left\{ \left( \mathbf{V}(\mathcal{M}(v^i, v^{I\setminus\{i\}})); (v^i, v^{I\setminus\{i\}})) - \mathbf{V}(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}})); (\bar{v}^i, v^{I\setminus\{i\}})) \right) 
- \left( u^i(\mathcal{M}(v^i, v^{I\setminus\{i\}})); v^i) - u^i(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); \bar{v}^i) \right) \right\} \right] \leq \epsilon.
\]

This condition requires that, in expectation, for any investment choices \( v^{I\setminus\{i\}} \in V^{I\setminus\{i\}} \),
the maximum gap between the social and private returns to a change in \( i \)'s valuation distribution is at most \( \epsilon \). Note that, as in Definition 13, the choices of \( v^i \) and \( \bar{v}^i \) for the supremum in Definition 14 may depend on the realized valuation distributions \( v^{I\setminus\{i\}} \) of the other agents, not just the investment choices \( v^{I\setminus\{i\}} \) of the other agents.

Lastly, we introduce our concept of efficiency within \( \eta \) in expectation.

**Definition 15.** A mechanism \( \mathcal{M} \) is **efficient within \( \eta \) in expectation** if, for all \( i \in I \), for all \( v^{I\setminus\{i\}} \in V^{I\setminus\{i\}} \),

\[
\mathbb{E}_{v^{I\setminus\{i\}}} \left[ \sup_{v^i, \omega \in \Omega} \left\{ \mathbf{V}(\omega; (v^i, v^{I\setminus\{i\}})) - \mathbf{V}(\mathcal{M}(v^i, v^{I\setminus\{i\}})); (v^i, v^{I\setminus\{i\}})) \right\} \right] \leq \eta.
\]

This condition requires that, for each \( i \in I \), in expectation, for any investment choices \( v^{I\setminus\{i\}} \in V^{I\setminus\{i\}} \), the maximum efficiency loss that can be induced by the mechanism is small. Note that, analogously to Definitions 13 and 14, the choices of \( v^i \) and \( \omega \) for the supremum in Definition 14 may depend on the realized valuation distributions \( v^{I\setminus\{i\}} \) of the other agents, not just the investment choices \( v^{I\setminus\{i\}} \) of the other agents. Additionally, we consider for each agent \( i \) the efficiency loss in expectation over \( v^{I\setminus\{i\}} \) when \( i \)'s valuation \( v^i \) maximizes the efficiency loss of the mechanism given \( v^{I\setminus\{i\}} \) (unlike our earlier definitions of efficiency, where we consider the maximum efficiency loss induced by allowing \( v \), the valuation distribution profile of all agents, to vary freely).\(^{30}\)

\(^{30}\)This formalization of the concept of efficiency in expectation is important for ensuring that a mechanism which provides marginal rewards within \( \epsilon \) in expectation to \( i \) and is efficient within \( \eta \) in expectation will be strategy-proof within \( \epsilon + \eta \) in expectation for \( i \) and, similarly, that a mechanism which is strategy-proof within \( \epsilon \) in expectation for \( i \) and is efficient within \( \eta \) in expectation will provide marginal rewards within \( |\Omega^i| (\epsilon + \eta) \) in expectation to \( i \).
5.5.2 Results

We first show that if the failure of a mechanism to provide marginal rewards in expectation is bounded by $\epsilon$ then the failure of that mechanism to induce efficient investment is also bounded by the same $\epsilon$.

**Lemma 4.** Consider a mechanism $\mathcal{M}$ and an agent $i \in I$ in the setting of Section 5.1. For any $\epsilon \geq 0$, if the mechanism $\mathcal{M}$ provides marginal rewards within $\epsilon$ in expectation to $i$, then the mechanism $\mathcal{M}$ induces efficient investment within $\epsilon$ for $i$.

Note that, unlike in Lemma 3, we do not have a converse to Lemma 4. This is because we no longer assume that $V^i = \Delta(V^i)$: Inducing efficient investment by $i$ implies providing marginal rewards to $i$ only given a certain amount of richness in the space $V^i$ of investment choices. Thus, without additional assumptions on $V^i$ (e.g., that $\delta_{v^i} \in V^i$ for all $v^i \in V^i$) we can not prove a converse to Lemma 4.

In analogy to Theorem 4, we now show that if a mechanism provides marginal rewards within $\epsilon$ in expectation and is efficient within $\eta$ in expectation, then the mechanism is strategy-proof in expectation within $\epsilon + \eta$.

**Theorem 6.** Consider a mechanism $\mathcal{M}$ and an agent $i \in I$ in the setting of Section 5.1. For any $\epsilon \geq 0$ and $\eta \geq 0$, if the mechanism $\mathcal{M}$ provides marginal rewards within $\epsilon$ in expectation to $i$ and is efficient within $\eta$ in expectation, then the mechanism $\mathcal{M}$ is strategy-proof within $\epsilon + \eta$ in expectation for $i$.

Finally, we show an analogue of Theorem 5: we obtain a bound on the failure of a mechanism to provide marginal rewards in expectation as a function of the degree to which that mechanism fails to be strategy-proof and efficient in expectation.

**Theorem 7.** Consider a mechanism $\mathcal{M}$ and an agent $i \in I$ in the setting of Section 5.1. For any $\epsilon \geq 0$ and $\eta \geq 0$, if the mechanism $\mathcal{M}$ is strategy-proof within $\epsilon$ in expectation for $i$ and
efficient within \( \eta \) in expectation, then the mechanism \( \mathcal{M} \) provides marginal rewards within
\[ |\Omega^i|(\epsilon + \eta) \] in expectation to \( i \).\(^{31,32}\)

6 Applications

In this section, we provide four economic applications of our results. We first consider the labor market matching setting of Kelso and Crawford (1982) with quasilinear utility in Section 6.1: We show that the worker-optimal stable mechanism induces efficient investment by workers and is strategy-proof for workers in the presence of continuous transfers. Furthermore, our results allow us to bound the inefficiency in workers’ investments that arise under the worker-optimal stable mechanism when transfers are discrete. Then, in Sections 6.3 and 6.4, we apply our results on settings with uncertainty to show that in large markets, uniform-price auctions and double auctions approximately induce efficient investment in expectation.\(^{33}\)

6.1 Labor Market Matching

A classic question in labor economics asks how market structure affects the acquisition of human capital. In a companion paper (Hatfield, Kojima, and Kominers, 2014), we used Theorem 3 to show that the worker-optimal stable mechanism induces approximately efficient human capital investment by workers in the Kelso and Crawford (1982) model of job matching

\[^{31}\text{Recall that, as defined in Section 5.4, the set of equivalence classes of outcomes for } i \text{ is now defined with respect to the space of valuation distributions } \mathcal{V}^i \text{ instead of the space of valuations } V^i, \text{ i.e.,} \]

\[ \Omega^i \equiv \{ \Psi \subseteq \Omega : \forall \psi \in \Psi, \forall \omega \in \Omega, [\forall v^i \in \mathcal{V}^i, v^i(\psi) = v^i(\omega)] \Rightarrow \omega \in \Psi \}. \]

\[^{32}\text{Recall that, as in Theorem 5 (and unlike Theorems 1 and 3) we do not need to explicitly assume that } \mathcal{V}^i \text{ is path-connected, as its path-connectedness follows immediately from the fact that } \mathcal{V}^i = \Delta(V^i). \]

\[^{33}\text{In Appendix C, we provide one further application, showing that in the trading network setting of Hatfield et al. (2013), the seller-optimal stable mechanism provides marginal rewards for each unit-supply seller. Lemma 1 then implies that the seller-optimal stable mechanism induces efficient investment by unit-supply sellers, and Theorem 1 implies that the seller-optimal stable mechanism is strategy-proof for unit-supply sellers.} \]

\[^{34}\text{In work in progress, Hatfield et al. (2016) consider auctions with endogenous entry, using our results for settings with uncertainty to obtain both efficient and revenue-maximizing mechanisms, building on the work of McAfee and McMillan (1987), Engelbrecht-Wiggans (1993), and Jehiel and Lamy (2015).} \]
with discrete transfers. Here, we sharpen the result of Hatfield, Kojima, and Kominers (2014) by improving the relevant bound. Additionally, we provide the first analysis of workers’ ex ante investment incentives in the Kelso and Crawford (1982) model of job matching with continuous transfers.

6.1.1 The Kelso–Crawford (1982) Matching Model

There are finite sets $W$ and $F$ of workers and firms, respectively; hence, the set of agents is $I = W \cup F$. A pairing $(w, f)$ specifies that firm $f$ employs worker $w$; the set of pairings is given by $W \times F$. In this setting, an alternative is a matching specifying a set of employment pairings $\omega \subseteq W \times F$ such that each worker is assigned to at most one firm.

The space of valuations for a worker $w \in W$ is given by

$$V^w = \left\{ v^w(\omega) = \sum_{f \in F} a^w_f \mathbb{1}_{(w,f) \in \omega} : a^w_f \in \mathbb{R} \text{ for all } f \in F \right\};$$

intuitively, $a^w_f$ represents the utility $w$ obtains by working for $f$.

To define the space of valuations for a firm $f \in F$, first consider the set

$$\hat{V}^f = \left\{ v^f(\omega) = \sum_{\bar{W} \subseteq W} a^f_{\bar{W}} \mathbb{1}_{\bar{W} = \{w \in W : (w,f) \in \omega\}} : a^f_{\bar{W}} \in \mathbb{R} \text{ for all } \bar{W} \subseteq W \right\};$$

intuitively, $a^f_{\bar{W}}$ represents the utility $f$ from employing the set of workers $\bar{W}$.

There is a set of possible salaries $P \subseteq \mathbb{R}$; we assume that $0 \in P$. A salary vector $\pi$ is a vector of salaries in $\Pi \equiv P^{W \times F}$. The valuation function $v^f(\omega) = \sum_{\bar{W} \subseteq W} a^f_{\bar{W}} \mathbb{1}_{\bar{W} = \{w \in W : (w,f) \in \omega\}}$ for firm $f$ induces a demand correspondence

$$D^f(\pi; v^f) \equiv \arg\max_{\bar{W} \subseteq W} \left\{ a^f_{\bar{W}} - \sum_{w \in \bar{W}} \pi^w \right\}.$$

We say that a valuation $v^f$ induces a substitutable demand correspondence for $f$ if, for any

\[ \pi, \bar{\pi} \in \Pi \text{ such that } \pi^f \leq \bar{\pi}^f, \text{ if } W' \in D^f(\pi; v^f) \text{ then there exists } \bar{W}' \in D^f(\bar{\pi}; v^f) \text{ such that } \{w \in W' : \pi^w.f = \bar{\pi}^w.f \} \subseteq \bar{W}'. \]

That is, the firm’s demand correspondence is substitutable when the firm keeps worker \( w \) when the wages of other workers rise (while the wage of \( w \) is unchanged). We let \( V^f = \{v^f \in \hat{\nu}^f : v^f \text{ induces a substitutable demand correspondence} \} \);

thus, we say that the set of valuations for \( f \) is substitutable.

A transfer vector \( t \in \mathbb{R}^I \) is \( \omega \)-compatible if

1. for each worker \( w \in W \), we have\(^{36} -t^w \in P \) and, if there does not exist a firm \( f \in F \) such that \( (w, f) \in \omega \), then \( t^w = 0 \), and

2. for each firm \( f \in F \),

\[ t^f = - \sum_{w \in \{\bar{w} \in W : (\bar{w}, f) \in \omega\}} t^w. \]

A transfer vector is \( \omega \)-compatible if each worker receives a salary that is in \( P \), workers who are unemployed under \( \omega \) receive a salary of 0, and each firm pays an amount equal to the sum of the salaries received by the workers it employs under \( \omega \).

A *matching mechanism* \( M \) takes as input a vector of valuations and returns an alternative \( \omega \) and an \( \omega \)-compatible transfer vector \( t \). An outcome \( (\omega, t) \) is *stable* if it is

1. *individually rational*, i.e., \( u^i((\omega, t); v^i) \geq 0 \) for all \( i \in I \), and

2. *unblocked*, i.e., there exist no firm \( f \in F \) and a set of workers \( \hat{W} \subseteq W \) such that, for the alternative \( \hat{\omega} = \bigcup_{w \in \hat{W}} \{(w, f)\} \) there is some \( \hat{\omega} \)-compatible transfer vector \( \hat{t} \) such that

\[ u^i((\hat{\omega}, \hat{t}); v^i) \geq u^i((\omega, t); v^i) \]

for all \( i \in \hat{W} \cup \{f\} \), with at least one inequality holding strictly.

\(^{35}\)This condition is equivalent to the *gross substitutes* condition of *Kelso and Crawford* (1982). This condition also corresponds to a special case of the definition of full substitutability given in our discussion of trading networks: see Definition C.1 in Appendix C.

\(^{36}\)Note that we have defined \( u^w((\omega, t); v^w) = v^w(\omega) - t^w \); hence, the salary paid to worker \( w \) takes the form of a negative transfer.
Kelso and Crawford (1982) showed that if all firms’ preferences are substitutable then there exists a stable outcome. For the case with discrete transfers, Hatfield and Milgrom (2005) showed that there exists a worker-optimal stable outcome, that is, a stable outcome that all workers weakly prefer to any other stable outcome. Meanwhile, the results of Hatfield et al. (2013) imply that there exists a worker-optimal stable outcome for the Kelso and Crawford (1982) setting with continuous transfers.

We now examine workers’ investment incentives in the Kelso and Crawford (1982) framework.

6.1.2 Investment Incentives in Labor Markets with Continuous Transfers

First, we consider the case where $P$ is a convex subset of $\mathbb{R}$—in particular, we assume that continuous transfers are allowed. We further assume that the domain of salaries $P$ is large enough that there are sufficiently high and low wages $p^\text{max}, p^\text{min} \in P$ such that the marginal change in an agent’s valuation from a change in that agent’s assigned partners is always between $p^\text{max}$ and $p^\text{min}$.

We now show that the worker-optimal stable mechanism provides marginal rewards to each worker.

**Proposition 1.** If the set of valuations is substitutable for all $f \in F$, and the domain of salaries $P$ is a convex and sufficiently large subset of $\mathbb{R}$, then the worker-optimal stable mechanism provides marginal rewards to each worker $w \in W$.

Demange (1982) and Leonard (1983) showed that the seller-optimal stable mechanism provides marginal rewards in the one-to-one assignment game of Shapley and Shubik (1971). Gul and Stacchetti (1999) showed the same result in their model of an exchange economy with substitutable preferences. Proposition 1 generalizes all of these results to the setting of Kelso and Crawford (1982).

---

A sufficient condition for this is that both $-\inf P$ and $\sup P$ are larger than $\sup_{v \in V, \omega \in \Omega} \{ V(\omega; v) \}$; weaker sufficient conditions exist, but are more cumbersome to formalize.
Corollary 3. If the set of valuations is substitutable for all $f \in F$, and the domain of salaries $P$ is a convex and sufficiently large subset of $\mathbb{R}$, then the worker-optimal stable mechanism is strategy-proof for each worker $w \in W$.

Corollary 4. If the set of valuations is substitutable for all $f \in F$, and the domain of salaries $P$ is a convex and sufficiently large subset of $\mathbb{R}$, then the worker-optimal stable mechanism induces efficient investment for each worker $w \in W$.

Corollary 3 follows from combining Proposition 1 with Theorem 1 (and noting that the worker-optimal stable mechanism is efficient). Corollary 4 follows from combining Proposition 1 with Lemma 1.

Corollary 3 generalizes results by Demange (1982) and Leonard (1983), who showed that the seller-optimal stable mechanism is strategy-proof for sellers in the one-to-one assignment game. It has long been presumed that the worker-optimal stable mechanism is strategy-proof for workers in the setting of Kelso and Crawford (1982) with continuous transfers. However, to the best of our knowledge, Corollary 3 is the first formal demonstration of this fact.

Taking our results to a marriage market context, we can provide a partial answer to a price theory question posed by Becker and Murphy (2014), whom we paraphrase here:

*Suppose a particular woman $w$ has to choose her level of education, while everyone else’s education level is held fixed. How does the social gain from $w$’s investment differ from the private gain?*

Proposition 1 shows that under the woman-optimal stable mechanism, the social gain from $w$’s investment exactly equals the private gain. Consequently, Corollary 4 shows that $w$ is incentivized to make the socially optimal investment.\(^{38}\)

\(^{38}\)In Appendix C, we generalize Proposition 1 and Corollaries 3 and 4 to the trading network setting of Hatfield et al. (2013).
6.1.3 Discussion

*Ex ante* investment incentives have been previously studied in matching market settings different from ours. For instance, Azevedo (2014) considered a stable matching model in which firms choose hiring capacities. Azevedo (2014) showed that firms attain higher profits by systematically underinvesting in capacity relative to the socially optimal level of investment; this finding builds on the well-known result that firms have incentives to engage in capacity manipulation under stable matching mechanisms (Sönmez, 1997). A similar result holds in our setting: As there does not exist a stable and strategy-proof mechanism for firms with multi-unit demand, Theorem 1 implies that firms are not incentivized to make socially efficient investments under any stable mechanism.\(^{39}\)

Also closely related to the labor market matching model we study here is the work of Cole, Mailath, and Postlewaite (2001a), who considered a noncooperative game in which finite sets of firms and workers make investments and then match with each other.\(^{40}\) Cole, Mailath, and Postlewaite (2001a) showed that even in a situation in which agents cannot contract on investment (and thus investment is “sunk” when negotiating the split of surplus), there exists an *ex post contracting equilibrium* (which corresponds to a stable outcome in our labor market matching setting) that achieves a fully efficient outcome, that is, an efficient investment profile and an efficient matching given investment.

In the Cole et al. (2001a) *ex post* contracting equilibrium, both workers *and* firms invest efficiently. This is in sharp contrast to our results, which imply that stable matching mechanisms can induce efficient investment for only one side of the market—it is well-known that no stable mechanism is strategy-proof for both firms and workers so, by Theorem 1, no stable mechanism can induce efficient investment by both firms and workers. The disparity between our findings and those of Cole et al. (2001a) is due to a few differences between the modeling approaches of Cole et al. (2001a) and our work. First, in the setting of Cole et al.

\(^{39}\)The Azevedo (2014) result is not directly comparable to ours as Azevedo assumes a continuum of agents.

\(^{40}\)Cole et al. (2001b) studied a similar setting, but in a large economy with a continuum of agents.
(2001a), types are assumed to be single-dimensional and the surplus function is assumed to be supermodular, while our model places no such structure on types and surplus functions. Perhaps more importantly, Cole et al. (2001a) constructed an equilibrium with full efficiency by specifying a transfer rule that depends on the target equilibrium. While this construction allows for severe punishment of a unilateral deviation from the target investment by either a worker or a firm, it requires that the mechanism designer has precise knowledge of each agent’s cost function. By contrast, our inducement of efficient investment condition requires that a mechanism incentivizes efficient investment regardless of agents’ cost functions. Thus, the requirements we impose on mechanisms are stronger than the restrictions of Cole et al. (2001a); these differences account for the disparity in findings.

6.1.4 Investment Incentives in Labor Markets with Discrete Transfers

We now consider the case in which $P$ is a finite set of points in $\mathbb{R}$. In this case, any stable outcome is Pareto efficient in the ordinal sense. However, stable outcomes need not be efficient in the sense of the present paper—that is, they need not maximize total welfare—as the set of possible salaries is discrete. Hence, our exact equivalence result for efficient mechanisms, Theorem 1, cannot be applied.

However, we show that any stable outcome is approximately efficient, with the degree of approximation determined by the size of the salary increment. We suppose that the salary increment is at most $\epsilon > 0$, i.e., $\bar{p} - p \leq \epsilon$ for any $p, \bar{p} \in P$ such that $\bar{p} = \min\{\bar{p} \in P : \bar{p} > p\}$. As before, we also assume that the domain of salaries is sufficiently large.

Proposition 2. If the set of valuations is substitutable for all $f \in F$, the salary increment is at most $\epsilon$, and the domain of salaries is sufficiently large, then for any stable outcome $(\omega, t)$,

\[ p_{\text{max}}, p_{\text{min}} \in P \text{ such that in any individually rational outcome no worker can be paid either } p_{\text{max}} \text{ or } p_{\text{min}}. \]

\[ ^{41} \text{As in Section 6.1.2, we assume that there are sufficiently high and low wages } p_{\text{max}}, p_{\text{min}} \in P \text{ such that in any individually rational outcome no worker can be paid either } p_{\text{max}} \text{ or } p_{\text{min}}. \]
the alternative $\omega$ is efficient within $|W|\epsilon$, i.e.,

$$V(\omega) + |W|\epsilon \geq \max_{\psi \in \Omega} \{V(\psi)\}.$$ \(^{42}\)

Moreover, when transfers are discrete, Hatfield and Milgrom (2005) showed that the worker-optimal stable mechanism is strategy-proof for workers.\(^{43}\) Combining Proposition 2 with Theorem 3 shows that the worker-optimal stable mechanism induces workers to make approximately efficient investments, with the degree of approximation is determined by the size of the salary increment.

**Corollary 5.** If the set of valuations is substitutable for all $f \in F$, the salary increment is at most $\epsilon$, the domain of salaries is sufficiently large, and $V^w$ is path-connected, then the worker-optimal stable mechanism induces efficient investment within

$$(|F| + 1)|W|\epsilon \leq |I|^2 \epsilon$$

by worker $w \in W$.

Corollary 5 sharpens an analogous result we presented in earlier work, which found that the worker-optimal stable mechanism induces efficient investment by workers within

$$(|F| + 1)|I|\epsilon \geq (|F| + 1)|W|\epsilon \quad \text{(Hatfield, Kojima, and Kominers, 2014).}$$ \(^{44}\)

\(^{42}\)In our companion paper (Hatfield, Kojima, and Kominers (2014)), we proved the weaker bound $V(\omega) + |I|\epsilon \geq \max_{\psi \in \Omega} \{V(\psi)\}$.

\(^{43}\)Hatfield and Milgrom (2005) assumed that the demand of any firm or worker is unique at any price vector in $\Pi$. Hence, we assume for simplicity that the set of workers demanded by firm $f$ is unique for every salary vector $\pi \in \Pi$ and valuation $v^f \in V^f$. However, in order to ensure that the valuation space of each worker $w$ is path-connected, we must allow for the possibility that the difference in $w$’s valuation for particular firms may exactly equal some difference in salary values, thereby creating the possibility of indifferences. Hence, we assume that the mechanism treats each worker as having strict preferences induced by a fixed tie-breaking rule. With these assumptions, the worker-optimal stable mechanism of Hatfield and Milgrom (2005) is strategy-proof for workers and produces a stable outcome.

\(^{44}\)Our work can be extended to a setting in which a matching not only assigns worker–firm pairings but also specifies non-pecuniary contractual terms from a finite set $E$. (For a general discussion of worker–firm matching with contracts, see the work of Crawford and Knoer (1981), Kelso and Crawford (1982), Fleiner (2003), and Hatfield and Milgrom (2005).) In this case, the proofs of Proposition 2 and Corollary 5 can easily be extended to show that the worker-optimal stable mechanism induces efficient investment within $(|F| \cdot |E| + 1)|W|\epsilon$ by workers.
6.2 Computationally Tractable Multiunit Procurement Auctions

It is well-known that computing Vickrey-Clarke-Groves mechanism outcomes is NP-complete in many settings, including that of the multiunit procurement auction (see, e.g., Dash et al. (2003)). However, Kothari, Parkes, and Suri (2003, 2005) provide a computationally tractable mechanism for multiunit procurement. They do so by relaxing both the efficiency and strategy-proofness constraints. Here, we use Theorem 3, combined with the approximation bounds derived by Kothari et al. (2003), to bound the how much the Kothari et al. (2003) mechanism distorts incentives away from ex ante efficient investment.

Consider a procurement auction with an auctioneer $a$ who has a value $U \geq 0$ for obtaining $n \in \mathbb{Z}_{\geq 0}$ units of a good.\footnote{We use $\mathbb{Z}_{\geq 0}$ to denote the set of non-negative integers.} Here, we have $I = \{a\} \cup B$, where $B$ is a set of bidders. The set of alternatives specifies how many units of the good each bidder provides: $\Omega = \{\zeta \in (\mathbb{Z}_{\geq 0})^B : \forall b \in B, \zeta^b \leq n\}$.

The valuation function of the auctioneer for an alternative $\zeta$ is given by

$$v^a(\zeta) = \begin{cases} U & \sum_{b \in B} \zeta^b \geq n \\ 0 & \text{otherwise.} \end{cases}$$

(We assume that the valuation function of the auctioneer is fixed, i.e., $V^a = \{v^a\}$.)

Each bidder $b$ faces a minimal feasible level of production $m^b$, decreasing marginal costs for production, and a capacity constraint $y^b$. More specifically, $V^b$ is a compact, path-connected set of functions $v^b$ from $\Omega$ to $(-\infty, 0]$ such that:

1. Each additional item is costly for bidder $b$ to produce (equivalently, total costs are increasing in the number of goods produced), i.e., $v^b(\zeta) \geq v^b(\zeta^b + 1, \zeta^{B\setminus\{b\}})$ for all $\zeta \in \Omega$ such that $m^b \leq \zeta^b < y^b$.

2. Marginal costs are decreasing for bidder $b$ (from, e.g., economies of scale or learning-by-doing), that is, each additional unit is less costly to produce than the one before—
\[ v^b(\zeta + 1, \zeta^B \setminus \{b\}) - v^b(\zeta) \geq v^b(\zeta^b, \zeta^B \setminus \{b\}) - v^b(\zeta^b - 1, \zeta^B \setminus \{b\}) \text{ for all } \zeta \in \Omega \text{ such that } \]
\[ m^b < \zeta^b < y^b. \]

3. Bidder \( b \in B \) faces a minimal feasible production constraint \( m^b \)—any positive level of production below \( m^b \) is more costly than the total value in the economy—\( v^b(\zeta) = -2U \) if \( 0 < \zeta^b < m^b \).

4. Bidder \( b \in B \) faces a capacity constraint \( y^b \)—any level of production above \( y^b \) is more costly than the total value in the economy—\( v^b(\zeta) = -2U \) if \( \zeta^b > y^b \).

5. The valuation of bidder \( b \) depends only on that bidder’s own production, i.e., \( v^b(\zeta) = v^b(\theta) \) whenever \( \zeta^b = \theta^b \).

6. The valuation of bidder \( b \) is normalized so that producing 0 units of the good gives value 0, i.e., \( v^b(\zeta) = 0 \) for all \( \zeta \) such that \( \zeta^b = 0 \).

Kothari et al. (2003, 2005) provide a mechanism for this procurement auction environment that is computationally tractable, approximately strategy-proof, and approximately efficient.46

**Proposition 3** (Kothari et al. (2003, 2005)). There exists a multiunit procurement auction mechanism that is

1. efficient within \( \left( \frac{\epsilon}{1 + \epsilon} \right) U \) and
2. strategy-proof within \( \left( \frac{\epsilon}{1 + \epsilon} \right) U \) for each bidder \( b \in B \),

and has computational complexity \( O\left( \frac{|B|^3}{\epsilon} \log \left( \frac{|B|}{\epsilon} \right) \right) \).

Combining Proposition 3 with our Theorem 3 and Lemma 2, we obtain the following result bounding the distortion in bidders’ *ex ante* investment incentives under the Kothari et al. (2003, 2005) mechanism.

**Corollary 6.** The Kothari et al. (2003, 2005) multiunit procurement auction mechanism induces efficient investment within \( 2(n + 1) \left( \frac{\epsilon}{1 + \epsilon} \right) U \) by each bidder.

---

46The Kothari et al. (2003, 2005) result also relies on finding an upper bound for the ratio of the cost of production for a subset of \(|B| - 1\) bidders to the cost of production for the entire set of bidders.
6.3 Uniform-Price Auctions

We now consider a uniform-price auction setting in which there are $k$ identical objects for sale; moreover, we let agents have multi-unit demand. In such an auction, each agent may demand multiple objects and thus reports a multi-dimensional vector of valuations: the value for the first object received, the value for the second object received, and so forth. The auction then distributes the objects in an efficient manner with respect to the reported valuations; the price charged for each object is equal to the $k$th-highest value for an object reported by the agents. The uniform-price auction is efficient. However, the uniform-price auction is not strategy-proof, as the price for an object obtained by an agent may depend on his report. Nevertheless, in this section, we show that the uniform-price auction becomes approximately strategy-proof in expectation as the market grows large, and thus also approximately induces efficient investment as the market grows large.

In the uniform-price auction setting, each agent $i \in I$ is a buyer. We assume that each agent desires at most $M$ objects, where $M$ is a fixed integer. Moreover, we assume that for each agent, the marginal utility of each object is always a value in $[0, 1]$, and that each agent’s utility is (weakly) concave in the number of objects. There are $k(I)$ total objects available; we assume that the total number of objects is no more than the number of objects (weakly) demanded, i.e., $|I|M \geq k(I)$, but put no further restrictions on how $k(I)$ varies with $I$. Thus, the space of alternatives is given by $\Omega = \{\omega \in \{0, 1, \ldots, M\}^I : \sum_{i \in I} \omega^i = k(I)\}$. A valuation function for agent $i$ is given by

$$v^i(\omega) = \sum_{m=1}^{M} a^i_m \mathbb{1}_{\{\omega^i \geq m\}}, \quad (4)$$

where $a^i_m$ represents the value of the $m^{th}$ object to $i$. Hence, the space of valuation functions for $i$ is given by

$$V^i = \left\{v^i(\omega) = \sum_{m=1}^{M} a^i_m \mathbb{1}_{\{\omega^i \geq m\}} : 0 \leq a^i_M \leq \cdots \leq a^i_1 \leq 1 \right\}.$$
For a given \( v^i \in V^i \), we let \( a_m(v^i) \) be the associated value of the \( m^{th} \) object to \( i \) under \( v^i \).

The space of valuation distributions for each agent \( i \) is then given by \( V^i \equiv \Delta(V^i) \). Since the outcome of the uniform-price auction only depends on agents' expected values, it is helpful to define, for each \( v^i \in V^i \), the associated vector of expected values \( (a_1(v^i), \ldots, a_M(v^i))_{i \in I} \), where

\[
a_m(v^i) \equiv \int a_m(v^i) \, dv^i(v^i),
\]

i.e., \( a_m(v^i) \) is the expected value of the \( m^{th} \) object for \( i \) under \( v^i \); we define \( a_0(v^i) \equiv 1 \) and \( a_{M+1}(v^i) \equiv 0 \) for all \( i \in I \).

Our \textit{uniform-price auction} mechanism has an allocation rule \( \mu \) that allocates the \( k(I) \) objects to their highest-value uses, with ties across agents broken arbitrarily. Formally, to define an allocation rule \( \mu \) for a uniform-price auction, we let the \textit{auction price} \( p(v) \) be the \( k(I)^{th} \)-largest element of \( (a_1(v^i), \ldots, a_M(v^i))_{i \in I} \), and let \( \mu(v) \) be such that

1. \( \sum_{i \in I} \mu^i(v) = k(I) \),
2. for all \( i \in I \), \( a_{\mu^i(v)}(v^i) \geq p(v) \), and
3. for all \( i \in I \), \( a_{\mu^i(v)+1}(v^i) \leq p(v) \).

The first condition on \( \mu \) ensures that exactly \( k(I) \) objects are distributed; the second condition ensures that the value of each object received by each agent (if any) is at least \( p(v) \); and the third condition ensures that the value of an additional object (beyond what is allocated by the mechanism) to each agent is no more than \( p(v) \). The transfer rule \( r \) of the uniform-price auction is given by

\[
r^i(v) = p(v)\mu^i(v);
\]

that is, each agent pays \( p(v) \) for each object he receives.

To define the set \( V^i \) of investment choices for \( i \), consider a set of distributions \( \mathcal{F} \), where each \( F \in \mathcal{F} \) has a corresponding density function \( f \) over \([0,1] \); we assume that

\[
\inf_{F \in \mathcal{F}, x \in [0,1]} \{ f(x) \} > 0.
\]

An investment choice \( v^i \) then corresponds to a vector \((f_1, f_2, \ldots, f_M), \)
where each distribution $f_m$ is chosen from $\mathcal{F}$; letting $z^i_m$ be the value drawn by $f_m$, we construct the valuation function $v^i$ of agent $i$ according to (4), where $a^i_\ell$ is the $\ell$th-largest element of $(z^i_1, z^i_2, \ldots, z^i_M)$. (The associated valuation distribution is then constructed by letting $v^i = \delta_{v^i}$.)

We first show that a uniform-price auction becomes approximately strategy-proof in expectation as the market grows large. Intuitively, as the number of agents increases, the amount any given agent can move the price becomes smaller (in expectation): In a uniform-price auction, the price is given by the $k(I)$th-largest reported value, so the largest effect that an individual agent can have on the price is to increase it from the $k(I)$th-largest value reported by other agents to the $(k(I) - M)$th-largest value reported by other agents. But, as the number of agents grows large, the expected difference between the $k(I)$th-largest order statistic and the $(k(I) - M)$th-largest order statistic becomes small. Thus, as the number of agents becomes large, from the perspective of an individual agent the uniform-price auction is as if the price that agent faces does not depend on the report of that agent, and thus the uniform-price auction becomes approximately strategy-proof in expectation.

Proposition 4. For every $\delta > 0$, there exists a constant $K$ such that, for all sets of agents $I$, the uniform-price auction is strategy-proof within $\frac{K}{|I|^{1-\delta}}$ in expectation for each agent $i \in I$. In particular, for any $\epsilon > 0$, the uniform-price auction is strategy-proof within $\epsilon$ in expectation for each agent for every sufficiently large $|I|$.

Combining Proposition 4 with Theorem 7, and noting that the uniform-price auction is efficient, we have the following corollary.

Corollary 7. For every $\delta > 0$, there exists a constant $K$ such that, for all sets of agents $I$, the uniform-price auction induces efficient investment within $\frac{K}{|I|^{1-\delta}}$ for each agent $i \in I$. In particular, for any $\epsilon > 0$, the uniform-price auction induces efficient investment within $\epsilon$ for each agent for every sufficiently large $|I|$.

Friedman (1960, 1991) famously advocated using uniform-price auctions instead of “pay-
as-bid” auctions in which buyers pay their bids for each unit they obtain. Friedman’s (1960, 1991) argument relies on the claim that the former auction induces truth-telling by bidders while the latter auction does not; Proposition 4 provides one formalization of this claim, by showing that the uniform-price auction is approximately strategy-proof in expectation (even though it is not exactly strategy-proof). Our results show that there is an additional benefit of using uniform-price auctions, as stated in Corollary 7: the uniform-price auction is not only nearly strategy-proof in the limit, but also provides nearly efficient incentives for investment in the limit. Thus, our work complements and strengthens Friedman’s argument in favor of the uniform-price auction.

6.4 Double Auctions

In this section we study investment incentives in double auctions. Our formal discussion focuses on a variant of the double auction mechanism introduced by McAfee (1992). In our auction, each buyer reports his value for an object, and each seller reports his cost of supplying an object. The auction then executes every non-negative surplus trade but one that is “marginal,” in the sense that it generates the least surplus of all non-negative surplus trades. Our double auction is strategy-proof; however, it is not perfectly efficient, as it fails to execute one nonnegative-surplus trade. We show, however, that our double auction becomes nearly efficient in expectation—and thus also nearly induces efficient investment—as the market grows large.48

Now, we suppose that there are finite sets $B$ and $S$ of buyers and sellers, respectively; the set of agents is $I = B \cup S$. Each buyer $b \in B$ has unit demand; each seller $s \in S$ has unit supply; and goods are homogenous. The set of alternatives takes the form $\Omega = \{\omega \in$

47 We are far from being the first to recognize that the uniform-price auction is roughly incentive compatible in large markets—indeed, Friedman’s original argument is based on an intuition to this effect. The work of Azevedo and Budish (2015) provides a recent formal analysis of the uniform-price auction’s strategic properties, showing that the uniform-price auction is strategy-proof in the large.

48 Our auction is a simplified version of the McAfee (1992) double auction. Our double auction mechanism is less efficient than that of McAfee (1992); thus, our conclusions hold for the McAfee (1992) auction as well.
\( \varphi(B \cup S) : |B \cap \omega| = |S \cap \omega| \), for a given \( \omega \in \Omega \), having \( b \in B \cap \omega \) represents that buyer \( b \) obtains a good, and having \( s \in S \cap \omega \) represents that seller \( s \) supplies the good.

We assume that each buyer’s value for a good is in \([0, 1]\), and, similarly, each seller’s cost of supplying a good is in \([0, 1]\). Thus, a valuation function for buyer \( b \in B \) is given by

\[
v^b(\omega) = a^b \mathbb{1}_{\{b \in \omega\}},
\]

where \( a^b \) represents the value of a good to \( b \). The space of valuation functions for \( b \) is then given by

\[
V^b = \{ v^b(\omega) = a^b \mathbb{1}_{\{b \in \omega\}} : 0 \leq a^b \leq 1 \}.
\]

Similarly, a valuation function for seller \( s \in S \) is given by

\[
v^s(\omega) = -e^s \mathbb{1}_{\{s \in \omega\}},
\]

where \( e^s \) represents the cost of supplying the good for \( s \). The space of valuation functions for \( s \) is then given by

\[
V^s = \{ v^s(\omega) = -e^s \mathbb{1}_{\{s \in \omega\}} : 0 \leq e^s \leq 1 \}.
\]

The space of valuation distributions for each agent \( i \) is \( \mathcal{V}^i \equiv \Delta(V^i) \). Since the outcome of the double auction only depends on agents’ expected values, it is helpful to define, for each \( b \in B \) and \( v^b \in \mathcal{V}^b \), the associated expected value \( a(v^b) \) where

\[
a(v^b) \equiv \int a(v^b) \, dv^b(v^b),
\]

i.e., \( a(v^b) \) is the expected value of the good for \( b \) under \( v^b \). We introduce analogous notation for sellers:

\[
e(v^s) \equiv \int e(v^s) \, dv^s(v^s).
\]

\(^{49}\)Recall that \( \varphi(A) \) denotes the powerset of \( A \).
Our double auction mechanism has an allocation rule $\mu$ defined by the following rule: let $a^{(\ell)}$ be the $\ell^{th}$-highest value in $(a(v^b))_{b \in B}$, and let $e^{(\ell)}$ be the $\ell^{th}$-lowest value in $(e(v^s))_{s \in S}$, and define $a^{(0)} = 1$ and $e^{(0)} = 0$. Letting $k$ be the largest integer such that $a^{(k)} \geq e^{(k)}$, let $\hat{b}$ be an arbitrary buyer such that $a(v^{\hat{b}}) = a^{(k)}$ and $\hat{s}$ be an arbitrary seller such that $e(v^{\hat{s}}) = e^{(k)}$. We then let $\mu(v)$ be the alternative such that

1. $b \in \mu(v)$ if and only if $a(v^b) \geq a^{(k)}$ and $b \neq \hat{b}$, and

2. $s \in \mu(v)$ if and only if $e(v^s) \leq e^{(k)}$ and $s \neq \hat{s}$.

The first condition states that every buyer who values the good at least as much as $a^{\hat{b}}$ (except for $\hat{b}$) receives a good, and the second states that every seller whose cost is no more than $e^{\hat{s}}$ (except for $\hat{s}$) supplies a good. The transfer rule is then given by

$$r^b(v) = a^{\hat{b}} \mathbb{1}_{\{b \in \mu(v)\}}$$

$$r^b(v) = -e^{\hat{s}} \mathbb{1}_{\{s \in \mu(v)\}};$$

that is, each buyer who obtains a good pays $a^{\hat{b}}$ and each seller who supplies a good is paid $e^{\hat{s}}$.

To define the set $V^b$ of investment choices for $b \in B$, consider a set of distributions $\mathcal{F}$, where each $F \in \mathcal{F}$ has a corresponding density function $f$ over $[0,1]$; we assume that $\inf_{F \in \mathcal{F}, x \in [0,1]} \{f(x)\} > 0$. An investment choice $\nu^i$ then corresponds to a distribution $f$ chosen from $\mathcal{F}$; letting $a^b$ be the value drawn by $f$, the corresponding valuation function $v^b(\omega) = a^b \mathbb{1}_{\{b \in \omega\}}$. (The associated valuation distribution is then constructed by taking $v^b = \delta_{a^b}$.)

Similarly, to define the set $V^s$ of investment choices for $s \in S$, consider a set of distributions $\mathcal{G}$, where each $G \in \mathcal{G}$ has a corresponding density function $g$ over $[0,1]$; we assume that $\inf_{G \in \mathcal{G}, x \in [0,1]} \{g(x)\} > 0$. An investment choice $\nu^s$ then corresponds to a distribution $g$ chosen from $\mathcal{G}$; letting $e^s$ be the value drawn by $g$, the corresponding valuation function $v^s(\omega) = -e^s \mathbb{1}_{\{s \in \omega\}}$. (The associated valuation distribution is then constructed by taking $v^s = \delta_{e^s}$.)
We first show that our double auction becomes nearly efficient in expectation as the market grows large. Intuitively, as the number of agents increases, the surplus of the one unexecuted non-negative surplus trade grows smaller in expectation. Note that this claim only holds in expectation: For any number of agents, there always exists a realization such that the efficiency loss from omitting the marginal transaction is large (in the sense that the surplus lost by omitting that trade is on same scale as the average agent’s returns from entering the market); however, the chance of such an adverse realization approaches 0 as the number of agents grows large.

**Proposition 5.** For every $\delta > 0$, there exists a constant $K$ such that, for all sets of agents $I$, the double auction is efficient within $K\left(\frac{1}{|B|^{1-\delta}} + \frac{1}{|S|^{1-\delta}}\right)$ in expectation. In particular, for any $\epsilon > 0$, the double auction is efficient within $\epsilon$ in expectation for every sufficiently large $|B|$ and $|S|$.

Proposition 5 shows that our double auction becomes nearly efficient in expectation as the market grows; in particular, we obtain a bound on the efficiency loss as a function of market size that goes to 0 as the market grows.

Moreover, our double auction is strategy-proof: As the price is set according to the value of the marginal, non-executed trade, no buyer who receives a good can affect his price by mis-reporting his value, and any buyer who does not receive a good would prefer not to buy at the prevailing price. Meanwhile, if the marginal buyer were to increase his reported value, then he would face trade at a price weakly higher than his value—and thus a (weakly) negative payoff. An analogous argument shows that no sellers would like to misreport their cost of supplying goods.

Combining the preceding observations with Theorem 7, we obtain the following corollary.

**Corollary 8.** For every $\delta > 0$, there exists a constant $K$ such that, for all sets of agents $I$, the double auction induces efficient investment within $K\left(\frac{1}{|B|^{1-\delta}} + \frac{1}{|S|^{1-\delta}}\right)$ for each agent $i \in I$.

---

50Proposition 5 is related to Theorem 3 of McAfee (1992), which states that when the buyers’ and sellers’ valuations are conditionally independent and identically distributed, there exists a constant $K$ such that the expected efficiency loss of the McAfee (1992) double auction is at most $K\left(\frac{1}{|B|} + \frac{1}{|S|}\right)$. 

50
In particular, for any $\epsilon > 0$, the double auction induces efficient investment within $\epsilon$ for each agent for every sufficiently large $|B|$ and $|S|$.

Corollary 8 demonstrates that double auctions induce nearly efficient investment in large markets; in particular, we obtain a bound on the distortion of agents’ ex ante investment incentives away from the social optimum as a function of market size that goes to 0 as the market grows. Thus, buyers in double auctions are encouraged to make efficient investments in complementary goods; similarly, sellers are encouraged to efficiently invest in production technologies.

This section focused on a variant of the McAfee (1992) double auction mechanism, in order to highlight the role of approximate—as opposed to exact—efficiency in expectation on agents’ incentives to invest. Under more standard double auction mechanisms, such as the “buyer’s bid” double auction or the more general “$k$-double auction,” all mutually profitable trades (with respect to reported valuations) take place, leading to exactly efficient outcomes; however, such double auction mechanisms are not strategy-proof. These mechanisms do, however, become nearly strategy-proof in expectation as the number of market participants grows sufficiently large. Therefore, these standard double auction mechanisms also nearly induce efficient investment in large markets, just as the McAfee (1992) double auction mechanism does.

7 Conclusion

“Human capital analysis starts with the assumption that individuals decide on their education, training, medical care, and other additions to knowledge and health by weighing the benefits and costs.”

—Gary S. Becker (1993)

Human capital theory emphasizes the incentives markets create for investment in education and skills. Much of market and mechanism design, however, focuses only on the
market-clearing stage, treating agents’ types—which in practice are determined by \textit{ex ante} investments—as exogenously given.

Our results show that incentives (or lack thereof) to make efficient \textit{ex ante} investments are closely linked to how much of the agent’s marginal social contributions accrue to her at the market-clearing stage. Moreover, these properties are closely related to (the degrees of deviation from) \textit{ex post} efficiency and strategy-proofness. The linkages between incentivizing \textit{ex ante} efficient investment, rewarding agents with their marginal contributions, and incentivizing truthful revelation hold even when the returns to investment are uncertain, and extend to settings where our key conditions hold only partially.

Our work implies that a mechanism that achieves two key goals of labor market design—encouraging efficiency and strategy-proofness—perforce also incentivizes nearly efficient human capital acquisition. Likewise, nearly efficient and strategy-proof market-clearing mechanisms induce nearly efficient investment in production capacity, incentivize nearly efficient information acquisition, and promote nearly efficient network formation. Thus, we see that recognizing the importance of \textit{ex ante} efficient behavior further strengthens the case for designing mechanisms that achieve the classical economic ideal of an efficient market free from strategic gaming.
References


A Alternate Investment Efficiency Conditions

A.1 Limit-Inducing Efficient Investment

We restate here our alternate definition of inducing efficient investment given in Footnote 11.

Definition A.1. A mechanism $\mathcal{M}$ limit-Induces efficient investment by $i \in I$ if, for all $v^{I \setminus \{i\}} \in V^{I \setminus \{i\}}$ and any cost function $c^i : V^i \to \mathbb{R}$, for any sequence of valuations $\{\bar{v}_{\ell}^i\}_{\ell=1}^\infty$, the sequence $\{u^i(\mathcal{M}(\bar{v}_{\ell}^i, v^{I \setminus \{i\}}); \bar{v}_{\ell}^i) - c^i(\bar{v}_{\ell}^i)\}_{\ell=1}^\infty$ approaches

$$\underset{\bar{v}^i \in V^i}{\sup} \left\{ u^i(\mathcal{M}(\bar{v}^i, v^{I \setminus \{i\}}); \bar{v}^i) - c^i(\bar{v}^i) \right\}$$

if and only if the sequence $\{V(\mathcal{M}(\bar{v}_{\ell}^i, v^{I \setminus \{i\}}); (\bar{v}_{\ell}^i, v^{I \setminus \{i\}})) - c^i(\bar{v}_{\ell}^i)\}_{\ell=1}^\infty$ approaches

$$\underset{\bar{v}^i \in V^i}{\sup} \left\{ V(\mathcal{M}(\bar{v}^i, v^{I \setminus \{i\}}); (\bar{v}^i, v^{I \setminus \{i\}})) - c^i(\bar{v}^i) \right\}.$$

Proposition A.1. A mechanism induces efficient investment by $i$ if and only if it limit-Induces efficient investment by $i$.

Proof. First, suppose that a mechanism $\mathcal{M}$ limit-Induces efficient investment by $i$. Fix $v^{I \setminus \{i\}} \in V^{I \setminus \{i\}}$ and $c^i$ arbitrarily. Let

$$\hat{v}^i \in \arg\max_{\bar{v}^i \in V^i} \left\{ u^i(\mathcal{M}(\bar{v}^i, v^{I \setminus \{i\}}); \bar{v}^i) - c^i(\bar{v}^i) \right\},$$

and consider a sequence of valuations $\{\bar{v}_{\ell}^i\}_{\ell=1}^\infty$ such that $\bar{v}_{\ell}^i = \hat{v}^i$ for all $\ell$, that is, a constant sequence at $\hat{v}^i$. Then the sequence $\{u^i(\mathcal{M}(\bar{v}_{\ell}^i, v^{I \setminus \{i\}}); \bar{v}_{\ell}^i) - c^i(\bar{v}_{\ell}^i)\}_{\ell=1}^\infty$ trivially approaches

$$\underset{\bar{v}^i \in V^i}{\sup} \left\{ u^i(\mathcal{M}(\bar{v}^i, v^{I \setminus \{i\}}); \bar{v}^i) - c^i(\bar{v}^i) \right\}.$$ This and the assumption that $\mathcal{M}$ limit-Induces efficient investment by $i$ imply that the sequence $\{V(\mathcal{M}(\bar{v}_{\ell}^i, v^{I \setminus \{i\}}); (\bar{v}_{\ell}^i, v^{I \setminus \{i\}})) - c^i(\bar{v}_{\ell}^i)\}$ converges to $\underset{\bar{v}^i \in V^i}{\sup} \left\{ V(\mathcal{M}(\bar{v}^i, v^{I \setminus \{i\}}); (\bar{v}^i, v^{I \setminus \{i\}})) - c^i(\bar{v}^i) \right\}$. By the construction of the sequence $\{\bar{v}_{\ell}^i\}_{\ell=1}^\infty$, the sequence $\{V(\mathcal{M}(\bar{v}_{\ell}^i, v^{I \setminus \{i\}}); (\bar{v}_{\ell}^i, v^{I \setminus \{i\}})) - c^i(\bar{v}_{\ell}^i)\}$ is a constant sequence, and thus
Together, it follows that
\[ V(\mathcal{M}(\tilde{v}^i, u^{I\setminus\{i\}}); (\hat{v}^i, v^{I\setminus\{i\}})) - c^i(\hat{v}^i) = \sup_{\tilde{v}^i \in V^i} \left\{ V(\mathcal{M}(\tilde{v}^i, u^{I\setminus\{i\}}); (\tilde{v}^i, v^{I\setminus\{i\}})) - c^i(\tilde{v}^i) \right\}, \]
which shows that
\[ \hat{v}^i \in \arg \max_{\tilde{v}^i \in V^i} \left\{ V(\mathcal{M}(\tilde{v}^i, u^{I\setminus\{i\}}); (\tilde{v}^i, v^{I\setminus\{i\}})) - c^i(\tilde{v}^i) \right\}. \]
Thus, we have
\[ \arg \max_{\tilde{v}^i \in V^i} \left\{ u^i(\mathcal{M}(\tilde{v}^i, u^{I\setminus\{i\}}); \tilde{v}^i) - c^i(\tilde{v}^i) \right\} \subseteq \arg \max_{\tilde{v}^i \in V^i} \left\{ V(\mathcal{M}(\tilde{v}^i, u^{I\setminus\{i\}}); (\tilde{v}^i, v^{I\setminus\{i\}})) - c^i(\tilde{v}^i) \right\}. \]
Using an analogous argument, we find that
\[ \arg \max_{\tilde{v}^i \in V^i} \left\{ u^i(\mathcal{M}(\tilde{v}^i, u^{I\setminus\{i\}}); \tilde{v}^i) - c^i(\tilde{v}^i) \right\} \supseteq \arg \max_{\tilde{v}^i \in V^i} \left\{ V(\mathcal{M}(\tilde{v}^i, u^{I\setminus\{i\}}); (\tilde{v}^i, v^{I\setminus\{i\}})) - c^i(\tilde{v}^i) \right\}. \]
Together, (5) and (6) imply that \( \mathcal{M} \) induces efficient investment by \( i \).

Second, suppose that mechanism \( \mathcal{M} \) induces efficient investment by \( i \). Then by Lemma 1, \( \mathcal{M} \) provides marginal rewards to \( i \). Therefore, for any \( \tilde{v}^i \) and an arbitrarily fixed \( v \in V \), we have
\[
\begin{align*}
&u^i(\mathcal{M}(\tilde{v}^i, u^{I\setminus\{i\}}); \tilde{v}^i) - c^i(\tilde{v}^i) \\
&= \left( u^i(\mathcal{M}(\tilde{v}^i, u^{I\setminus\{i\}}); v^i) - u^i(\mathcal{M}(v^i, u^{I\setminus\{i\}}); v^i) - c^i(\tilde{v}^i) \right) + u^i(\mathcal{M}(v^i, u^{I\setminus\{i\}}); v^i) \\
&= V(\mathcal{M}(\tilde{v}^i, v^{I\setminus\{i\}}); (\tilde{v}^i, v^{I\setminus\{i\}})) - V(\mathcal{M}(v^i, v^{I\setminus\{i\}}); (v^i, v^{I\setminus\{i\}})) - c^i(\tilde{v}^i) + u^i(\mathcal{M}(v^i, u^{I\setminus\{i\}}); v^i) \\
&= V(\mathcal{M}(\tilde{v}^i, v^{I\setminus\{i\}}); (\tilde{v}^i, v^{I\setminus\{i\}})) - c^i(\tilde{v}^i) + \left( u^i(\mathcal{M}(v^i, u^{I\setminus\{i\}}); v^i) - V(\mathcal{M}(v^i, v^{I\setminus\{i\}}); (v^i, v^{I\setminus\{i\}})) \right),
\end{align*}
\]
where the second equality follows from the fact that \( \mathcal{M} \) provides marginal rewards. This
equality then implies that $u^i(M(\bar{v}^i, v^{I-\{i\}}); \bar{v}^i) - c^i(\bar{v}^i)$ and $V(M(\bar{v}^i, v^{I-\{i\}}); (\bar{v}^i, v^{I-\{i\}})) - c^i(\bar{v}^i)$ are different by a constant amount. Thus, we have that the requirement for limit-inducement of efficient investment holds.

\[ \]

A.2 Making Efficient Investment Satisfice

**Definition A.2.** A mechanism $M$ makes efficient investment satisfice within $\epsilon$ for $i \in I$ if, for all $v^{I-\{i\}} \in V^{I-\{i\}}$, if

\[
\hat{v}^i \in \arg \max_{\tilde{v}^i \in V^i} \left\{ V(M(\tilde{v}^i, v^{I-\{i\}}); (\bar{v}^i, v^{I-\{i\}})) - c^i(\bar{v}^i) \right\},
\]

then

\[
u^i(M(\hat{v}^i, v^{I-\{i\}}); \hat{v}^i) - c^i(\hat{v}^i) + \epsilon \geq \sup_{\tilde{v}^i \in V^i} \left\{ u^i(M(\tilde{v}^i, v^{I-\{i\}}); \tilde{v}^i) - c^i(\tilde{v}^i) \right\}\]

for all cost functions $c^i : V^i \to \mathbb{R}$.

**Proposition A.2.** A mechanism makes efficient investment satisfice within $\epsilon$ for $i \in I$ if and only if it induces efficient investment within $\epsilon$ by $i$.

**Proof.** By Lemma 2, it suffices to show that a mechanism $M$ makes efficient investment satisfice within $\epsilon$ for $i$ if and only if it provides marginal rewards within $\epsilon$ to $i$.

First, we show that any mechanism $M$ that provides marginal rewards within $\epsilon$ will make efficient investment satisfice within $\epsilon$. Fix $v^{I-\{i\}} \in V^{I-\{i\}}$, fix a cost function $c^i$, and consider\(^{51}\)

\[
\hat{v}^i \in \arg \max_{\tilde{v}^i \in V^i} \left\{ V(M(\tilde{v}^i, v^{I-\{i\}}); (\bar{v}^i, v^{I-\{i\}})) - c^i(\bar{v}^i) \right\},
\]

Then, we have

\[
V(M(\hat{v}^i, v^{I-\{i\}}); (\hat{v}^i, v^{I-\{i\}})) - c^i(\bar{v}^i) \geq V(M(v^i, v^{I-\{i\}}); (v^i, v^{I-\{i\}})) - c^i(v^i) \tag{7}
\]

\(^{51}\)Note that if the set of maximizers is empty, then the condition of Definition A.2 is vacuously satisfied.
for all $v^i \in V^i$. Since $\mathcal{M}$ provides marginal rewards within $\epsilon$,

$$
\epsilon \geq \left( u^i(\mathcal{M}(v); v^i) - u^i(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); \hat{v}^i)) \right) - \left( V(\mathcal{M}(v); v) - V(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); (\hat{v}^i, v^{I\setminus\{i\}})) \right).
$$

Hence, we have

$$
u^i(\mathcal{M}(\hat{v}^i, v^{I\setminus\{i\}}); \hat{v}^i) - V(\mathcal{M}(\hat{v}^i, v^{I\setminus\{i\}}); (\hat{v}^i, v^{I\setminus\{i\}})) + \epsilon \geq u^i(\mathcal{M}(v); v^i) - V(\mathcal{M}(v); v).
$$

Combining this with (7), we obtain

$$
u^i(\mathcal{M}(\hat{v}^i, v^{I\setminus\{i\}}); \hat{v}^i) - c^i(\hat{v}^i) + \epsilon \geq u^i(\mathcal{M}(v); v^i) - c^i(v^i).
$$

Hence, $\mathcal{M}$ makes efficient investment satisfice within $\epsilon$ for $i$.

Second, we show that if $\mathcal{M}$ makes efficient investment satisfice within $\epsilon$ for $i$, then $\mathcal{M}$ provides marginal rewards within $\epsilon$ to $i$: Suppose that $\mathcal{M}$ does not provide marginal rewards within $\epsilon$ to $i$. Then, there exist $v \in V$ and $\bar{v}^i \in V^i$ such that

$$
V(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); (\bar{v}^i, v^{I\setminus\{i\}})) - V(\mathcal{M}(v); v) > u^i(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); \bar{v}^i) - u^i(\mathcal{M}(v); v^i) + \epsilon. \quad (8)
$$

We now consider the cost function

$$
c^i(\bar{v}^i) = \begin{cases} 
V(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); (\bar{v}^i, v^{I\setminus\{i\}})) - \delta & \text{if } \bar{v}^i = \bar{v}^i \\
V(\mathcal{M}(v); v) & \text{if } \bar{v}^i = v^i \\
\bar{A} & \text{otherwise},
\end{cases}
$$

where $\delta > 0$ and $\bar{A} = 1 + \sup_{\bar{v}^i \in V^i} \{V(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); (\bar{v}^i, v^{I\setminus\{i\}}))\}. \quad 52$

We have

$$
\{\bar{v}^i\} = \arg\max_{\bar{v}^i \in V^i} \left\{ V(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); (\bar{v}^i, v^{I\setminus\{i\}})) - c^i(\bar{v}^i) \right\}. 
$$

\[52\text{Note that } \bar{A} \text{ is finite, because } \Omega \text{ is finite and } V^i \text{ is compact for each } i.\]
However, for $\delta$ sufficiently small, we have

$$u^i(\mathcal{M}(\bar{v}^i, v^{I \setminus \{i\}}); \bar{v}^i) - c^i(\bar{v}^i) + \epsilon < \sup_{\bar{v}^i \in V_i} \left\{ u^i(\mathcal{M}(\bar{v}^i, v^{I \setminus \{i\}}); \bar{v}^i) - c^i(\bar{v}^i) \right\},$$

as we have

$$u^i(\mathcal{M}(\bar{v}^i, v^{I \setminus \{i\}}); \bar{v}^i) - c^i(\bar{v}^i) + \epsilon$$

$$= u^i(\mathcal{M}(\bar{v}^i, v^{I \setminus \{i\}}); \bar{v}^i) - V(\mathcal{M}(\bar{v}^i, v^{I \setminus \{i\}}); (\bar{v}^i, v^{I \setminus \{i\}})) + \delta + \epsilon$$

$$< u^i(\mathcal{M}(v); v^i) - V(\mathcal{M}(v); v)$$

$$= u^i(\mathcal{M}(v); v^i) - c^i(v^i).$$

(9)

The first line of (9) follows from the definition of the cost function $c^i$; the second line follows from (8) (given $\delta$ sufficiently small); and the final line follows from the definition of the cost function $c^i$. Thus, we see that $\mathcal{M}$ does not make efficient investment satisfice within $\epsilon$ for $i$. \hfill \Box

B Examples Omitted from the Main Text

B.1 A Simple Example: The Second-Price Auction

Consider a standard second-price auction for one item, where the set of agents $I$ is the set of bidders for that single item. An alternative represents the assignment of the item to one of the agents $i \in I$; hence, the space of alternatives $\Omega$ is isomorphic to $I$, which we denote as $\Omega \cong I$. The valuation function of each agent is simply the value of the item to that agent, a real number between 0 and 1, so that $V^i = \{a \mathbb{1}_{\omega=i} : a \in [0, 1]\}$ for each $i \in I$.\footnote{For any logical proposition $p$, the indicator function $\mathbb{1}_{\{p\}}$ is defined by

$$\mathbb{1}_{\{p\}} = \begin{cases} 1 & \text{p is true} \\ 0 & \text{p is false}. \end{cases}$$}
As the second-price auction is a Groves mechanism, it is easy to see that the second-price auction provides marginal rewards: If the value of the item increases for agent \( i \), the payoff of \( i \) does not increase unless \( i \) is (or becomes) the winning bidder. However, if \( i \) is the winning bidder, any increase in the value of the item for \( i \) increases the utility of \( i \) by exactly the same amount. Lemma 1 then implies that the second-price auction will lead agents to invest efficiently given the investments of other agents.\(^{54}\)

We consider a simple case with two bidders, \( i \) and \( j \). We suppose that agent \( i \) has the cost function
\[
c^i(a \mathbb{1}_{\omega=i}) = a^2,
\]
while agent \( j \) has the cost function
\[
c^j(v^j) = \begin{cases} 0 & v^j(\omega) = \bar{a}^j \mathbb{1}_{\omega=j} \\ 2017 & \text{otherwise}, \end{cases}
\]
where \( \bar{a}^j \) is a constant in \([0, 1]\) and 2017 can be replaced with any sufficiently large constant.

It is clear from the construction of \( c^j \) that agent \( j \) will always choose the valuation \( \bar{a}^j \mathbb{1}_{\omega=j} \). The optimal choice of the valuation function for agent \( i \) is \( v^i(\omega) = \frac{1}{2} \mathbb{1}_{\omega=i} \) if \( \bar{a}^j < \frac{1}{4} \); in this case, \( i \) will have to pay \( \bar{a}^j < \frac{1}{4} \) in the auction and an investment cost of \( \frac{1}{4} \), but will obtain a gross profit of \( \frac{1}{2} \), which is better for \( i \) than choosing \( v^i(\omega) = 0 \).\(^{55}\) This investment choice is also \( \text{ex ante} \) socially optimal when \( \bar{a}^j < \frac{1}{4} \): Choosing \( v^i(\omega) \notin \{0, \frac{1}{2} \mathbb{1}_{\omega=i}\} \) or \( v^j(\omega) \notin \{\bar{a}^j \mathbb{1}_{\omega=j}\} \) is never socially efficient,\(^{56}\) and when \( \bar{a}^j < \frac{1}{4} \) we have
\[
\left(\frac{1}{2} \cdot 1 - \frac{1}{4}\right) + \left(\bar{a}^j \cdot 0 - 0\right) = \frac{1}{4} > \max\{0, \bar{a}^j\} = \bar{a}^j,
\]
\(^{54}\)This result was first noted by Arozamena and Cantillon (2004).
\(^{55}\)Note that \( i \) will always find it optimal to choose a valuation of 0 or \( \frac{1}{2} \mathbb{1}_{\omega=i} \), since the former is optimal for \( i \) when he does not obtain the item while the latter is optimal for \( i \) when he does obtain the item.
\(^{56}\)To see this, it suffices to note that it is socially optimal for \( i \) to choose a valuation that balances the marginal cost and marginal benefit of the item for \( i \); when \( i \) does not receive the item, it is socially optimal for \( i \) to choose a valuation of 0.
so \( v^i(\omega) = \frac{1}{2} \mathbb{1}_{\{\omega = i\}} \) is the \textit{ex ante} socially efficient investment for agent \( i \).

On the other hand, if \( \bar{a}^j > \frac{1}{4} \), then \( i \) can not obtain a positive profit from choosing \( v^i(\omega) = \frac{1}{2} \mathbb{1}_{\{\omega = i\}} \), as \( i \) will pay \( \bar{a}^j \) if he obtains the item. In this case, \( i \) chooses \( v^i(\omega) = 0 \); this is \textit{ex ante} efficient because choosing \( v^i(\omega) \notin \{0, \frac{1}{2} \mathbb{1}_{\{\omega = i\}}\} \) or \( v^j(\omega) \notin \{\bar{a}^j \mathbb{1}_{\{\omega = j\}}\} \) is never socially efficient, and

\[
\left(\frac{1}{2} \cdot 1 - \frac{1}{4}\right) + \left(\bar{a}^j \cdot 0 - 0\right) = \frac{1}{4} < \max\{0, \bar{a}^j\} = \bar{a}^j.
\]

Hence, regardless of the value of \( \bar{a}^j \), agent \( i \) will maximize his \textit{ex ante} utility by choosing the socially optimal level of investment.\footnote{If \( \bar{a}^j = \frac{1}{4} \), then the set of optimal valuation choices by \( i \) is \( \{0, \frac{1}{2} \mathbb{1}_{\{\omega = i\}}\} \); moreover, this is the set of socially optimal valuation choices.}

### B.2 Efficiency Is Necessary for Strategy-Proofness to Imply the Provision of Marginal Rewards

Consider a setting with two agents, a buyer \( b \) and a seller \( s \). There are two alternatives: one in which the buyer obtains the item, and one in which the seller keeps the item. Hence, the set of alternatives is \( \Omega \cong I = \{b, s\} \). The valuation function of the buyer is \( v^b = a^b \mathbb{1}_{\{\omega = b\}} \); \( a^b \) is private information but falls within the interval \([0, 1]\). Similarly, the valuation function of the seller is \( v^s = a^s \mathbb{1}_{\{\omega = s\}} \); \( a^s \) is private information but again falls within the interval \([0, 1]\).

A \textit{posted-price mechanism} is a mechanism in which, first, an exogenous price \( p \) is posted, and then both agents report their valuations. Trade occurs if and only if the seller’s valuation is below the posted price and the buyer’s valuation is above the posted price; if both of these conditions are satisfied, trade takes place at the price \( p \), and otherwise, no trade occurs and no transfer is made.

This mechanism is clearly strategy-proof, but it does \textit{not} induce efficient investment (or
provide marginal rewards). To see this, suppose that \( p = \frac{1}{2} \), the cost function for the buyer is

\[
c^b(v^b) = \begin{cases} 
0 & v^b(\omega) = 0 \\
\frac{1}{2} & v^b(\omega) = \frac{2}{3} 1_{\{\omega = b\}} \\
2017 & \text{otherwise},
\end{cases}
\]

and the cost function for the seller is

\[
c^s(v^s) = \begin{cases} 
0 & v^s(\omega) = 0 \\
2017 & \text{otherwise}.
\end{cases}
\]

Then the socially efficient alternative is for the buyer to choose \( v^b(\omega) = \frac{2}{3} 1_{\{\omega = b\}} \); however, such a choice is not individually optimal for \( b \), since it leads to an ex ante utility of \(-\frac{1}{3}\), while choosing \( v^b(\omega) = 0 \) leads to an ex ante utility of 0. Hence, while the posted price mechanism is strategy-proof, it does not induce efficient investment.\(^{58}\)

**B.3 Path-Connectedness Is Necessary for Strategy-Proofness to Imply the Provision of Marginal Rewards**

Consider a setting with two agents, \( I = \{i, j\} \). There is one good to be allocated, so that we have \( \Omega \cong I \). Suppose that the set of valuation functions for \( i \) is \( V^i = \{a 1_{\{\omega = i\}} : a \in [0, 2] \cup [5, 7]\} \), while the valuation of \( j \) is fixed at \( v^j(\omega) = 3 1_{\{\omega = j\}} \). Consider a mechanism \( \mathcal{M} \) such that

\[
\mathcal{M}((a 1_{\{\omega = i\}}, 3 1_{\{\omega = j\}})) = \begin{cases} 
(i, (4, 0)) & a \in [5, 7] \\
(j, (0, 0)) & a \in [0, 2],
\end{cases}
\]

\(^{58}\)It is also possible to construct an inefficient mechanism which is not strategy-proof but does provide marginal rewards.
where we list the transfer from $i$ before the transfer from $j$. This mechanism is clearly strategy-proof and efficient, but it does not provide marginal rewards; for example, the utility of $i$ only increases by 1 if his type changes from $2 \mathbb{1}_{\{\omega=i\}}$ to $5 \mathbb{1}_{\{\omega=i\}}$, while total social welfare increases by 2.

### B.4 Investment Miscoordination in a Nash Equilibrium of the Investment Game under a Second-Price Auction

Consider the setting of Appendix B.1, and again suppose that we use a second-price auction to allocate the good. Agent $i$ still has the cost function

$$c^i(a \mathbb{1}_{\{\omega=i\}}) = a^2,$$

while agent $j$ now has the cost function

$$c^j(a \mathbb{1}_{\{\omega=j\}}) = \frac{3}{2} a^2.$$

In this case, there exist two pure strategy Nash equilibria of the game induced by the second-price auction: $(\frac{1}{2} \mathbb{1}_{\{\omega=i\}}, 0)$ and $(0, \frac{1}{3} \mathbb{1}_{\{\omega=j\}})$. While the former of these equilibria maximizes ex ante social welfare, the latter equilibrium does not—while each agent is choosing a valuation so as to maximize social surplus given the choice of the other agent, social welfare can still be increased by simultaneously changing the investment choices of both agents.

### C Applications to Trading Networks

**Hatfield et al. (2013)** introduced a general framework for modeling trading networks with indivisible goods, which can be used to model economic settings such as industrial supply chains, reinsurance markets, and the dealer market for used cars.

In the setting of **Hatfield et al. (2013)**, agents can trade indivisible goods or services via
a network of bilateral contracts. Formally, there is a finite set of bilateral trades $\Omega$; each trade $\omega \in \Omega$ is associated with a distinct buyer $b(\omega)$ and seller $s(\omega)$. For each agent $i \in I$, we let $\Psi_{\rightarrow i} \equiv \{\psi \in \Psi : b(\psi) = i\}$ be the set of trades in $\Psi$ for which $i$ is a buyer, let $\Psi_{\leftarrow i} \equiv \{\psi \in \Psi : s(\psi) = i\}$ be the set of trades in $\Psi$ for which $i$ is a seller, and let $\Psi_i \equiv \Psi_{\rightarrow i} \cup \Psi_{\leftarrow i}$.

Each agent $i \in I$ has a valuation $v^i$ over sets of trades; hence, a set of trades $\Psi$ in the Hatfield et al. (2013) model corresponds to an alternative $\omega$ in our general framework. We impose the condition that $v^i(\Psi) = v^i(\Psi_i)$ for all $\Psi \subseteq \Omega$ and for all $i \in I$, i.e., agents only derive utility from trades with which they are associated. We also normalize the valuation functions so that $v^i(\emptyset) = 0$ for all $i \in I$. For each $i \in I$, we assume that there exists $v^i \in V^i$ such that, for all $\Psi$ such that $\Psi_i \neq \emptyset$ and for all $v^{i \setminus \{i\}} \in V^{i \setminus \{i\}}$, we have that $v^i(\Psi) < -\sum_{j \in I \setminus \{i\}} v^j(\Psi)$. This assumption simply allows the possibility that $i$ has a valuation under which he will never trade at any price acceptable to other agents.

A trade represents the nonpecuniary aspects of a transaction. We let $p \in \mathbb{R}^\Omega$ denote the vector of prices for all trades in the economy. An arrangement $[\Psi; p]$ specifying a set of executed trades and a price vector $p$ corresponds to the outcome

$$\left(\Psi, \left(\sum_{\psi \in \Psi_{\rightarrow i}} p_{\psi} - \sum_{\psi \in \Psi_{\leftarrow i}} p_{\psi}\right)_{i \in I}\right).$$

In (10), the transfer from agent $i$ is the sum of the prices of the executed trades for which $i$ is the buyer, minus the sum of the prices of the executed trades for which $i$ is the seller.

The demand of agent $i$ at prices $p$ is given by

$$D^i(p) \equiv \arg \max_{\Psi \subseteq \Omega_i} \left\{v^i(\Psi) - \sum_{\psi \in \Psi_{\rightarrow i}} p_{\psi} + \sum_{\psi \in \Psi_{\leftarrow i}} p_{\psi}\right\}.$$  

A competitive equilibrium is an arrangement $[\Psi; p]$ such that $\Psi_i \in D^i(p)$ for all $i \in I$.

---

59 Note that we use the notations $\omega$ and $\Omega$ to mean an alternative and the set of all possible alternatives, respectively, while we use the notations $\omega$ and $\Omega$ to mean a trade and the set of all possible trades, respectively.
Hatfield et al. (2013, 2016) show that the following substitutability condition is sufficient for the existence of competitive equilibrium.

**Definition C.1.** The valuation function $v^i$ is fully substitutable if, for all price vectors $p, \bar{p} \in \mathbb{R}^\Omega$ such that $\{\Psi\} = D^i(p)$ for some $\Psi \subseteq \Omega$, and $\{\bar{\Psi}\} = D^i(\bar{p})$ for some $\bar{\Psi} \subseteq \Omega$, if $p \leq \bar{p}$, then $\{\psi \in \Psi_{\rightarrow i} : p_\psi = \bar{p}_\psi\} \subseteq \bar{\Psi}$ and $\{\psi \in \bar{\Psi}_{\rightarrow i} : p_\psi = \bar{p}_\psi\} \subseteq \Psi$.\(^{60}\)

Intuitively, full substitutability can be understood by considering each trade as transferring a specific “item” from the seller to the buyer; in this case, full substitutability for agent $i$ states that no two items act as complements for agent $i$. That is, whenever prices increase from $p$ to $\bar{p}$, agent $i$ still demands all of the items that he demanded before whose prices did not change, both in the sense of still demanding as a buyer any previously-demanded trade whose price did not increase (i.e., $\{\psi \in \Psi_{\rightarrow i} : p_\psi = \bar{p}_\psi\} \subseteq \bar{\Psi}$), and still not participating as a seller in any trade previously not demanded whose price did not increase (i.e., $\{\psi \in \bar{\Psi}_{\rightarrow i} : p_\psi = \bar{p}_\psi\} \subseteq \Psi$).

We say that the valuations of $i$ are fully substitutable if each $v^i \in V^i$ is fully substitutable. Combining Theorems 1–4 of Hatfield et al. (2013) shows the following result.

**Proposition C.1 (Hatfield et al. (2013)).** If the valuations of each agent are fully substitutable, then, for any valuation profile $v \in V$, there exists a vector of prices $\bar{p}$ such that:

1. For any efficient set of trades $\Psi$, $[\Psi; \bar{p}]$ is a competitive equilibrium.

2. If $[\Psi; p]$ is a competitive equilibrium, then $\Psi$ is efficient and $p \leq \bar{p}$.

When the valuations of each agent are fully substitutable, Proposition C.1 implies the existence of a seller-optimal stable mechanism $\mathcal{M}^* = (\mu, r)$, i.e., a mechanism that selects an outcome corresponding to an efficient set of trades and the highest competitive equilibrium

\(^{60}\)The condition in Definition C.1 is a demand-theoretic phrasing of full substitutability. It is different from—but equivalent to—the Hatfield et al. (2013) full substitutability condition (see Hatfield et al. (2016)).
price vector $\tilde{p}$.\footnote{The term “seller-optimal stable mechanism” comes from the fact (Theorems 5 and 6 of Hatfield et al. (2013)) that there is a close correspondence (when preferences are fully substitutable) between competitive equilibria and stable sets of contracts.} In particular, under $\mathcal{M}^* = (\mu, r)$, for all $v \in V$, we have

$$\mu(v) \in \arg \max_{\Psi \in \Omega} \{ \mathbf{V}(\Psi; v) \}$$

$$r^i(v) = \sum_{\omega \in [\mu(v)] \rightarrow i} \tilde{p}_\omega - \sum_{\omega \in [\mu(v)] \leftarrow i} \tilde{p}_\omega$$

(11)

for each $i \in I$. Note that $\mathcal{M}^*$ is budget-balanced, that is, $\sum_{i \in I} r^i(v) = 0$ for all $v \in V$.

A unit-supply seller is an agent $s$ such that, for all $v^s \in V^s$, $v^s(\Psi) < -\sum_{j \in I \setminus \{s\}} v^j(\Psi)$ for all $v^{I \setminus \{s\}} \in V^{I \setminus \{s\}}$ if $|\Psi_s| > 1$ or $|\Psi_{\rightarrow s}| > 0$, i.e., $s$ is unwilling to engage in more than one trade or engage in any trade as a buyer at any price acceptable to other agents. We now show that the seller-optimal stable mechanism provides marginal rewards to each unit-supply seller.

**Proposition C.2.** If the valuations of each agent are fully substitutable, then the seller-optimal stable mechanism $\mathcal{M}^*$ provides marginal rewards to each unit-supply seller $s \in I$.

We now show that the seller-optimal stable mechanism is both strategy-proof for and induces efficient investment by unit-supply sellers.

**Corollary C.1.** If the preferences of each agent are fully substitutable, then the seller-optimal stable mechanism $\mathcal{M}^*$ is strategy-proof for each unit-supply seller $s \in I$.

**Corollary C.2.** If the preferences of each agent are fully substitutable, then the seller-optimal stable mechanism $\mathcal{M}^*$ induces efficient investment by each unit-supply seller $s \in I$.

Proposition C.2 and Corollaries C.1 and C.2 generalize Proposition 1 and Corollaries 3 and 4, respectively. Corollary C.1 follows from combining Proposition C.2 with Theorem 1 (and noting that the seller-optimal stable mechanism is efficient). Corollary C.2 follows from combining Proposition C.2 with Lemma 1.
We may extend Corollaries C.1 and C.2 to settings with uncertainty. In this case, the seller-optimal stable mechanism takes as input a valuation distribution for each agent, and is computed by determining the outcome of the seller-optimal stable mechanism in an economy without uncertainty where, for each alternative \( \omega \in \Omega \) and each agent \( i \in I \), the value of \( \omega \) to \( i \) is given by its expected value, \( \int v^i(\omega) \, dv^i(v^i) \). Combining Proposition C.2 with Lemma 3 and Theorem 4 then yields that the seller-optimal stable mechanism is strategy-proof and induces efficient investment in trading network settings with uncertainty. This latter result immediately implies the efficient network formation result (Proposition 1) of Kranton and Minehart (2001), who consider a model with unit-demand buyers who only learn the value of a seller’s good if they invest \textit{ex ante} in a relationship with that seller, and where the outcome is found by a mechanism that is equivalent to the buyer-optimal stable mechanism.\footnote{The symmetric nature of the trading networks model implies that the buyer-optimal stable mechanism is strategy-proof and induces efficient investment by unit-demand buyers.}

By contrast, when buyers have multi-unit demand, many mechanism designs used in practice, such as pay-as-bid auctions, are not strategy-proof for buyers (Ausubel et al., 2016). Our results imply that such mechanisms also fail to induce buyers to invest efficiently \textit{ex ante}.

\section*{D Proofs Omitted from the Text}

\textbf{Proof of Lemma 1}

First, we show that if \( \mathcal{M} \) provides marginal rewards to \( i \), then \( \mathcal{M} \) induces efficient investment by \( i \): Suppose that \( \mathcal{M} \) provides marginal rewards. Fix \( v^{I \setminus \{i\}} \in V^{I \setminus \{i\}} \), fix a cost function \( c^i \), and consider an arbitrary valuation \( v^i \in V^i \). The set of optimal investment choices for agent
\[ i \text{ is given by} \]
\[
\arg\max_{\bar{v}^i \in V^i} \left\{ u^i(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); \bar{v}^i) - c^i(\bar{v}^i) \right\} 
= \arg\max_{\bar{v}^i \in V^i} \left\{ u^i(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); \bar{v}^i) - u^i(\mathcal{M}(v^i, v^{I\setminus\{i\}}); v^i) - c^i(\bar{v}^i) \right\}
\]
\[
= \arg\max_{\bar{v}^i \in V^i} \left\{ V(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); (\bar{v}^i, v^{I\setminus\{i\}})) - V(\mathcal{M}(v^i, v^{I\setminus\{i\}}); (v^i, v^{I\setminus\{i\}})) - c^i(\bar{v}^i) \right\}
\]
\[
= \arg\max_{\bar{v}^i \in V^i} \left\{ V(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); (\bar{v}^i, v^{I\setminus\{i\}})) - c^i(\bar{v}^i) \right\}.
\]

In (12), the first equality follows because \( u^i(\mathcal{M}(v^i, v^{I\setminus\{i\}}); v^i) \) is a constant, the second equality follows from the fact that \( \mathcal{M} \) provides marginal rewards, and the third equality follows from the fact that \( V(\mathcal{M}(v^i, v^{I\setminus\{i\}}); (v^i, v^{I\setminus\{i\}})) \) is a constant. Hence, we see from (12) that \( \mathcal{M} \) induces efficient investment by \( i \).

Second, we show that if \( \mathcal{M} \) induces efficient investment by \( i \), then \( \mathcal{M} \) provides marginal rewards to \( i \): Suppose that \( \mathcal{M} \) does not provide marginal rewards to \( i \). Then it follows from Definition 2 that there exist \( v \in V \) and \( \bar{v}^i \in V^i \) such that\(^{63}\)

\[
u^i(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); \bar{v}^i) - u^i(\mathcal{M}(v); v^i) > V(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); (\bar{v}^i, v^{I\setminus\{i\}})) - V(\mathcal{M}(v); v). \quad (13)\]

Now, we consider the cost function

\[
c^i(\bar{v}^i) = \begin{cases} 
  u^i(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); \bar{v}^i) - \delta & \text{if } \bar{v}^i = \bar{v}^i \\
  u^i(\mathcal{M}(v); v^i) & \text{if } \bar{v}^i = v^i \\
  \bar{A} & \text{otherwise},
\end{cases}
\]

\(^{63}\)Note that it is without loss of generality to assume the direction of the inequality given in (13) as we can swap \( v^i \) and \( \bar{v}^i \) if necessary to reverse the inequality.
where $\delta > 0$ and $\bar{A} = 1 + \sup_{\tilde{v}^i \in V^i} \{u^i(M(\tilde{v}^i, v^{I \setminus \{i\}}); \tilde{v}^i)\}$.\footnote{Note that $\bar{A}$ is finite, as $\bar{A} = 1 + \sup_{\tilde{v}^i \in V^i} \{u^i(M(\tilde{v}^i, v^{I \setminus \{i\}}); \tilde{v}^i)\} \leq 1 + \sup_{\omega \in \Omega} \{u^i(\omega; \tilde{v}^i)\} - \inf_{\tilde{v}^i \in V^i} \{r^i(\tilde{v}^i, v^{I \setminus \{i\}})\}$, where $\Omega$ is finite, $V^i$ is compact, and $s$ is bounded by assumption.} We have

$$\{\tilde{v}^i\} = \arg \max_{\tilde{v}^i \in V^i} \left\{u^i(M(\tilde{v}^i, v^{I \setminus \{i\}}); \tilde{v}^i) - c^i(\tilde{v}^i)\right\}.$$ 

However, for $\delta$ sufficiently small, we have

$$\tilde{v}^i \notin \arg \max_{\tilde{v}^i \in V^i} \left\{V(M(\tilde{v}^i, v^{I \setminus \{i\}}); (\tilde{v}^i, v^{I \setminus \{i\}})) - c^i(\tilde{v}^i)\right\},$$

as we have

$$V(M(\tilde{v}^i, v^{I \setminus \{i\}}); (\tilde{v}^i, v^{I \setminus \{i\}})) - c^i(\tilde{v}^i) = V(M(\tilde{v}^i, v^{I \setminus \{i\}}); (\tilde{v}^i, v^{I \setminus \{i\}})) - u^i(M(\tilde{v}^i, v^{I \setminus \{i\}}); \tilde{v}^i) + \delta$$

$$< V(M(v); v) - u^i(M(v); \tilde{v}^i)$$

$$= V(M(v); v) - c^i(v^i), \quad (14)$$

where the first line of (14) follows from the definition of the cost function $c^i$, the second line follows from (13) (given $\delta$ sufficiently small), and the final line follows from the definition of the cost function $c^i$. We see from (14) that $M$ does not induce efficient investment by $i$.

**Proof of Theorem 1**

The equivalence of the first two statements follows from Lemma 1. Thus, we show the equivalence of the second and third statements.

We first show that any efficient mechanism that provides marginal rewards for $i \in I$ is also strategy-proof for $i$.\footnote{In fact, this result follows as a special case of Theorem 3. We provide a direct proof to clarify the key intuition behind the argument.} Consider the valuation $v \in V$ and some $\tilde{v}^i \in V^i$; for ease of notation, we let $\tilde{v} = (\tilde{v}^i, v^{I \setminus \{i\}})$. We calculate the change in the utility of $i$ from reporting his actual
valuation $v^i$ instead of $\bar{v}^i$ as follows:

$$
u^i(M(v); v^i) - u^i(M(\bar{v}); v^i) = u^i(M(v); v^i) - u^i(M(\bar{v}); \bar{v}^i) + u^i(M(\bar{v}); v^i) - u^i(M(\bar{v}); v^i)
\quad = V(M(v); v) - V(M(\bar{v}); \bar{v}) + v^i(M(\bar{v})) - v^i(M(\bar{v}))
\quad = V(M(v); v) - V(M(\bar{v}); \bar{v}) + \sum_{j \in I} v^j(M(\bar{v})) - \sum_{j \in I} v^j(M(\bar{v}))
\quad = V(M(v); v) - V(M(\bar{v}); \bar{v}) + V(M(\bar{v}); \bar{v}) - V(M(\bar{v}); v)
\quad = V(M(v); v) - V(M(\bar{v}); v)
\quad \geq 0.
$$

Here, the second equality follows from the fact that $M$ provides marginal rewards to $i$; the third equality follows from the fact that transfers are unchanged if reports are unchanged; the fourth equality follows from the fact that $v^j = \bar{v}^j$ for all $j \in I \setminus \{i\}$; the fifth equality follows from the definition of welfare (1); and the final inequality follows from the fact that $M$ is efficient.

The fact that strategy-proofness implies for $i$ provision of marginal rewards to $i$ follows as a special case of Theorem 3.

**Proof of Lemma 2**

First, we show that any mechanism $M$ that provides marginal rewards within $\epsilon$ induces efficient investment within $\epsilon$. Fix $v^{I \setminus \{i\}} \in V^{I \setminus \{i\}}$ and a cost function $c^i$, and consider\textsuperscript{66}

$$\hat{v}^i \in \arg \max_{\tilde{v}^i \in V^i} \left\{ u^i(M(\tilde{v}^i, v^{I \setminus \{i\}}); \tilde{v}^i) - c^i(\tilde{v}^i) \right\}.$$  

\textsuperscript{66}Note that if the set of maximizers is empty, then the condition in Definition 3 is vacuously satisfied.
Then, we have

\[ u^i(M(\hat{v}^i, v^{I \setminus \{i\}}); \hat{v}^i) - c^i(\hat{v}^i) \geq u^i(M(v^i, v^{I \setminus \{i\}}); v^i) - c^i(v^i) \] (15)

for all \( v^i \in V^i \). Since \( M \) provides marginal rewards within \( \epsilon \),

\[ \epsilon \geq \left( V(M(v); v) - V(M(\hat{v}^i, v^{I \setminus \{i\}}); (\hat{v}^i, v^{I \setminus \{i\}})) \right) - \left( u^i(M(v); v^i) - u^i(M(\hat{v}^i, v^{I \setminus \{i\}}); \hat{v}^i) \right). \]

Hence, we have

\[ V(M(\hat{v}^i, v^{I \setminus \{i\}}); (\hat{v}^i, v^{I \setminus \{i\}})) - u^i(M(\hat{v}^i, v^{I \setminus \{i\}}); \hat{v}^i) + \epsilon \geq V(M(v); v) - u^i(M(v); v^i). \]

Combining this with (15), we obtain

\[ V(M(\hat{v}^i, v^{I \setminus \{i\}}); (\hat{v}^i, v^{I \setminus \{i\}})) - c^i(\hat{v}^i) + \epsilon \geq V(M(v); v) - c^i(v^i). \]

Hence, \( M \) induces efficient investment within \( \epsilon \) by \( i \).

Second, we show that if \( M \) induces efficient investment within \( \epsilon \) by \( i \), then \( M \) provides marginal rewards within \( \epsilon \) to \( i \): Suppose that \( M \) does not provide marginal rewards within \( \epsilon \) to \( i \). Then, there exist \( v \in V \) and \( \tilde{v}^i \in V^i \) such that

\[ u^i(M(\tilde{v}^i, v^{I \setminus \{i\}}); \tilde{v}^i) - u^i(M(v); v^i) > V(M(\tilde{v}^i, v^{I \setminus \{i\}}); (\tilde{v}^i, v^{I \setminus \{i\}})) - V(M(v); v) + \epsilon. \] (16)

We now consider the cost function

\[ c^i(\tilde{v}^i) = \begin{cases} 
  u^i(M(\tilde{v}^i, v^{I \setminus \{i\}}); \tilde{v}^i) - \delta & \text{if } \tilde{v}^i = \bar{v}^i \\
  u^i(M(v); v^i) & \text{if } \tilde{v}^i = v^i \\
  \bar{A} & \text{otherwise,} 
\end{cases} \]
where \( \delta > 0 \) and \( \bar{A} = 1 + \sup_{\bar{v}^i \in V^i} \{ u^i(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); \bar{v}^i) \} \). We have

\[
\{ \bar{v}^i \} = \arg \max_{\bar{v}^i \in V^i} \left\{ u^i(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); \bar{v}^i) - c^i(\bar{v}^i) \right\}.
\]

However, for \( \delta \) sufficiently small, we have

\[
V(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); (\bar{v}^i, v^{I\setminus\{i\}})) - c^i(\bar{v}^i) + \epsilon < \sup_{\bar{v}^i \in V^i} \left\{ V(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); (\bar{v}^i, v^{I\setminus\{i\}})) - c^i(\bar{v}^i) \right\}
\]
as we have

\[
V(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); (\bar{v}^i, v^{I\setminus\{i\}})) - c^i(\bar{v}^i) + \epsilon = V(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); (\bar{v}^i, v^{I\setminus\{i\}})) - u^i(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); \bar{v}^i) + \delta + \epsilon
\]
\[
< V(\mathcal{M}(v); v) - u^i(\mathcal{M}(v); v^i)
\]
\[
= V(\mathcal{M}(v); v) - c^i(v^i).
\]

The first equality of (17) follows from the definition of the cost function \( c^i \); the inequality follows from (16) (given \( \delta \) sufficiently small); and the final equality follows from the definition of the cost function \( c^i \). Thus, we see that \( \mathcal{M} \) does not induce efficient investment within \( \epsilon \) by \( i \).

**Proof of Theorem 2**

Consider the valuation profile \( v \in V \) and some false report \( \bar{v}^i \in V^i \) for \( i \). For ease of notation, we let \( \bar{v} = (\bar{v}^i, v^{I\setminus\{i\}}) \). We calculate the change in the utility of \( i \) from reporting his actual

Theorem 2

\[
\text{Proof of Theorem 2}
\]

Consider the valuation profile \( v \in V \) and some false report \( \bar{v}^i \in V^i \) for \( i \). For ease of notation, we let \( \bar{v} = (\bar{v}^i, v^{I\setminus\{i\}}) \). We calculate the change in the utility of \( i \) from reporting his actual

\[\text{Note that } \bar{A} \text{ is finite, as}\]

\[\bar{A} = 1 + \sup_{\bar{v}^i \in V^i} \{ u^i(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); \bar{v}^i) \} \leq 1 + \sup_{\omega \in \Omega} \{ u^i(\omega; \bar{v}^i) \} - \inf_{\bar{v}^i \in V^i} \{ r^i(\bar{v}^i, v^{I\setminus\{i\}}) \},\]

\[\Omega \text{ is finite, } V^i \text{ is compact, and } r \text{ is bounded by assumption.}\]
valuation $v^i$ instead of $\bar{v}^i$ as follows:

\[
\begin{align*}
& u^i(M(v); v^i) - u^i(M(\bar{v}); \bar{v}^i) = u^i(M(v); v^i) - u^i(M(\bar{v}); \bar{v}^i) + u^i(M(\bar{v}); \bar{v}^i) - u^i(M(\bar{v}); v^i) \\
& \quad \geq V(M(v); v) - V(M(\bar{v}); \bar{v}) - \epsilon + u^i(M(\bar{v}); \bar{v}^i) - u^i(M(\bar{v}); v^i) \\
& \quad = V(M(v); v) - V(M(\bar{v}); \bar{v}) - \epsilon + \bar{v}^i(M(\bar{v})) - \bar{v}^i(M(\bar{v})) \\
& \quad = V(M(v); v) - V(M(\bar{v}); \bar{v}) - \epsilon + V(M(\bar{v}); \bar{v}) - V(M(\bar{v}); v) \\
& \quad = V(M(v); v) - V(M(\bar{v}); \bar{v}) - \epsilon \\
& \quad \geq -\epsilon - \eta.
\end{align*}
\]

Here, the second line follows from the fact that $\mathcal{M}$ provides marginal rewards within $\epsilon$ to $i$; the third line follows from the fact that identical reports induce identical transfers; the fourth line follows from the fact that $v^j = \bar{v}^j$ for all $j \in I \setminus \{i\}$; the fifth line follows from the definition of welfare (1); and the final inequality follows from the fact that $\mathcal{M}$ is efficient within $\eta$. Hence, we see that

\[
u^i(M(v); v^i) + \epsilon + \eta \geq u^i(M(\bar{v}); v^i),
\]

so that $\mathcal{M}$ is strategy-proof within $\epsilon + \eta$ for $i$.

**Proof of Theorem 3**

Fix $v^{i \setminus \{i\}} \in V^{I \setminus \{i\}}$, and consider two valuations $v^i$ and $\bar{v}^i$ for agent $i$; denote $v = (v^i, v^{I \setminus \{i\}})$ and $\bar{v} = (\bar{v}^i, v^{I \setminus \{i\}})$. For each $\omega \in \Omega^i$, let $V^i_\omega \equiv \{\bar{v}^i \in V^i : \mu(\bar{v}^i, v^{I \setminus \{i\}}) \in \omega\}$. Since the type space $V^i$ is path-connected, there is a path $\gamma(\cdot)$ between $v^i$ and $\bar{v}^i$, i.e., a continuous function $\gamma : [0, 1] \to V^i$ such that $\gamma(0) = v^i$ and $\gamma(1) = \bar{v}^i$. We fix $\delta > 0$.

Let $\omega_1$ be the equivalence class such that $\mu(v) \in \omega_1$, and let $\tilde{v}^i_{\omega_1} \equiv v^i$ and $\tilde{\bar{v}}_{\omega_1} \equiv$
\[ v^i = \hat{v}^i_{\omega_1} = \gamma(0) \quad \cdots \quad \hat{v}^i_{\omega_n} \gamma(y_{\omega_n}) \hat{v}^i_{\omega_{n+1}} \quad \hat{v}^i_{\omega_{n+1}} \cdots \hat{v}^i = \hat{v}^i_{\omega_N} = \gamma(1) \]

\[ \mu(\cdot) \in \omega_n \quad \mu(\cdot) \in \omega_{n+1} \]

Figure 1: Depiction of the path \( \gamma(\cdot) \). Distances are marked above the path, while the alternative chosen at the identified preferences is marked below. (Note that there may be valuations between \( \hat{v}^i_{\omega_n} \) and \( \gamma(y_{\omega_n}) \) on the path such that an alternative outside of \( \omega_n \) is chosen; the key point is that \( \mu(\hat{v}^i_{\omega_n}), \mu(\hat{v}^i_{\omega_n}) \in \omega_n \).)

\((\hat{v}^i_{\omega_1}, v^{I \setminus \{i\}})\). We now inductively define sequences of equivalence classes, path steps, and preference profiles, indexed by \( n = 1, \ldots, N \).

- Let \( y_{\omega_n} \equiv \sup\{ y \in [0, 1] : \gamma(y) \in V^i_{\omega_n} \} \).

- If \( y_{\omega_n} < 1 \):
  - Let \( \hat{y}_{\omega_n} \in \{ y \in [0, 1] : \| \gamma(y_{\omega_n}) - \gamma(y) \| \leq \delta \) and \( \mu(\gamma(y), v^{I \setminus \{i\}}) \in \omega_n \}; \) (18)

    note that \( \hat{y}_{\omega_n} \) must exist as, by the definition of \( y_{\omega_n} \), there exists \( y \) arbitrarily close to \( y_{\omega_n} \) such that \( \mu(\gamma(y), v^{I \setminus \{i\}}) \in \omega_n \).\(^{68}\) Let \( \hat{v}^i_{\omega_n} \equiv \gamma(\hat{y}_{\omega_n}) \) and \( \hat{v}_{\omega_n} \equiv (\hat{v}^i_{\omega_n}, v^{I \setminus \{i\}}) \).

  - Consider some

    \[ \tilde{y}_{\omega_{n+1}} \in \{ y \in [0, 1] : \| \gamma(y) - \gamma(y_{\omega_n}) \| \leq \delta \) and \( y > y_{\omega_n} \}; \]

    note that, by definition, \( \mu(\gamma(\tilde{y}_{\omega_{n+1}}), v^{I \setminus \{i\}}) \notin \bigcup_{m=1}^n \omega_m \), as \( \tilde{y}_{\omega_{n+1}} > y_{\omega_n} \) for all \( m \in \{1, \ldots, n\} \). Let \( \omega_{n+1} \) be the equivalence class such that \( \mu(\gamma(\tilde{y}_{\omega_{n+1}}), v^{I \setminus \{i\}}) \in \omega_{n+1} \). Let \( \tilde{v}^i_{\omega_{n+1}} \equiv \gamma(\tilde{y}_{\omega_{n+1}}) \) and \( \tilde{v}_{\omega_{n+1}} \equiv (\tilde{v}^i_{\omega_{n+1}}, v^{I \setminus \{i\}}) \).

- Otherwise (i.e., \( y_{\omega_n} = 1 \)):

\(^{68}\)The norm of a function \( x : \Omega \to \mathbb{R} \) is given by \( \sqrt{\sum_{\omega \in \Omega} (x(\omega))^2} \).
- If $\mu(\gamma(\omega_n), v^{i-\{i\}}) \in \omega_n$, then let $\hat{v}_n^i = \check{v}_n^i$ and $\check{v}_n = (\check{v}_n^i, v^{i-\{i\}})$, and let $N = n$.
- If $\mu(\gamma(\omega_n), v^{i-\{i\}}) \notin \omega_n$, then let $\hat{v}_n^i \equiv \gamma(\hat{\omega}_n)$, where $\hat{\omega}_n$ is as in (18), and let $\check{v}_n \equiv (\check{v}_n^i, v^{i-\{i\}})$. Finally, let $\hat{v}_{\omega_{n+1}} = \hat{v}_{\omega_{n+1}}^i = \check{v}_n$ and $\check{v}_{\omega_{n+1}} = (\hat{v}_{\omega_{n+1}}^i, v^{i-\{i\}})$, and let $N = n + 1$.

This construction is illustrated in Figure 1. Note that $N \leq |\Omega^i|$, as the equivalence class for each $n$ is distinct and $\mu(\gamma(\omega_{n+1}), v^{i-\{i\}}) \notin \cup_{m=1}^{n} \omega_m$ by construction for all $n \in \{1, \ldots, N\}$.

As $\mathcal{M}$ is strategy-proof within $\epsilon$ for $i$, we have

$$u^i(\mathcal{M}(\hat{v}_n); \check{v}_n^i) - u^i(\mathcal{M}(\check{v}_n); \check{v}_n^i) \leq \epsilon.$$ 

As $u^i(\mathcal{M}(\hat{v}_1); \check{v}_1^i) = u^i(\mathcal{M}(\hat{v}_1); \check{v}_1^i) - \left(\check{v}_1^i(\mu(\hat{v}_1)) - \check{v}_1^i(\mu(\check{v}_1))\right)$, we have that

$$u^i(\mathcal{M}(\hat{v}_1); \check{v}_1^i) - \left(\check{v}_1^i(\mu(\hat{v}_1)) - \check{v}_1^i(\mu(\check{v}_1))\right) - u^i(\mathcal{M}(\check{v}_1); \check{v}_1^i) \leq \epsilon.$$ 

Again as $\mathcal{M}$ is strategy-proof within $\epsilon$ for $i$, we obtain

$$u^i(\mathcal{M}(\hat{v}_1); \check{v}_1^i) - \left(\check{v}_1^i(\mu(\hat{v}_1)) - \check{v}_1^i(\mu(\check{v}_1))\right) - u^i(\mathcal{M}(\check{v}_1); \check{v}_1^i) \leq 2\epsilon.$$ 

Since $\|\hat{v}_1^i - \check{v}_1^i\| \leq 2\delta$, we have

$$u^i(\mathcal{M}(\hat{v}_1); \hat{v}_1^i) - \left(\check{v}_1^i(\mu(\hat{v}_1)) - \check{v}_1^i(\mu(\check{v}_1))\right) - u^i(\mathcal{M}(\check{v}_1); \check{v}_1^i) \leq 2\epsilon + 2\delta.$$ 

Hence, we obtain

$$u^i(\mathcal{M}(\hat{v}_1); \hat{v}_1^i) - \sum_{n=1}^{2} \left(\check{v}_n^i(\mu(\hat{v}_n)) - \check{v}_n^i(\mu(\check{v}_n))\right) - u^i(\mathcal{M}(\check{v}_1); \check{v}_1^i) \leq 2\epsilon + 2\delta.$$ 

80
Repeatedly applying the preceding three steps, we find that

\[ u^i(M(\hat{\omega}_N); \hat{\omega}_N) - \sum_{n=1}^{N} \left( \hat{v}^i_{\omega_n}(\mu(\hat{\omega}_n)) - \hat{v}^i_{\omega_n}(\mu(\hat{\omega}_n)) \right) - u^i(M(\hat{\omega}_1); \hat{\omega}_1) \leq N \epsilon + 2(N-1) \delta. \]

(19)

As \( M \) is efficient within \( \eta \), we have

\[ V(M(\hat{\omega}_N); \hat{\omega}_N) - V(M(\hat{\omega}_N); \hat{\omega}_N) \leq \eta. \]

This implies that

\[ V(M(\hat{\omega}_N); \hat{\omega}_N) - \left( \hat{v}^i_{\omega_N}(\mu(\hat{\omega}_N)) - \hat{v}^i_{\omega_N}(\mu(\hat{\omega}_N)) \right) - V(M(\hat{\omega}_N); \hat{\omega}_N) \leq \eta. \]

Again, as \( M \) is efficient within \( \eta \), we obtain

\[ V(M(\hat{\omega}_{N-1}); \hat{\omega}_{N-1}) - \left( \hat{v}^i_{\omega_{N-1}}(\mu(\hat{\omega}_{N-1})) - \hat{v}^i_{\omega_{N-1}}(\mu(\hat{\omega}_{N-1})) \right) - V(M(\hat{\omega}_{N-1}); \hat{\omega}_{N-1}) \leq 2 \eta. \]

Since \( \|\hat{\omega}_{N-1} - \hat{\omega}_N\| \leq 2 \delta \), we have

\[ V(M(\hat{\omega}_{N-1}); \hat{\omega}_{N-1}) - \left( \hat{v}^i_{\omega_{N-1}}(\mu(\hat{\omega}_{N-1})) - \hat{v}^i_{\omega_{N-1}}(\mu(\hat{\omega}_{N-1})) \right) - V(M(\hat{\omega}_{N-1}); \hat{\omega}_{N-1}) \leq 2 \eta + 2 \delta. \]

Hence, we obtain

\[ V(M(\hat{\omega}_{N-1}); \hat{\omega}_{N-1}) - \sum_{n=N-1}^{N} \left( \hat{v}^i_{\omega_n}(\mu(\hat{\omega}_n)) - \hat{v}^i_{\omega_n}(\mu(\hat{\omega}_n)) \right) - V(M(\hat{\omega}_{N-1}); \hat{\omega}_{N-1}) \leq 2 \eta + 2 \delta. \]

Repeatedly applying the preceding three steps, we get

\[ V(M(\hat{\omega}_1); \hat{\omega}_1) - \sum_{n=1}^{N} \left( \hat{v}^i_{\omega_n}(\mu(\hat{\omega}_n)) - \hat{v}^i_{\omega_n}(\mu(\hat{\omega}_n)) \right) - V(M(\hat{\omega}_N); \hat{\omega}_N) \leq N \eta + 2(N-1) \delta. \]
Note that $v^i(\mu(\tilde{\omega}_n)) = v^i(\mu(\nu_n))$ for all $v^i \in V^i$ as $\mu(\nu_n), \mu(\tilde{\omega}_n) \in \omega_n$ by definition. Hence,

$$V(M(\tilde{\omega}_1); \tilde{\omega}_1) - \sum_{n=1}^{N} \left( \tilde{v}^i_n(\mu(\tilde{\omega}_n)) - \tilde{v}^i_n(\mu(\nu_n)) \right) - V(M(\tilde{\omega}_N); \tilde{\omega}_N) \leq N\eta + 2(N-1)\delta.$$  

(20)

Adding together (19) and (20), we obtain

$$u^i(M(\tilde{\omega}_N); \tilde{\omega}_N) - u^i(M(\tilde{\omega}_1); \tilde{\omega}_1) - (V(M(\tilde{\omega}_N); \tilde{\omega}_N) - V(M(\tilde{\omega}_1); \tilde{\omega}_1)) \leq N(\epsilon + \eta) + 4(N-1)\delta.$$  

Noting that $\tilde{\omega}_N = \tilde{\omega}$ and $\tilde{\omega}_1 = \omega$, and that $N \leq |\Omega^i|$, we obtain:

$$u^i(M(\tilde{\omega}); \tilde{\omega}) - u^i(M(\nu); \nu) - (V(M(\tilde{\omega}); \tilde{\omega}) - V(M(\nu); \nu)) \leq N(\epsilon + \eta) + 4(N-1)\delta$$

$$\leq |\Omega^i|(\epsilon + \eta) + 4(|\Omega^i| - 1)\delta.$$  

Finally, taking $\delta \to 0$, we see that:

$$(u^i(M(\tilde{\omega}); \tilde{\omega}) - u^i(M(\nu); \nu)) - (V(M(\tilde{\omega}); \tilde{\omega}) - V(M(\nu); \nu)) \leq |\Omega^i|(\epsilon + \eta).$$

The proof that

$$(u^i(M(\nu); \nu) - u^i(M(\tilde{\omega}); \tilde{\omega})) - (V(M(\nu); \nu) - V(M(\tilde{\omega}); \tilde{\omega})) \leq |\Omega^i|(\epsilon + \eta)$$

is analogous.
Proof of Lemma 3

First, we show that any mechanism $M$ that provides marginal rewards within $\epsilon$ will induce efficient investment within $\epsilon$. Fix $v^{I \setminus \{i\}} \in \mathcal{V}^{I \setminus \{i\}}$ and a cost function $c^i$, and consider $69$

$$\hat{v}^i \in \text{arg max}_{\tilde{v}^i \in \mathcal{V}^i} \left\{ \left( \int \int u^i(M(v^i, v^{I \setminus \{i\}}); v^i) \, dv^{I \setminus \{i\}}(v^{I \setminus \{i\}}) \, d\tilde{v}^i(v^i) \right) - c^i(\hat{v}^i) \right\}. $$

We have

$$\left( \int \int u^i(M(v^i, v^{I \setminus \{i\}}); v^i) \, dv^{I \setminus \{i\}}(v^{I \setminus \{i\}}) \, d\tilde{v}^i(v^i) \right) - c^i(\hat{v}^i) \geq \left( \int \int u^i(M(v^i, v^{I \setminus \{i\}}); v^i) \, dv^{I \setminus \{i\}}(v^{I \setminus \{i\}}) \, d\bar{v}^i(v^i) \right) - c^i(\bar{v}^i)$$

for all $\bar{v}^i \in \mathcal{V}^i$. So

$$\left( \int \int u^i(M(v^i, v^{I \setminus \{i\}}); v^i) \, dv^{I \setminus \{i\}}(v^{I \setminus \{i\}}) \, d\hat{v}^i(v^i) \right) - \left( \int \int u^i(M(v^i, v^{I \setminus \{i\}}); v^i) \, dv^{I \setminus \{i\}}(v^{I \setminus \{i\}}) \, d\bar{v}^i(v^i) \right) \geq c^i(\hat{v}^i) - c^i(\bar{v}^i). \quad (21)$$

By Theorem 8.5.4 of Bogachev (2007), there is a lottery over $[0, 1]$ that “induces” lotteries $\hat{v}^i$ and $\bar{v}^i$; that is, there are functions $\hat{f}$ and $\bar{f}$ from $[0, 1]$ to $\mathcal{V}^i$ such that

$$\forall W^i \subseteq \mathcal{V}^i, \lambda(\hat{f}^{-1}(W^i)) = \hat{v}^i(W^i) \quad \text{and} \quad \lambda(\bar{f}^{-1}(W^i)) = \bar{v}^i(W^i),$$

where, for each measurable subset $z$ of $[0, 1]$, $\lambda(z)$ is the Lebesgue measure of $z$. $70$

Now, fix any $t \in [0, 1]$, and let $\hat{v}^i = \hat{f}(t)$ and $\bar{v}^i = \bar{f}(t)$. Fix $v^{I \setminus \{i\}}$ arbitrarily. Because

---

$69$ Note that if the set of maximizers is empty, then the condition in Definition 9 is vacuously satisfied.

$70$ We are grateful to Wei He, Yeneng Sun, and Satoru Takahashi for calling this result to our attention.
mechanism $\mathcal{M}$ provides marginal rewards within $\epsilon$ to $i \in I$, it follows that

$$\left| \langle \hat{V}(\mathcal{M}(\hat{x}^i, v^{I\setminus\{i\}}); (\hat{x}^i, v^{I\setminus\{i\}})) - \hat{V}(\mathcal{M}(\hat{x}^i, v^{I\setminus\{i\}}); (\hat{x}^i, v^{I\setminus\{i\}})) \rangle - u^i(\mathcal{M}(\hat{x}^i, v^{I\setminus\{i\}}); \hat{x}^i) - u^i(\mathcal{M}(\hat{x}^i, v^{I\setminus\{i\}}); \bar{x}^i) \right| \leq \epsilon.$$

Therefore,

$$\int u^i(\mathcal{M}(\hat{x}^i, v^{I\setminus\{i\}}); \hat{x}^i) d\bar{x}^i(\hat{x}^i) - \int u^i(\mathcal{M}(\hat{x}^i, v^{I\setminus\{i\}}); \hat{x}^i) d\bar{x}^i(\hat{x}^i) = \int_{[0,1]} u^i(\mathcal{M}(\hat{f}(t), v^{I\setminus\{i\}}); \hat{f}(t)) - u^i(\mathcal{M}(\hat{f}(t), v^{I\setminus\{i\}}); \bar{f}(t)) dt$$

$$\leq \int_{[0,1]} \left( V(\mathcal{M}(\hat{f}(t), v^{I\setminus\{i\}}); (\hat{f}(t), v^{I\setminus\{i\}})) - V(\mathcal{M}(\hat{f}(t), v^{I\setminus\{i\}}); (\bar{f}(t), v^{I\setminus\{i\}})) + \epsilon \right) dt$$

$$= \int_{[0,1]} V(\mathcal{M}(\hat{f}(t), v^{I\setminus\{i\}}); (\hat{f}(t), v^{I\setminus\{i\}})) dt - \int_{[0,1]} V(\mathcal{M}(\hat{f}(t), v^{I\setminus\{i\}}); (\bar{f}(t), v^{I\setminus\{i\}})) dt + \epsilon$$

$$= \int V(\mathcal{M}(\hat{x}^i, v^{I\setminus\{i\}}); \hat{x}^i, v^{I\setminus\{i\}}) d\bar{x}^i(\hat{x}^i) - \int V(\mathcal{M}(\hat{x}^i, v^{I\setminus\{i\}}); \hat{x}^i, v^{I\setminus\{i\}}) d\bar{x}^i(\hat{x}^i) + \epsilon.$$

Integrating this inequality over $v^{I\setminus\{i\}}$ with respect to $v^{I\setminus\{i\}}$, we obtain

$$\int u^i(\mathcal{M}(\hat{x}^i, v^{I\setminus\{i\}}); \hat{x}^i) d\bar{x}^i(\hat{x}^i) d\bar{v}^{-i}(\bar{v}^{-i}) - \int u^i(\mathcal{M}(\hat{x}^i, v^{I\setminus\{i\}}); \hat{x}^i) d\bar{x}^i(\hat{x}^i) d\bar{v}^{-i}(\bar{v}^{-i})$$

$$\leq \int V(\mathcal{M}(\hat{x}^i, v^{I\setminus\{i\}}); \hat{x}^i, v^{I\setminus\{i\}}) d\bar{x}^i(\hat{x}^i) d\bar{v}^{-i}(\bar{v}^{-i}) - \int V(\mathcal{M}(\hat{x}^i, v^{I\setminus\{i\}}); \hat{x}^i, v^{I\setminus\{i\}}) d\bar{x}^i(\hat{x}^i) d\bar{v}^{-i}(\bar{v}^{-i}) + \epsilon.$$

Substituting (21) into this inequality and rearranging, we obtain

$$\int V(\mathcal{M}(\hat{x}^i, v^{I\setminus\{i\}}); \hat{x}^i, v^{I\setminus\{i\}}) d\bar{x}^i(\hat{x}^i) d\bar{v}^{-i}(\bar{v}^{-i}) - \int V(\mathcal{M}(\hat{x}^i, v^{I\setminus\{i\}}); \hat{x}^i, v^{I\setminus\{i\}}) d\bar{x}^i(\hat{x}^i) d\bar{v}^{-i}(\bar{v}^{-i}) = c^i(\bar{v}^i) + \epsilon$$

Hence, $\mathcal{M}$ induces efficient investment within $\epsilon$ by $i$.

Second, we show that if $\mathcal{M}$ induces efficient investment within $\epsilon$ by $i$, then $\mathcal{M}$ provides
marginal rewards within $\epsilon$ to $i$: Suppose that the mechanism does not provide marginal rewards within $\epsilon$ to $i$. Then there exists $v \in V$ and $\bar{v}^i \in V^i$ such that

$$ u^i(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); \bar{v}^i) - u^i(\mathcal{M}(v); v^i) > V(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); (\bar{v}^i, v^{I\setminus\{i\}})) - V(\mathcal{M}(v); v) + \epsilon. $$

We consider the cost function

$$ c^i(v^i) = \begin{cases} 
    u^i(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); \bar{v}^i) - \delta & \text{if } v^i = \delta_{\bar{v}^i} \\
    u^i(\mathcal{M}(v); v^i) & \text{if } v^i = \delta_{v^i} \\
    \bar{A} & \text{otherwise},
\end{cases} $$

where $\delta > 0$ is sufficiently small and $\bar{A} = 1 + \sup_{v^i \in V^i} u^i(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); \bar{v}^i)$ is a large constant similar to that described in Footnote 64. Letting $v^{I\setminus\{i\}} = \delta_{v^{I\setminus\{i\}}}$, we have

$$ \{\delta_{\bar{v}^i}\} = \arg \max_{v^i \in V^i} \left\{ \iint u^i(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); \bar{v}^i) d\bar{v}^i d(\bar{v}^{I\setminus\{i\}}) - c^i(\bar{v}^i) \right\}. $$

However, we have

$$ \iint V(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); \bar{v}^i) d\delta_{\bar{v}^i}(\bar{v}^i) d\bar{v}^{I\setminus\{i\}}(\bar{v}^{I\setminus\{i\}}) - c^i(\bar{v}^i) + \epsilon < \sup_{v^i \in V^i} \left\{ \iint V(\mathcal{M}(\bar{v}^i, v^{I\setminus\{i\}}); \bar{v}^i) d\bar{v}^i d\bar{v}^{I\setminus\{i\}}(\bar{v}^{I\setminus\{i\}}) - c^i(\bar{v}^i) \right\}, $$

85
as we have

\[
\int \int V(\mathcal{M}(\tilde{v}^i, \tilde{v}^{I \setminus \{i\}}); \bar{v}^i) \, d\delta_{\bar{v}^i}(\tilde{v}^i) \, dv^{I \setminus \{i\}}(\tilde{v}^{I \setminus \{i\}}) - c^i(\delta_{\bar{v}^i}) + \epsilon
\]

\[
= V(\mathcal{M}(\tilde{v}^i, v^{I \setminus \{i\}}); \bar{v}^i) - c^i(\delta_{\bar{v}^i}) + \epsilon
\]

\[
= V(\mathcal{M}(\tilde{v}^i, v^{I \setminus \{i\}}); \bar{v}^i) - u^i(\mathcal{M}(\tilde{v}^i, v^{I \setminus \{i\}}); \tilde{v}^i) + \delta + \epsilon
\]

\[
< V(\mathcal{M}(v^i, v^{I \setminus \{i\}}); v^i) - u^i(\mathcal{M}(v^i, v^{I \setminus \{i\}}); v^i)
\]

\[
= V(\mathcal{M}(v^i, v^{I \setminus \{i\}}); v^i) - c^i(\delta_{\bar{v}^i})
\]

\[
= \int \int V(\mathcal{M}(\tilde{v}^i, \tilde{v}^{I \setminus \{i\}}); \bar{v}^i) \, d\delta_{\bar{v}^i}(\tilde{v}^i) \, dv^{I \setminus \{i\}}(\tilde{v}^{I \setminus \{i\}}) - c^i(\delta_{\bar{v}^i}),
\]

where the first equality follows from simple integration, the second equality follows from the definition of the cost function \(c^i\), the inequality follows from (22) (for \(\delta\) sufficiently small), the next equality follows from the definition of the cost function, and the last equality follows from simple integration. Hence, we see that the mechanism \(\mathcal{M}\) does not induce efficient investment by \(i\).

**Proof of Theorem 4**

Consider the valuation \(v \in \mathcal{V}\) and some other report \(\bar{v}^i \in \mathcal{V}^i\); for ease of notation, let \(\tilde{v} = (\tilde{v}^i, v^{I \setminus \{i\}})\). We calculate the change in the expected utility of \(i\) from reporting his actual
valuation \( v^i \) instead of \( \bar{v}^i \), i.e.,
\[
\begin{align*}
&u^i(M(v); v^i) - u^i(M(\bar{v}); \bar{v}^i) + u^i(M(\bar{v}); \bar{v}^i) - u^i(M(\bar{v}); v^i) \\
&\geq V(M(v); v) - V(M(\bar{v}); \bar{v}^i) - \epsilon + u^i(M(\bar{v}); \bar{v}^i) - u^i(M(\bar{v}); v^i) \\
&= V(M(\bar{v}); v) - V(M(\bar{v}); \bar{v}^i) - \epsilon + v^i(M(\bar{v})) - \sum_{j \in I} v^j(M(\bar{v})) \\
&= V(M(\bar{v}); v) - V(M(\bar{v}); \bar{v}^i) - \epsilon + V(M(\bar{v}); \bar{v}) - V(M(\bar{v}); v) \\
&= V(M(\bar{v}); v) - V(M(\bar{v}); v) - \epsilon \\
&\geq -\epsilon - \eta,
\end{align*}
\]

where the first inequality follows from the fact that \( M \) provides marginal rewards within \( \epsilon \) to \( i \), the first equality follows from the fact that identical reports induce identical transfers, the second equality follows from the fact that \( v^j = \bar{v}^j \) for all \( j \in I \setminus \{i\} \), the third equality follows from the definition of expected welfare (3), and the last inequality follows from the fact that \( M \) is efficient within \( \eta \). Hence,

\[
u^i(M(v); v^i) + \epsilon + \eta \geq u^i(M(\bar{v}); v^i)
\]

and so \( M \) is strategy-proof within \( \epsilon + \eta \) for \( i \).

**Proof of Theorem 5**

Fix \( v^{I \setminus \{i\}} \in \mathcal{V}^{I \setminus \{i\}} \), and consider \( v^i, \bar{v}^i \in \mathcal{V}^i \) for agent \( i \); denote \( \nu = (v^i, v^{I \setminus \{i\}}) \) and \( \nu = (\bar{v}^i, v^{I \setminus \{i\}}) \). For each \( \omega \in \Omega^i \), let \( \mathcal{V}^i_\omega \equiv \{ \nu^i \in \mathcal{V}^i : \mu(\nu^i, v^{I \setminus \{i\}}) \in \omega \} \). Since \( \nu^i = \Delta(V^i) \) is convex, the line segment \( \gamma(\cdot) \) between \( v^i \) and \( \bar{v}^i \) lies in \( \mathcal{V}^i \), i.e., the function \( \gamma : [0, 1] \to \mathcal{V}^i \) defined by
\[
\gamma(y) = (1 - y)v^i + y\bar{v}^i \quad \text{for all } y \in [0, 1]
\]
satisfies $\gamma(0) = \varphi^i$ and $\gamma(1) = \tilde{\varphi}^i$ and $\gamma(y) \in \mathcal{V}^i$ for every $y \in [0, 1]$. 

Fix $\delta > 0$.

$$
\begin{array}{cccc}
\varphi^i & = & \tilde{\varphi}^i_{\omega_1} = \gamma(0) & \cdots & \tilde{\varphi}^i_{\omega_n} & \gamma(y_{\omega_n}) & \tilde{\varphi}^i_{\omega_{n+1}} & \cdots & \tilde{\varphi}^i = \tilde{\varphi}^i_{\omega_N} = \gamma(1)
\end{array}
$$

Figure 2: Depiction of the line segment $\gamma(y)$. Distances in the parameter space are marked above the path, while the alternative chosen at the identified preferences is marked below. (Note that there may be valuations between $\tilde{\varphi}^i_{\omega_n}$ and $\gamma(y_{\omega_n})$ on the path such that an alternative outside of $\omega_n$ is chosen; the key point is that $\mu(\tilde{\varphi}^i_{\omega_n}), \mu(\tilde{\varphi}^i_{\omega_n}) \in \omega_n$.)

Let $\omega_1$ be the equivalence class such that $\mu(\varphi) \in \omega_1$, and let $\tilde{\varphi}^i_{\omega_1} \equiv \varphi^i$ and $\tilde{\varphi}_{\omega_1} \equiv (\tilde{\varphi}^i_{\omega_1}, \varphi^{I \setminus \{i\}})$. We now inductively define sequences of equivalence classes, path steps, and preference profiles, indexed by $n = 1, \ldots, N$.

- Let $y_{\omega_n} \equiv \sup\{y \in [0, 1] : \gamma(y) \in \mathcal{V}^i_{\omega_n}\}$.

- If $y_{\omega_n} < 1$:
  - Let
    $$
    \hat{y}_{\omega_n} \in \{y \in [0, 1] : y_{\omega_n} - \delta < y \leq y_{\omega_n} \text{ and } \mu(\gamma(y), \varphi^{I \setminus \{i\}}) \in \omega_n\};
    $$
    (23)
    note that $\hat{y}_{\omega_n}$ must exist as, by the definition of $y_{\omega_n}$, there exists $y$ arbitrarily close to $y_{\omega_n}$ such that $\mu(\gamma(y), \varphi^{I \setminus \{i\}}) \in \omega_n$. Let $\hat{\varphi}^i_{\omega_n} \equiv \gamma(\hat{y}_{\omega_n})$ and $\hat{\varphi}_{\omega_n} \equiv (\hat{\varphi}^i_{\omega_n}, \varphi^{I \setminus \{i\}})$.
  - Consider some
    $$
    \hat{y}_{\omega_{n+1}} \in \{y \in [0, 1] : y_{\omega_n} < y < y_{\omega_n} + \delta\};
    $$
    note that, by definition, $\mu(\gamma(\hat{y}_{\omega_{n+1}}), \varphi^{I \setminus \{i\}}) \notin \bigcup_{m=1}^n \omega_m$, as $\hat{y}_{\omega_{n+1}} > y_{\omega_m}$ for all $m \in \{1, \ldots, n\}$. Let $\omega_{n+1}$ be the equivalence class such that $\mu(\gamma(\hat{y}_{\omega_{n+1}}), \varphi^{I \setminus \{i\}}) \in \omega_{n+1}$. Let $\tilde{\varphi}^i_{\omega_{n+1}} \equiv \gamma(\hat{y}_{\omega_{n+1}})$ and $\tilde{\varphi}_{\omega_{n+1}} \equiv (\tilde{\varphi}^i_{\omega_{n+1}}, \varphi^{I \setminus \{i\}})$. 

88
• Otherwise (i.e., if \( y_{\omega_n} = 1 \)):
  
  - If \( \mu(\gamma(y_{\omega_n}), \nu^{i \setminus \{i\}}) \in \omega_n \), then let \( \hat{v}_i^{\omega_n} = \hat{v}^i \) and \( \hat{v}_{\omega_n} = (\hat{v}_i^{\omega_n}, \nu^{i \setminus \{i\}}) \), and let \( N = n \).
  
  - If \( \mu(\gamma(y_{\omega_n}), \nu^{i \setminus \{i\}}) \notin \omega_n \), then let \( \hat{v}_i^{\omega_n} \equiv \gamma(\hat{y}_{\omega_n}) \), where \( \hat{y}_{\omega_n} \) is as in (23), and let \( \hat{v}_{\omega_n} \equiv (\hat{v}_i^{\omega_n}, \nu^{i \setminus \{i\}}) \). Finally, let \( \hat{v}_{\omega_{n+1}}^{i} = \hat{v}_i^{\omega_{n+1}} = \hat{v}^i \) and \( \hat{v}_{\omega_{n+1}} = (\hat{v}_i^{\omega_{n+1}}, \nu^{i \setminus \{i\}}) \), and let \( N = n + 1 \).

This construction is illustrated in Figure 2. Note that \( N \leq |\Omega| \), as the equivalence class for each \( n \) is distinct, since \( \mu(\gamma(\hat{y}_{\omega_{n+1}}), \nu^{i \setminus \{i\}}) \notin \bigcup_{m=1}^{n} \omega_m \) by construction for all \( n \in \{1, \ldots, N\} \).

By strategy-proofness within \( \epsilon \), we have

\[
u^{i}(\mathcal{M}(\hat{v}_{\omega_1}); \hat{v}_i^{\omega_1}) - \nu^{i}(\mathcal{M}(\hat{v}_{\omega_1}; \hat{v}_i^{\omega_1}) \leq \epsilon.
\]

As we have

\[
u^{i}(\mathcal{M}(\hat{v}_{\omega_1}; \hat{v}_i^{\omega_1}) = \nu^{i}(\mathcal{M}(\hat{v}_{\omega_1}; \hat{v}_i^{\omega_1}) - \hat{v}_i^{\omega_1}(\mu(\hat{v}_{\omega_1})) \right) - \nu^{i}(\mathcal{M}(\hat{v}_{\omega_1}; \hat{v}_i^{\omega_1}) \leq \epsilon.
\]

we have that

\[
u^{i}(\mathcal{M}(\hat{v}_{\omega_1}; \hat{v}_i^{\omega_1}) - \left( \hat{v}_i^{\omega_1}(\mu(\hat{v}_{\omega_1})) - \hat{v}_i^{\omega_1}(\mu(\hat{v}_{\omega_1})) \right) - \nu^{i}(\mathcal{M}(\hat{v}_{\omega_1}; \hat{v}_i^{\omega_1}) \leq \epsilon.
\]

As \( \mathcal{M} \) is strategy-proof within \( \epsilon \) for \( i \), we obtain

\[
u^{i}(\mathcal{M}(\hat{v}_{\omega_2}; \hat{v}_i^{\omega_2}) - \left( \hat{v}_i^{\omega_1}(\mu(\hat{v}_{\omega_1})) - \hat{v}_i^{\omega_1}(\mu(\hat{v}_{\omega_1})) \right) - \nu^{i}(\mathcal{M}(\hat{v}_{\omega_1}; \hat{v}_i^{\omega_1}) \leq 2\epsilon.
\]

Since \( |\hat{y}_{\omega_2} - \hat{y}_{\omega_1}| \leq 2\delta \), we have

\[
u^{i}(\mathcal{M}(\hat{v}_{\omega_2}; \hat{v}_i^{\omega_2}) - \left( \hat{v}_i^{\omega_1}(\mu(\hat{v}_{\omega_1})) - \hat{v}_i^{\omega_1}(\mu(\hat{v}_{\omega_1})) \right) - \nu^{i}(\mathcal{M}(\hat{v}_{\omega_1}; \hat{v}_i^{\omega_1}) \leq 2\epsilon + 2\delta \overline{A},
\]

89
where $\bar{A} \equiv \sup_{\tilde{v}^i \in V^i, \omega \in \Omega} \{v^i(\omega; \tilde{v}^i)\} - \inf_{\tilde{v}^i \in V^i, \omega \in \Omega} \{v^i(\omega; \tilde{v}^i)\}$.\footnote{Note that $\bar{A}$ is finite by an argument analogous to the one described in Footnote 64.}

Hence, we obtain

$$ u^i(\mathcal{M}(\hat{\omega}_2); \hat{v}^i_{\omega_2}) - \sum_{n=1}^{2} \left( \hat{\omega}_n (\mu(\hat{\omega}_n)) - \hat{v}^i_{\omega_n} (\mu(\hat{\omega}_n)) \right) - u^i(\mathcal{M}(\hat{\omega}_1); \hat{v}^i_{\omega_1}) \leq 2\epsilon + 2\delta \bar{A}. $$

Repeatedly applying the last three steps, we get

$$ u^i(\mathcal{M}(\hat{\omega}_N); \hat{v}^i_{\omega_N}) - \sum_{n=1}^{N} \left( \hat{\omega}_n (\mu(\hat{\omega}_n)) - \hat{v}^i_{\omega_n} (\mu(\hat{\omega}_n)) \right) - u^i(\mathcal{M}(\hat{\omega}_1); \hat{v}^i_{\omega_1}) $$

$$ \leq N\epsilon + 2(N-1)\delta \bar{A}. \quad (24) $$

As $\mathcal{M}$ is efficient within $\eta$, we have

$$ V(\mathcal{M}(\hat{\omega}_N); \hat{v}^i_{\omega_N}) - V(\mathcal{M}(\hat{\omega}_N); \hat{v}^i_{\omega_N}) \leq \eta. $$

This implies that

$$ V(\mathcal{M}(\hat{\omega}_N); \hat{v}^i_{\omega_N}) - \left( \hat{\omega}_N (\mu(\hat{\omega}_N)) - \hat{v}^i_{\omega_N} (\mu(\hat{\omega}_N)) \right) - V(\mathcal{M}(\hat{\omega}_N); \hat{v}^i_{\omega_N}) \leq \eta. $$

Now, as $\mathcal{M}$ is efficient within $\eta$, we obtain

$$ V(\mathcal{M}(\hat{\omega}_{N-1}); \hat{v}^i_{\omega_{N-1}}) - \left( \hat{\omega}_N (\mu(\hat{\omega}_N)) - \hat{v}^i_{\omega_N} (\mu(\hat{\omega}_N)) \right) - V(\mathcal{M}(\hat{\omega}_N); \hat{v}^i_{\omega_N}) \leq 2\eta. $$

Since $|\bar{\omega}_{N-1} - \bar{\omega}_{N}| \leq 2\delta$, we have

$$ V(\mathcal{M}(\bar{\omega}_{N-1}); \hat{v}^i_{\omega_{N-1}}) - \left( \hat{\omega}_N (\mu(\hat{\omega}_N)) - \hat{v}^i_{\omega_N} (\mu(\hat{\omega}_N)) \right) - V(\mathcal{M}(\hat{\omega}_N); \hat{v}^i_{\omega_N}) \leq 2\eta + 2\delta \bar{A}. $$

71 Note that $\bar{A}$ is finite by an argument analogous to the one described in Footnote 64.
Hence, we obtain

\[ V(\mathcal{M}(\tilde{v}_{\omega_{N-1}}); \tilde{v}_{\omega_{N-1}}) - \sum_{n=N-1}^{N} \left( \tilde{v}_{\omega_n}^i (\mu(\tilde{v}_{\omega_n})) - \hat{v}_{\omega_n}^i (\mu(\tilde{v}_{\omega_n})) \right) - V(\mathcal{M}(\hat{v}_{\omega_N}); \hat{v}_{\omega_N}) \]

\[ \leq 2\eta + 2\delta \bar{A}. \]

Repeatedly applying the last three steps, we get

\[ V(\mathcal{M}(\tilde{v}_{\omega_1}); \tilde{v}_{\omega_1}) - \sum_{n=1}^{N} \left( \tilde{v}_{\omega_n}^i (\mu(\tilde{v}_{\omega_n})) - \hat{v}_{\omega_n}^i (\mu(\tilde{v}_{\omega_n})) \right) - V(\mathcal{M}(\hat{v}_{\omega_N}); \hat{v}_{\omega_N}) \]

\[ \leq N\eta + 2(N-1)\delta \bar{A}. \]

Note that \( v^i(\mu(\tilde{v}_{\omega_n})) = v^i(\mu(\hat{v}_{\omega_n})) \) for all \( v^i \in \mathcal{V} \) as \( \mu(\tilde{v}_{\omega_n}), \mu(\hat{v}_{\omega_n}) \in \omega_n \) by definition.

Hence,

\[ V(\mathcal{M}(\tilde{v}_{\omega_1}); \tilde{v}_{\omega_1}) - \sum_{n=1}^{N} \left( \tilde{v}_{\omega_n}^i (\mu(\tilde{v}_{\omega_n})) - \hat{v}_{\omega_n}^i (\mu(\tilde{v}_{\omega_n})) \right) - V(\mathcal{M}(\hat{v}_{\omega_N}); \hat{v}_{\omega_N}) \]

\[ \leq N\eta + 2(N-1)\delta \bar{A}. \quad (25) \]

Adding together (24) and (25) we obtain

\[ u^i(\mathcal{M}(\hat{v}_{\omega_N}); \hat{v}_{\omega_N}) - u^i(\mathcal{M}(\tilde{v}_{\omega_1}); \tilde{v}_{\omega_1}) - (V(\mathcal{M}(\hat{v}_{\omega_N}); \hat{v}_{\omega_N}) - V(\mathcal{M}(\tilde{v}_{\omega_1}); \tilde{v}_{\omega_1})) \leq N(\epsilon + \eta) + 4(N-1)\delta \bar{A}. \]

Noting that \( \hat{v}_{\omega_N} = \tilde{v}^i \) and \( \hat{v}_{\omega_1} = v^i \), and that \( N \leq |\Omega|^i \), we obtain:

\[ u^i(\mathcal{M}(\tilde{v}); \tilde{v}^i) - u^i(\mathcal{M}(v); v^i) - (V(\mathcal{M}(\tilde{v}); \tilde{v}) - V(\mathcal{M}(v); v)) \leq N(\epsilon + \eta) + 4(N-1)\delta \bar{A} \]

\[ \leq |\Omega|^i(\epsilon + \eta) + 4(|\Omega|^i - 1)\delta \bar{A}. \]
Finally, taking $\delta \to 0$, we see that:

$$(u^i(\mathcal{M}(\tilde{v}); \tilde{v}^i) - u^i(\mathcal{M}(v); v^i)) - (\mathcal{V}(\mathcal{M}(\tilde{v}); \tilde{v}) - \mathcal{V}(\mathcal{M}(v); v)) \leq |\Omega^i| (\epsilon + \eta).$$

The proof that

$$(u^i(\mathcal{M}(v); v) - u^i(\mathcal{M}(\tilde{v}); \tilde{v})) - (\mathcal{V}(\mathcal{M}(v); v) - \mathcal{V}(\mathcal{M}(\tilde{v}); \tilde{v})) \leq |\Omega^i| (\epsilon + \eta)$$

is analogous.

**Proof of Lemma 4**

We show that any mechanism $\mathcal{M}$ that provides marginal rewards within $\epsilon$ in expectation induces efficient investment within $\epsilon$. Fix $\mathcal{v}^{I\setminus\{i\}} \in \mathcal{V}^{I\setminus\{i\}}$ and a cost function $c^i$, and consider

$$\hat{v}^i \in \arg\max_{v^i \in \mathcal{V}^i} \{E(v^i, \mathcal{v}^{I\setminus\{i\}}) \left[ u^i(\mathcal{M}(v^i, \mathcal{v}^{I\setminus\{i\}}); v^i) - c^i(v^i) \right] \}.$$  

Then, we have

$$E(\hat{v}^i, \mathcal{v}^{I\setminus\{i\}}) \left[ u^i(\mathcal{M}(v^i, \mathcal{v}^{I\setminus\{i\}}); v^i) - c^i(\hat{v}^i) \right] \geq E(\mathcal{v}^i, \mathcal{v}^{I\setminus\{i\}}) \left[ u^i(\mathcal{M}(v^i, \mathcal{v}^{I\setminus\{i\}}); v^i) - c^i(\mathcal{v}^i) \right] - c^i(\mathcal{v}^i) \tag{26}$$

for each $\mathcal{v}^i \in \mathcal{V}^i$. Since $\mathcal{M}$ provides marginal rewards within $\epsilon$ in expectation,

$$\left( E_{\mathcal{v}^{I\setminus\{i\}}} \left[ V(\mathcal{M}(v^i, \mathcal{v}^{I\setminus\{i\}}); (v^i, v^{I\setminus\{i\}})) \right] - E_{\mathcal{v}^{I\setminus\{i\}}} \left[ V(\mathcal{M}(\tilde{v}^i, \mathcal{v}^{I\setminus\{i\}}); (\tilde{v}^i, v^{I\setminus\{i\}})) \right] \right) -$$

$$\left( E_{\mathcal{v}^{I\setminus\{i\}}} \left[ u^i(\mathcal{M}(v^i, \mathcal{v}^{I\setminus\{i\}}); (v^i, v^{I\setminus\{i\}})) \right] - E_{\mathcal{v}^{I\setminus\{i\}}} \left[ u^i(\mathcal{M}(\tilde{v}^i, \mathcal{v}^{I\setminus\{i\}}); (\tilde{v}^i, v^{I\setminus\{i\}})) \right] \right) \leq \epsilon$$

$^72$Note that if the set of maximizers is empty, then the condition in Definition 9 is vacuously satisfied.
for all $v^i, \bar{v}^i \in \mathcal{V}^i$. Taking an expectation by letting $v^i$ be drawn from the distribution $v^i$, we obtain

$$
\left( \mathbb{E}_{(v^i, v^{I \setminus \{i\}})} \left[ \mathcal{V}(\mathcal{M}(v^i, v^{I \setminus \{i\}}); (v^i, v^{I \setminus \{i\}})) \right] - \mathbb{E}_{v^{I \setminus \{i\}}} \left[ \mathcal{V}(\mathcal{M}(\bar{v}^i, v^{I \setminus \{i\}}); (\bar{v}^i, v^{I \setminus \{i\}})) \right] \right) - \\
\left( \mathbb{E}_{(v^i, v^{I \setminus \{i\}})} \left[ u^i(\mathcal{M}(v^i, v^{I \setminus \{i\}}); (v^i, v^{I \setminus \{i\}})) \right] - \mathbb{E}_{v^{I \setminus \{i\}}} \left[ u^i(\mathcal{M}(\bar{v}^i, v^{I \setminus \{i\}}); (\bar{v}^i, v^{I \setminus \{i\}})) \right] \right) \leq \epsilon.
$$

Now, taking an expectation by letting $\bar{v}^i$ be drawn from the distribution $\hat{v}^i$, we obtain

$$
\left( \mathbb{E}_{(v^i, v^{I \setminus \{i\}})} \left[ \mathcal{V}(\mathcal{M}(v^i, v^{I \setminus \{i\}}); (v^i, v^{I \setminus \{i\}})) \right] - \mathbb{E}_{(\hat{v}^i, \bar{v}^{I \setminus \{i\}}, v^{I \setminus \{i\}})} \left[ \mathcal{V}(\mathcal{M}(\hat{v}^i, v^{I \setminus \{i\}}); (\hat{v}^i, \bar{v}^{I \setminus \{i\}}, v^{I \setminus \{i\}})) \right] \right) - \\
\left( \mathbb{E}_{(v^i, v^{I \setminus \{i\}})} \left[ u^i(\mathcal{M}(v^i, v^{I \setminus \{i\}}); (v^i, v^{I \setminus \{i\}})) \right] - \mathbb{E}_{(\hat{v}^i, \bar{v}^{I \setminus \{i\}}, v^{I \setminus \{i\}})} \left[ u^i(\mathcal{M}(\hat{v}^i, v^{I \setminus \{i\}}); (\hat{v}^i, \bar{v}^{I \setminus \{i\}}, v^{I \setminus \{i\}})) \right] \right) \leq \epsilon.
$$

Combining this with (26) (and noting that $\bar{v}^i$ is just a dummy variable in taking an expectation), we obtain

$$
\mathbb{E}_{(v^i, v^{I \setminus \{i\}})} \left[ \mathcal{V}(\mathcal{M}(v^i, v^{I \setminus \{i\}}); v^i) \right] - c^i(\hat{v}^i) + \epsilon \geq \mathbb{E}_{(v^i, v^{I \setminus \{i\}})} \left[ \mathcal{V}(\mathcal{M}(v^i, v^{I \setminus \{i\}}); v^i) \right] - c^i(\hat{v}^i)
$$

Hence, $\mathcal{M}$ induces efficient investment within $\epsilon$ by $i$. 

93
Proof of Theorem 6

Suppose the other agents report $v^{i \setminus \{i\}}$ according to the distribution $v^{i \setminus \{i\}} \in \mathcal{V}^{i \setminus \{i\}}$. We calculate

$$
\mathbb{E}_{v^{i \setminus \{i\}}} \left[ \sup_{v^i, \bar{v}^i \in \mathcal{V}^i} \left\{ u^i(\mathcal{M}(\bar{v}^i, v^{i \setminus \{i\}}); v^i) - u^i(\mathcal{M}(v^i, v^{i \setminus \{i\}}); v^i) \right\} \right]
$$

$$
= \mathbb{E}_{v^{i \setminus \{i\}}} \left[ \sup_{v^i, \bar{v}^i \in \mathcal{V}^i} \left\{ u^i(\mathcal{M}(\bar{v}^i, v^{i \setminus \{i\}}); v^i) - u^i(\mathcal{M}(\bar{v}^i, v^{i \setminus \{i\}}); v^i) \\
+ u^i(\mathcal{M}(v^i, v^{i \setminus \{i\}}); \bar{v}^i) - u^i(\mathcal{M}(v^i, v^{i \setminus \{i\}}); \bar{v}^i) \\
+ \left( \mathcal{V}(\mathcal{M}(\bar{v}^i, v^{i \setminus \{i\}}); (\bar{v}^i, v^{i \setminus \{i\}})) - \mathcal{V}(\mathcal{M}(v^i, v^{i \setminus \{i\}}); (v^i, v^{i \setminus \{i\}})) \right) \\
+ \left( \mathcal{V}(\mathcal{M}(v^i, v^{i \setminus \{i\}}); (v^i, v^{i \setminus \{i\}})) - \mathcal{V}(\mathcal{M}(\bar{v}^i, v^{i \setminus \{i\}}); (\bar{v}^i, v^{i \setminus \{i\}})) \right) \\
+ u^i(\mathcal{M}(\bar{v}^i, v^{i \setminus \{i\}}); \bar{v}^i) - u^i(\mathcal{M}(v^i, v^{i \setminus \{i\}}); \bar{v}^i) \right\} \right]
$$

$$
\leq \mathbb{E}_{v^{i \setminus \{i\}}} \left[ \sup_{v^i, \bar{v}^i \in \mathcal{V}^i} \left\{ u^i(\mathcal{M}(\bar{v}^i, v^{i \setminus \{i\}}); v^i) - u^i(\mathcal{M}(\bar{v}^i, v^{i \setminus \{i\}}); \bar{v}^i) \\
+ \left( \mathcal{V}(\mathcal{M}(\bar{v}^i, v^{i \setminus \{i\}}); (\bar{v}^i, v^{i \setminus \{i\}})) - \mathcal{V}(\mathcal{M}(v^i, v^{i \setminus \{i\}}); (v^i, v^{i \setminus \{i\}})) \right) \\
+ \left( \mathcal{V}(\mathcal{M}(v^i, v^{i \setminus \{i\}}); (\bar{v}^i, v^{i \setminus \{i\}})) - \mathcal{V}(\mathcal{M}(\bar{v}^i, v^{i \setminus \{i\}}); (\bar{v}^i, v^{i \setminus \{i\}})) \right) \\
+ \left( \mathcal{V}(\mathcal{M}(v^i, v^{i \setminus \{i\}}); (v^i, v^{i \setminus \{i\}})) - \mathcal{V}(\mathcal{M}(\bar{v}^i, v^{i \setminus \{i\}}); (\bar{v}^i, v^{i \setminus \{i\}})) \right) \\
+ u^i(\mathcal{M}(\bar{v}^i, v^{i \setminus \{i\}}); \bar{v}^i) - u^i(\mathcal{M}(v^i, v^{i \setminus \{i\}}); \bar{v}^i) \right\} \right]
$$

$$
\leq \eta + \epsilon,
$$

where $\eta$ and $\epsilon$ are small constants.
where the first inequality follows from the subadditivity of the supremum operator and the additivity of the expectation operator, the second inequality follows from the fact that $\mathcal{M}$ provides marginal rewards within $\epsilon$ in expectation to $i$, the next equality follows from the fact that identical reports induce identical transfers, and the last inequality follows from the fact that $\mathcal{M}$ is efficient within $\eta$.\(^{73}\)

**Proof of Theorem 7**

For each $v^{I \setminus \{i\}} \in \mathcal{V}^{I \setminus \{i\}}$, let

$$\epsilon(v^{I \setminus \{i\}}) \equiv \sup_{\bar{v}^i, \tilde{v}^i \in \mathcal{V}^i} \left\{ u^i(M(\bar{v}^i, v^{I \setminus \{i\}}); v^i) - u^i(M(v^i, v^{I \setminus \{i\}}); v^i) \right\}. \quad (27)$$

Note that $E_{v^{I \setminus \{i\}}} \left[ \epsilon(v^{I \setminus \{i\}}) \right] \leq \epsilon$ for all $v^{I \setminus \{i\}} \in \mathcal{V}^{I \setminus \{i\}}$ as $\mathcal{M}$ is strategy-proof within $\epsilon$ in expectation for $i$.

Similarly, for each $v^{I \setminus \{i\}} \in \mathcal{V}^{I \setminus \{i\}}$, let

$$\eta(v^{I \setminus \{i\}}) \equiv \sup_{\bar{v}^i, \tilde{v}^i \in \mathcal{V}^i} \left\{ V(M(\bar{v}^i, v^{I \setminus \{i\}}); v^i) - V(M(v^i, v^{I \setminus \{i\}}); v^i) \right\}. \quad (28)$$

Note that $E_{v^{I \setminus \{i\}}} \left[ \eta(v^{I \setminus \{i\}}) \right] \leq \eta$ for all $v^{I \setminus \{i\}} \in \mathcal{V}^{I \setminus \{i\}}$ as $\mathcal{M}$ is efficient within $\eta$ in expectation for $i$.

Now consider $v^i, \tilde{v}^i \in \mathcal{V}^i$ for agent $i$; denote $v = (v^i, v^{I \setminus \{i\}})$ and $\tilde{v} = (\tilde{v}^i, v^{I \setminus \{i\}})$. For each $\omega \in \Omega$, let $\mathcal{V}^i_\omega \equiv \{ \tilde{v}^i \in \mathcal{V}^i : \mu(\tilde{v}^i, v^{I \setminus \{i\}}) \in \omega \}$. Since $\mathcal{V}^i = \Delta(V^i)$ is convex, the line segment $\gamma(\cdot)$ between $v^i$ and $\tilde{v}^i$ lies in $\mathcal{V}^i$, i.e., the function $\gamma : [0, 1] \to \mathcal{V}^i$ defined by

$$\gamma(y) = (1 - y)v^i + y\tilde{v}^i \text{ for all } y \in [0, 1]$$

satisfies $\gamma(0) = v^i$ and $\gamma(1) = \tilde{v}^i$ and $\gamma(y) \in \mathcal{V}^i$ for every $y \in [0, 1]$.

Fix $\delta > 0$.\(^{73}\) All other equalities follow from adding 0 to the expression.
Let $\omega_1$ be the equivalence class such that $\mu(v) \in \omega_1$, and let $\hat{v}_{\omega_1}^i \equiv v^i$ and $\tilde{v}_{\omega_1} \equiv (\hat{v}_{\omega_1}^i, v^{I \setminus \{i\}}$). We now inductively define sequences of equivalence classes, path steps, and preference profiles, indexed by $n = 1, \ldots, N$.

- Let $y_{\omega_n} \equiv \sup\{y \in [0, 1] : \gamma(y) \in \mathcal{P}_{\omega_n}^n\}$.

- If $y_{\omega_n} < 1$:
  
  - Let $\hat{y}_{\omega_n} \in \{y \in [0, 1] : y_{\omega_n} - \delta < y \leq y_{\omega_n} \text{ and } \mu(\gamma(y), v^{I \setminus \{i\}}) \in \omega_n\}; \quad (29)$

  note that $\hat{y}_{\omega_n}$ must exist as, by the definition of $y_{\omega_n}$, there exists $y$ arbitrarily close to $y_{\omega_n}$ such that $\mu(\gamma(y), v^{I \setminus \{i\}}) \in \omega_n$. Let $\hat{v}_{\omega_n}^i \equiv \gamma(\hat{y}_{\omega_n})$ and $\tilde{v}_{\omega_n} \equiv (\hat{v}_{\omega_n}^i, v^{I \setminus \{i\}}$).

  - Consider some $\tilde{y}_{\omega_{n+1}} \in \{y \in [0, 1] : y_{\omega_n} < y < y_{\omega_n} + \delta\}$;

  note that, by definition, $\mu(\gamma(\tilde{y}_{\omega_{n+1}}), v^{I \setminus \{i\}}) \notin \bigcup_{m=1}^n \omega_m$, as $\tilde{y}_{\omega_{n+1}} > y_{\omega_n}$ for all $m \in \{1, \ldots, n\}$. Let $\omega_{n+1}$ be the equivalence class such that $\mu(\gamma(\tilde{y}_{\omega_{n+1}}), v^{I \setminus \{i\}}) \in \omega_{n+1}$.

  Let $\hat{v}_{\omega_{n+1}}^i \equiv \gamma(\tilde{y}_{\omega_{n+1}})$ and $\tilde{v}_{\omega_{n+1}} \equiv (\hat{v}_{\omega_{n+1}}^i, v^{I \setminus \{i\}}$).

- Otherwise (i.e., if $y_{\omega_n} = 1$):
  
  - If $\mu(\gamma(y_{\omega_n}), v^{I \setminus \{i\}}) \in \omega_n$, then let $\hat{v}_{\omega_n} = \hat{v}^i$ and $\tilde{v}_{\omega_n} = (\hat{v}_{\omega_n}^i, v^{I \setminus \{i\}}$, and let $N = n$.

  - If $\mu(\gamma(y_{\omega_n}), v^{I \setminus \{i\}}) \notin \omega_n$, then let $\hat{v}_{\omega_n}^i \equiv \gamma(\hat{y}_{\omega_n})$, where $\hat{y}_{\omega_n}$ is as in (29), and let $\tilde{v}_{\omega_n} \equiv (\hat{v}_{\omega_n}^i, v^{I \setminus \{i\}}$). Finally, let $\hat{v}_{\omega_{n+1}}^i = \hat{v}_{\omega_{n+1}} = \tilde{v}^i$ and $\tilde{v}_{\omega_{n+1}} = \hat{v}_{\omega_{n+1}} = (\hat{v}_{\omega_{n+1}}^i, \tilde{v}^{I \setminus \{i\}})$, and let $N = n + 1$.

This construction is illustrated in Figure 2. (The construction of $\{\hat{v}_{\omega_n}^i\}_{n \in \{1, \ldots, N\}}$ and $\{\tilde{v}_{\omega_n}^i\}_{n \in \{1, \ldots, N\}}$ here is the same as in the proof of Theorem 5.) Note that $N \leq |\Omega^i|$, as the equivalence class for each $n$ is distinct, since $\mu(\gamma(\tilde{y}_{\omega_{n+1}}), v^{I \setminus \{i\}}) \notin \bigcup_{m=1}^n \omega_m$ by construction for all $n \in \{1, \ldots, N\}$.
By (27), we have

\[ u^i(M(\hat{\omega}_1); \hat{v}_i^{\hat{\omega}_1}) - u^i(M(\check{\omega}_1); \check{v}_i^{\check{\omega}_1}) \leq \epsilon(\nu^{I \setminus \{i\}}). \]

As we have

\[ u^i(M(\hat{\omega}_1); \hat{v}_i^{\hat{\omega}_1}) = u^i(M(\hat{\omega}_1); \hat{v}_i^{\hat{\omega}_1}) - \left( \hat{v}_i^{\hat{\omega}_1}(\mu(\hat{\omega}_1)) - \check{v}_i^{\check{\omega}_1}(\mu(\check{\omega}_1)) \right), \]

we have that

\[ u^i(M(\hat{\omega}_1); \hat{v}_i^{\hat{\omega}_1}) - \left( \hat{v}_i^{\hat{\omega}_1}(\mu(\hat{\omega}_1)) - \check{v}_i^{\check{\omega}_1}(\mu(\check{\omega}_1)) \right) - u^i(M(\check{\omega}_1); \check{v}_i^{\check{\omega}_1}) \leq \epsilon(\nu^{I \setminus \{i\}}). \]

Now, by (27), we obtain

\[ u^i(M(\hat{\omega}_2); \hat{v}_i^{\hat{\omega}_2}) - \left( \hat{v}_i^{\hat{\omega}_2}(\mu(\hat{\omega}_2)) - \check{v}_i^{\check{\omega}_1}(\mu(\check{\omega}_1)) \right) - u^i(M(\check{\omega}_1); \check{v}_i^{\check{\omega}_1}) \leq 2\epsilon(\nu^{I \setminus \{i\}}). \]

Since \(|\hat{\omega}_2 - \hat{\omega}_1| \leq 2\delta\), we have

\[ u^i(M(\hat{\omega}_2); \hat{v}_i^{\hat{\omega}_2}) - \left( \hat{v}_i^{\hat{\omega}_2}(\mu(\hat{\omega}_2)) - \check{v}_i^{\check{\omega}_1}(\mu(\check{\omega}_1)) \right) - u^i(M(\check{\omega}_1); \check{v}_i^{\check{\omega}_1}) \leq 2\epsilon(\nu^{I \setminus \{i\}}) + 2\delta \bar{A}, \]

where \( \bar{A} \equiv \sup_{\nu_i \in V^i, \omega \in \Omega} \{ u^i(\omega; \nu_i) \} - \inf_{\nu_i \in V^i, \omega \in \Omega} \{ u^i(\omega; \nu_i) \}. \)

Hence, we obtain

\[ u^i(M(\hat{\omega}_2); \hat{v}_i^{\hat{\omega}_2}) - \sum_{n=1}^{2} \left( \hat{v}_i^{\omega_n}(\mu(\hat{\omega}_n)) - \check{v}_i^{\omega_n}(\mu(\check{\omega}_n)) \right) - u^i(M(\check{\omega}_1); \check{v}_i^{\check{\omega}_1}) \leq 2\epsilon(\nu^{I \setminus \{i\}}) + 2\delta \bar{A}. \]

\[ ^{74} \text{Note that } \bar{A} \text{ is finite by an argument analogous to the one described in Footnote 64.} \]
Repeatedly applying the last three steps, we get

\[
u_i(\mathcal{M}(\hat{v}_\omega); \hat{v}_N) - \sum_{n=1}^{N} \left( \hat{v}_n^i(\mu(\hat{v}_n)) - \hat{v}_n^i(\mu(\tilde{v}_n)) \right) - u_i(\mathcal{M}(\hat{v}_\omega); \hat{v}_1) \\
\leq N \epsilon \left( \nu^{\left\{ i \right\}} \right) + 2(N - 1)\delta \tilde{A}. \quad (30)
\]

By (28), we have

\[
\mathbf{V}(\mathcal{M}(\hat{v}_\omega); \hat{v}_N) - \mathbf{V}(\mathcal{M}(\hat{v}_\omega); \hat{v}_N) \leq \eta \left( \nu^{\left\{ i \right\}} \right).
\]

This implies that

\[
\mathbf{V}(\mathcal{M}(\hat{v}_\omega); \hat{v}_N) - \left( \hat{v}_N^i(\mu(\hat{v}_n)) - \hat{v}_N^i(\mu(\tilde{v}_n)) \right) - \mathbf{V}(\mathcal{M}(\hat{v}_\omega); \hat{v}_N) \leq \eta \left( \nu^{\left\{ i \right\}} \right).
\]

Now, by (28), we obtain

\[
\mathbf{V}(\mathcal{M}(\hat{v}_{N-1}); \hat{v}_N) - \left( \hat{v}_{N-1}^i(\mu(\hat{v}_n)) - \hat{v}_{N-1}^i(\mu(\tilde{v}_n)) \right) - \mathbf{V}(\mathcal{M}(\hat{v}_\omega); \hat{v}_N) \leq 2\eta \left( \nu^{\left\{ i \right\}} \right).
\]

Since \(|\hat{y}_{N-1} - \hat{y}_N| \leq 2\delta\), we have

\[
\mathbf{V}(\mathcal{M}(\hat{v}_{N-1}); \hat{v}_{N-1}) - \left( \hat{v}_{N-1}^i(\mu(\hat{v}_n)) - \hat{v}_{N-1}^i(\mu(\tilde{v}_n)) \right) - \mathbf{V}(\mathcal{M}(\hat{v}_\omega); \hat{v}_N) \leq 2\eta \left( \nu^{\left\{ i \right\}} \right) + 2\delta \tilde{A}.
\]

Hence, we obtain

\[
\mathbf{V}(\mathcal{M}(\hat{v}_{N-1}); \hat{v}_{N-1}) - \sum_{n=N-1}^{N} \left( \hat{v}_n^i(\mu(\hat{v}_n)) - \hat{v}_n^i(\mu(\tilde{v}_n)) \right) - \mathbf{V}(\mathcal{M}(\hat{v}_\omega); \hat{v}_N) \leq 2\eta \left( \nu^{\left\{ i \right\}} \right) + 2\delta \tilde{A}.
\]
Repeatedly applying the last three steps, we get

\[
V(\mathcal{M}(\hat{v}_1); \hat{v}_1) - \sum_{n=1}^{N} \left( \hat{v}_n^i (\mu(\hat{v}_n)) - \hat{v}_n^i (\mu(\hat{v}_n)) \right) - V(\mathcal{M}(\hat{v}_N); \hat{v}_N)
\leq N\eta(v^{I \setminus \{i\}}) + 2(N-1)\delta A.
\]

Note that \(v^i(\mu(\hat{v}_n)) = v^i(\mu(\hat{v}_n))\) for all \(v^i \in \mathcal{V}^h\) as \(\mu(\hat{v}_n), \mu(\hat{v}_n) \in \mathcal{W}_n\) by definition. Hence,

\[
V(\mathcal{M}(\hat{v}_1); \hat{v}_1) - \sum_{n=1}^{N} \left( \hat{v}_n^i (\mu(\hat{v}_n)) - \hat{v}_n^i (\mu(\hat{v}_n)) \right) - V(\mathcal{M}(\hat{v}_N); \hat{v}_N)
\leq N\eta(v^{I \setminus \{i\}}) + 2(N-1)\delta A. \quad (31)
\]

Adding together (30) and (31) we obtain

\[
u^i(\mathcal{M}(\hat{v}_N); \hat{v}_N) - u^i(\mathcal{M}(\hat{v}_1); \hat{v}_1)
- (V(\mathcal{M}(\hat{v}_N); \hat{v}_N) - V(\mathcal{M}(\hat{v}_1); \hat{v}_1))
\leq N\left(\epsilon(v^{I \setminus \{i\}}) + \eta(v^{I \setminus \{i\}})\right) + 4(N-1)\delta A.
\]

Noting that \(\hat{v}_N = \tilde{v}^i\) and \(\hat{v}_1 = v^i\), and that \(N \leq |\Omega^i|\), we obtain:

\[
u^i(\mathcal{M}(\hat{v}); \tilde{v}^i) - u^i(\mathcal{M}(\tilde{v}); v^i) - (V(\mathcal{M}(\hat{v}); \tilde{v}) - V(\mathcal{M}(\tilde{v}); v))
\leq |\Omega^i|\left(\epsilon(v^{I \setminus \{i\}}) + \eta(v^{I \setminus \{i\}})\right) + 4(|\Omega^i| - 1)\delta A.
\]

Next, taking \(\delta \to 0\), we see that:

\[(u^i(\mathcal{M}(\tilde{v}); \tilde{v}^i) - u^i(\mathcal{M}(\tilde{v}); v^i)) - (V(\mathcal{M}(\tilde{v}); \tilde{v}) - V(\mathcal{M}(\tilde{v}); v)) \leq |\Omega^i|\left(\epsilon(v^{I \setminus \{i\}}) + \eta(v^{I \setminus \{i\}})\right).
\]
We can rewrite this as

\[
\left( V(\mathcal{M}(v^i, v^{I\setminus \{i\}}); (v^i, v^{I\setminus \{i\}})) - V(\mathcal{M}(\tilde{v}^i, v^{I\setminus \{i\}}); (\tilde{v}^i, v^{I\setminus \{i\}})) \right) \\
- \left( u^i(\mathcal{M}(v^i, v^{I\setminus \{i\}}); v^i) - u^i(\mathcal{M}(\tilde{v}^i, v^{I\setminus \{i\}}); \tilde{v}^i) \right) \\
\leq |\Omega^i| \left( \epsilon(v^{I\setminus \{i\}}) + \eta(v^{I\setminus \{i\}}) \right).
\]

Since this holds for all \( v^i, \tilde{v}^i \in \mathcal{V} \), we have that

\[
\sup_{v^i, \tilde{v}^i \in \mathcal{V}} \left\{ \left( V(\mathcal{M}(v^i, v^{I\setminus \{i\}}); (v^i, v^{I\setminus \{i\}})) - V(\mathcal{M}(\tilde{v}^i, v^{I\setminus \{i\}}); (\tilde{v}^i, v^{I\setminus \{i\}})) \right) \\
- \left( u^i(\mathcal{M}(v^i, v^{I\setminus \{i\}}); v^i) - u^i(\mathcal{M}(\tilde{v}^i, v^{I\setminus \{i\}}); \tilde{v}^i) \right) \right\} \\
\leq |\Omega^i| \left( \epsilon(v^{I\setminus \{i\}}) + \eta(v^{I\setminus \{i\}}) \right).
\]

Finally, taking an expectation with respect to \( v^{I\setminus \{i\}} \), we obtain

\[
\mathbb{E}_{v^{I\setminus \{i\}}} \left[ \sup_{v^i, \tilde{v}^i \in \mathcal{V}} \left\{ \left( V(\mathcal{M}(v^i, v^{I\setminus \{i\}}); (v^i, v^{I\setminus \{i\}})) - V(\mathcal{M}(\tilde{v}^i, v^{I\setminus \{i\}}); (\tilde{v}^i, v^{I\setminus \{i\}})) \right) \\
- \left( u^i(\mathcal{M}(v^i, v^{I\setminus \{i\}}); v^i) - u^i(\mathcal{M}(\tilde{v}^i, v^{I\setminus \{i\}}); \tilde{v}^i) \right) \right\} \right] \\
\leq |\Omega^i| (\epsilon + \eta).
\]

**Proof of Proposition 1**

The result follows immediately from Proposition C.2.

**Proof of Proposition 2**

Suppose that outcome \((\omega, t)\) is stable, and assume for the sake of contradiction that \( \omega \) is not efficient within \( |W|\epsilon \). Then there exists an alternative \( \psi \) such that

\[
V(\psi) > V(\omega) + |W|\epsilon. \tag{32}
\]

Consider a transfer vector \( \hat{t} \) defined as follows: For each \( w \in W \), if \((w, f) \in \psi \) for some \( f \in F \), let \(-\hat{t}^w \) be the smallest salary in \( \mathcal{P} \) such that

\[
v^w(\psi) + \hat{t}^w \geq u^w((\omega, t); v^w); \tag{33}
\]

100
otherwise, let \( \hat{t}^w = 0 \). For each \( f \in F \), let

\[
\hat{t}^f = - \sum_{w \in \{ \bar{w} \in W : (w, f) \in \psi \}} \hat{t}^w.
\]

Because the salary increment is at most \( \epsilon \), we have

\[
u^w((\psi, \hat{t}); v^w) \leq u^w((\omega, t); v^w) + \epsilon
\]

for any \( w \in W \).\(^{75}\) For each \( f \in F \) and \( \tilde{\omega} \in \Omega \), we denote by

\[
W_f(\tilde{\omega}) \equiv \{ w \in W : (w, f) \in \tilde{\omega} \}
\]

the set of workers matched with \( f \) at matching \( \tilde{\omega} \); we denote by \( W_\emptyset(\tilde{\omega}) \equiv \{ w \in W : \nexists f \in F, (w, f) \in \tilde{\omega} \} \) the set of workers unmatched at \( \tilde{\omega} \).

**Claim.** There exists \( f \in F \) such that

\[
v^f(\psi) + \sum_{w \in W_f(\psi)} v^w(\psi) > v^f(\omega) + \sum_{w \in W_f(\omega)} v^w(\omega) + |W_f(\psi)| \epsilon + \sum_{w \in W_f(\omega)} t^w - \sum_{w \in W_f(\psi)} t^w. \tag{35}\]

\(^{75}\)Note that as \( (\omega, t) \) is stable, it is individually rational. It follows that \( u^w((\omega, t); v^w) + \epsilon \geq u^w((\omega, t); v^w) \geq 0 \); hence, if \((w, f') \notin \psi \) for any \( f' \in F \), then we have \( u^w((\psi, \hat{t}); v^w) = 0 \leq u^w((\omega, t); v^w) + \epsilon. \)
Proof. Assume by way of contradiction that (35) fails for every $f \in F$. Then we have

$$V(\psi) = \sum_{f \in F} \left( v^f(\psi) + \sum_{w \in W_f(\psi)} v^w(\psi) \right)$$

$$\leq \sum_{f \in F} \left( v^f(\omega) + \sum_{w \in W_f(\psi)} v^w(\omega) + |W_f(\psi)| \epsilon + \sum_{w \in W_f(\omega)} t^w - \sum_{w \in W_f(\psi)} t^w \right)$$

$$\leq \sum_{f \in F} v^f(\omega) + \sum_{w \in W} v^w(\omega) - \sum_{w \in W_\emptyset(\psi)} v^w(\omega) + |W| \epsilon + \sum_{w \in W_\emptyset(\psi)} t^w$$

$$= V(\omega) + |W| \epsilon - \sum_{w \in W_\emptyset(\psi)} (v^w(\omega) - t^w)$$

$$\leq V(\omega) + |W| \epsilon,$$

(36)

where the first equality follows from the definition of $V(\psi)$, the first inequality follows from the assumption that (35) fails for every $f \in F$, the second inequality follows from rearranging terms, the definition of $W_\emptyset(\cdot)$, and noting that $\sum_{f \in F} |W_f(\psi)| \leq |W|$.

The second equality follows from the definition of $V(\omega)$ and rearranging terms, and the last inequality follows from the fact that $v^w(\omega) - t^w \geq 0$ for all $w$ (as $\omega$ is stable). Now, (36) contradicts (32). \qed

Claim. There exists $f \in F$ such that $u^f((\psi, \hat{t}); v^f) > u^f((\omega, t); v^f)$.

Proof. First note that (34) can be rewritten as

$$v^w(\psi) - \hat{t}^w \leq v^w(\omega) - t^w + \epsilon.$$  

(37)

For the second inequality in (36), we use the following identity:

$$\sum_{f \in F} \left( \sum_{w \in W_f(\omega)} t^w - \sum_{w \in W_f(\psi)} t^w \right) = \sum_{f \in F} \left( \sum_{w \in W_f(\omega)} t^w - \sum_{w \in W_f(\psi)} t^w \right) + \sum_{w \in W_\emptyset(\psi)} t^w - \sum_{w \in W_\emptyset(\psi)} t^w$$

$$= \sum_{f \in F} \left( \sum_{w \in W_f(\omega)} t^w \right) + \sum_{w \in W_\emptyset(\psi)} t^w - \sum_{w \in W_\emptyset(\psi)} t^w$$

$$= \sum_{w \in W_\emptyset(\psi)} t^w,$$

where the last equality comes from the fact that any worker who is unmatched under $\omega$ receives a transfer of 0, as $(\omega, t)$ is stable.
Let $f$ be a firm for which (35) holds; such a firm exists by the preceding claim. We obtain

\[
    u^f((\psi, \hat{t}); v^f) = v^f(\psi) - \hat{t}^f
    = v^f(\psi) + \sum_{w \in W_f(\psi)} v^w(\psi) - \sum_{w \in W_f(\psi)} (v^w(\psi) - \hat{t}^w)
    > v^f(\omega) + \sum_{w \in W_f(\omega)} v^w(\omega) + |W_f(\psi)|\epsilon + \sum_{w \in W_f(\omega)} t^w - \sum_{w \in W_f(\omega)} t^w - \sum_{w \in W_f(\psi)} (v^w(\omega) - t^w + \epsilon)
    = v^f(\omega) - t^f
    = u^f((\omega, t); v^f),
\]

where the first equality follows from the definition of $u^f$, the second equality is an identity, the inequality follows from (35) and (37), the third equality follows from rearranging terms, and the last equality follows from the definition of $u^f$.

Now, we see that $(\omega, t)$ is not stable because it is blocked. To see this, we consider the alternative $\hat{\psi} = W' \times \{f\}$, where $W' = \{w \in W : (w, f) \in \psi\}$, and the $\hat{\psi}$-compatible transfer $\hat{t}$ that is the restriction of $\hat{t}$ to $W' \cup \{f\}$, i.e., $\hat{t}^i = \hat{t}^i$ for all $i \in W' \cup \{f\}$ and $\hat{t}^i = 0$ for all $i \notin W' \cup \{f\}$. By (33) and our last claim, we see that $(W' \cup \{f\}, \hat{t})$ blocks $(\omega, t)$.

**Proof of Corollary 5**

By Theorem 11 of Hatfield and Milgrom (2005), the worker-optimal stable mechanism is strategy-proof for $w$, and, by Proposition 2, it is efficient within $|W|\epsilon$. Thus, Theorem 3 implies that the worker-optimal stable mechanism induces efficient investment within $|\Omega^w|(|W|\epsilon)$ by $w$, where $\Omega^w$ is the set of equivalence classes of alternatives for $w$.

As the valuation of a worker $w$ depends only on the identity of the firm to which he is paired (if any), we have $|\Omega^w| \leq |F| + 1$; combining this with our preceding observations shows the theorem.
Proof of Proposition 4

Before we prove Proposition 4, we first show a useful result on the speed of the convergence of differences in order statistics. Let $\mathcal{F}$ be a set of distributions over $[0, 1]$. There are $N$ random variables $X_1, X_2, \ldots, X_N$, with each $X_n$ following a (possibly different) distribution in $\mathcal{F}$. We assume that $X_1, \ldots, X_N$ are independently distributed. We assume moreover that each distribution in $\mathcal{F}$ admits a probability density function $f$; we slightly abuse notation and write $f \in \mathcal{F}$ to mean that $f$ is the probability density function for a distribution in $\mathcal{F}$. Finally, assume that $z \equiv \inf_{f \in \mathcal{F}, x \in [0, 1]} \{f(x)\} > 0$. Let $X(\ell)$ be the $\ell$th order statistic of $X_1, \ldots, X_N$, and let $X(0) \equiv 1$ and $X(N+1) \equiv 0$. Furthermore, for each $k$ and $\ell$, let $S(k, \ell) \equiv X(k) - X(\ell)$ be the difference between the $k$th and $\ell$th order statistics.

Proposition D.1. For any $m \in \mathbb{N}$ and any $\delta > 0$, $E\left[\sup_{\ell \in \{0, \ldots, N+1-m\}} \{S(\ell, \ell+m)\}\right] = O\left(\frac{1}{N^{1-\delta}}\right)$ as $N \to \infty$. In particular, $E\left[\sup_{\ell \in \{0, \ldots, N+1-m\}} \{S(\ell, \ell+m)\}\right] \to 0$ as $N \to \infty$.

Proof. We begin by partitioning the unit interval into $\left\lfloor \frac{N}{1-\delta} \right\rfloor$ segments of length $\frac{1}{\left\lfloor \frac{N}{1-\delta} \right\rfloor}$; let

$$J_k \equiv \left[ \frac{k-1}{\left\lfloor \frac{N}{1-\delta} \right\rfloor}, \frac{k}{\left\lfloor \frac{N}{1-\delta} \right\rfloor} \right),$$

where $k \in \{1, \ldots, \left\lfloor \frac{N}{1-\delta} \right\rfloor - 1\}$ and

$$J_{\left\lfloor \frac{N}{1-\delta} \right\rfloor} \equiv \left[ \frac{\left\lfloor \frac{N}{1-\delta} \right\rfloor - 1}{\left\lfloor \frac{N}{1-\delta} \right\rfloor}, 1 \right].$$

Lemma D.1. For each $k \in \{1, \ldots, \left\lfloor \frac{N}{1-\delta} \right\rfloor\}$,

$$\mathbb{P}\left[ \bigwedge_{n=1}^{N} (X_n \notin J_k) \right] \leq \left(1 - \frac{z}{N^{1-\delta}}\right)^N. \quad (38)$$

Proof. The probability, for any $n \in \{1, \ldots, N\}$ and $k \in \{1, \ldots, \left\lfloor \frac{N}{1-\delta} \right\rfloor\}$, that $X_n$ (with

\footnote{Note that if $N+1 < m$, the value of $\sup_{\ell \in \{0, \ldots, N+1-m\}} \{S(\ell, \ell+m)\}$ is undefined; however, here we are only concerned with behavior as $N \to \infty$.}
corresponding probability density function \( f_n \) is in \( \mathcal{I}_k \) is

\[
\int_{x \in \mathcal{I}_k} f_n(x) \, dx \geq \int_{x \in \mathcal{I}_k} z \, dx = z |\mathcal{I}_k| = \frac{z}{[N^{1-\delta}]} \geq \frac{z}{N^{1-\delta}}.
\]

Thus the probability that \( X_n \) is outside of \( \mathcal{I}_k \) is at most \( 1 - \frac{z}{N^{1-\delta}} \). As the \( N \) random variables \( X_1, \ldots, X_N \) are distributed independently, we obtain the desired bound \( (38) \).

Now, let \( \mathcal{E} \) be the event that there exists \( k \in \{1, \ldots, \lfloor N^{1-\delta} \rfloor \} \) such that, for all \( n \in \{1, \ldots, N\} \), we have that \( X_n \notin \mathcal{I}_k \).

**Lemma D.2.** The probability of \( \mathcal{E} \) is at most \( N \left( 1 - \frac{z}{N^{1-\delta}} \right)^N \).

**Proof.** The result follows immediately from the union bound (i.e., the fact that for any events \( \mathcal{A} \) and \( \mathcal{B} \), we have that \( \mathbb{P}[\mathcal{A} \cup \mathcal{B}] \leq \mathbb{P}[\mathcal{A}] + \mathbb{P}[\mathcal{B}] \)), the fact that there are weakly fewer than \( N \) intervals to consider, and Lemma D.1.

We now complete the proof of Proposition D.1. From Lemma D.2, we have that

\[
\mathbb{P}[\mathcal{E}] \leq N \left( 1 - \frac{z}{N^{1-\delta}} \right)^N,
\]

which converges to 0 at a rate faster than any inverse of a polynomial of \( N \); in particular \( \mathbb{P}[\mathcal{E}] = O\left( \frac{1}{N^{1-\delta}} \right). \)\(^{78}\) If the event \( \mathcal{E} \) fails, then any realizations of two consecutive order statistics of the random variables are apart from each other by at most \( \frac{2m}{[N^{1-\delta}]} \), so \( S(\ell, \ell+m) \leq \frac{2m}{[N^{1-\delta}]} = O\left( \frac{1}{N^{1-\delta}} \right) \) for any \( \ell \) and \( m \), conditional on the complement \( \bar{\mathcal{E}} \) of event \( \mathcal{E} \). Thus, we

\(^{78}\)Note that \( (1 - \frac{z}{N^{1-\delta}})^N \approx e^{-zN^\delta} \) as \( N \to \infty \).
We now prove Proposition 4. For the uniform-price auction and fixed $\delta > 0$, we calculate

$$
\mathbb{E}_{v^I \sim (\cdot)} \left[ \sup_{v^1, \bar{v}^1 \in \mathcal{V}} \left\{ u^i(\mathcal{M}(\bar{v}^i, v^I \setminus \{i\}); v^i) - u^i(\mathcal{M}(v^i, v^I \setminus \{i\}); v^i) \right\} \right]
$$

$$
= \mathbb{E}_{v^I \sim (\cdot)} \left[ \sup_{v^1, \bar{v}^1 \in \mathcal{V}} \left\{ \sum_{n=1}^{\mu(v)} \left( a^i_n - p(\bar{v}^i, v^I \setminus \{i\}) \right) - \sum_{n=1}^{\mu(v)} \left( a^i_n - p(v^i, v^I \setminus \{i\}) \right) \right\} \right]
$$

$$
= \mathbb{E}_{v^I \sim (\cdot)} \left[ \sup_{v^1, \bar{v}^1 \in \mathcal{V}} \left\{ \sum_{n=1}^{M} \max\left\{ a^i_n - p(\bar{v}^i, v^I \setminus \{i\}), 0 \right\} - \sum_{n=1}^{M} \max\left\{ a^i_n - p(v^i, v^I \setminus \{i\}), 0 \right\} \right\} \right]
$$

$$
\leq \mathbb{E}_{v^I \sim (\cdot)} \left[ \sup_{v^1, \bar{v}^1 \in \mathcal{V}} \left\{ \sum_{n=1}^{M} \max\left\{ a^i_n - p(\bar{v}^i, v^I \setminus \{i\}), 0 \right\} - \sum_{n=1}^{M} \max\left\{ a^i_n - p(v^i, v^I \setminus \{i\}), 0 \right\} \right\} \right]
$$

$$
= M \mathbb{E}_{v^I \sim (\cdot)} \left[ \sup_{v^1, \bar{v}^1 \in \mathcal{V}} \left\{ \sum_{n=1}^{M} \left| p(\bar{v}^i, v^I \setminus \{i\}) - p(v^i, v^I \setminus \{i\}) \right| \right\} \right]
$$

$$
= O\left( \frac{1}{|I|^{1-\delta}} \right),
$$

where:

1. the first equality follows from the definition of the uniform-price auction;

2. the second equality follows as the uniform-price auction only allocates an object to an
agent if his marginal value for that object is at least the price;

3. the first inequality follows as \( \max \{ a_i^n - p(\bar{v}^i, v^{I \setminus \{i\}}, 0) \} \geq 0; \)

4. the next equality follows from arithmetic manipulation;

5. the second inequality follows as, if \( p(\bar{v}^i, v^{I \setminus \{i\}}) \geq p(v^i, v^{I \setminus \{i\}}) \), then

\[
\max \{ a_i^n - p(\bar{v}^i, v^{I \setminus \{i\}}), 0 \} - \max \{ a_i^n - p(v^i, v^{I \setminus \{i\}}), 0 \} \leq 0
\]

for each \( n \in \{1, \ldots, M\} \) and, if \( p(\bar{v}^i, v^{I \setminus \{i\}}) \leq p(v^i, v^{I \setminus \{i\}}) \), then

\[
\max \{ a_i^n - p(\bar{v}^i, v^{I \setminus \{i\}}), 0 \} - \max \{ a_i^n - p(v^i, v^{I \setminus \{i\}}), 0 \} \leq p(v^i, v^{I \setminus \{i\}}) - p(\bar{v}^i, v^{I \setminus \{i\}})
\]

for each \( n \in \{1, \ldots, M\} \).

6. the next equality follows from arithmetic manipulation; and

7. the last inequality follows from Proposition D.1.

Thus, we see that the uniform-price auction is strategy-proof within \( \frac{K}{|I|^{1-\delta}} \) in expectation for each agent \( i \in I \).
Proof of Proposition 5

Define $a^{(|B|+1)} \equiv 0$ and $e^{(|S|+1)} \equiv 1$. We first consider $i \in B$. Let $\hat{a}^{(\ell)}$ denote the $\ell$th order statistic of $((a_{b})_{b \in B \setminus \{i\}}, 0, 1)$. We calculate that

$$\mathbb{E}_{\mathcal{I} \setminus \{i\}} \left[ \sup_{\nu' \in \mathcal{V}^n} \left\{ V(\omega; (\nu', \nu'^{\setminus \{i\}})) - V(\mathcal{M}(\nu', \nu^{\setminus \{i\}}); (\nu', \nu'^{\setminus \{i\}})) \right\} \right]
$$

$$= \mathbb{E}_{\mathcal{I} \setminus \{i\}} \left[ \sup_{\nu' \in \mathcal{V}^n} \left\{ a^{(k)} - e^{(k)} \right\} \right]
$$

$$= \mathbb{E}_{\mathcal{I} \setminus \{i\}} \left[ \sup_{\nu' \in \mathcal{V}^n} \left\{ a^{(k)} - a^{(k+1)} + a^{(k+1)} - e^{(k+1)} + e^{(k+1)} - e^{(k)} \right\} \right]
$$

$$\leq \mathbb{E}_{\mathcal{I} \setminus \{i\}} \left[ \sup_{\nu' \in \mathcal{V}^n} \left\{ a^{(k)} - a^{(k+1)} + e^{(k+1)} - e^{(k)} \right\} \right]
$$

$$\leq \mathbb{E}_{\mathcal{I} \setminus \{i\}} \left[ \sup_{\nu' \in \mathcal{V}^n} \left\{ a^{(k)} - a^{(k+1)} \right\} + \mathbb{E}_{\mathcal{I} \setminus \{i\}} \left[ \sup_{\nu' \in \mathcal{V}^n} \left\{ e^{(k+1)} - e^{(k)} \right\} \right] \right]
$$

$$\leq \mathbb{E}_{\mathcal{I} \setminus \{i\}} \left[ \sup_{\nu' \in \mathcal{V}^n} \left\{ \sum_{\ell \in [0, \ldots, |B| - 1]} \left\{ \hat{a}^{(\ell)} - \hat{a}^{(\ell+1)} \right\} \right\} \right] + \mathbb{E}_{\mathcal{I} \setminus \{i\}} \left[ \sup_{\nu' \in \mathcal{V}^n} \left\{ e^{(\ell+1)} - e^{(\ell)} \right\} \right]
$$

$$\leq O\left(\frac{1}{|B|^{1-\delta}}\right) + O\left(\frac{1}{|S|^{1-\delta}}\right)
$$

where:

1. the first equality follows from the fact that the double auction executes all non-negative surplus trades except one that has lowest-surplus among the non-negative surplus trades (i.e., the $k$th trade), while not executing any negative surplus trades;

2. the second equality is obtained by adding $0$;

3. the first inequality follows since $a^{(k+1)} - e^{(k+1)} \leq 0$ as only the first $k$ trades generate positive surplus;

4. the second inequality follows from the subadditivity of the supremum;

5. the third inequality follows as both $\sup_{\ell \in [0, \ldots, |B|]} \left\{ a^{(\ell)} - a^{(\ell+1)} \right\} \geq a^{(k)} - a^{(k+1)}$ and $\sup_{\ell \in [0, \ldots, |B|]} \left\{ e^{(\ell)} - e^{(\ell+1)} \right\} \geq e^{(k)} - e^{(k+1)}$, a priori;
6. the fourth inequality follows as the supremum over the differences between consecutive order statistics (weakly) increases when one *non-extremal realization* is removed; and

7. the last inequality follows from applying Proposition D.1 to each term.

The case in which $i$ is a seller is analogous. Thus, we see that the double auction is efficient within $K \left( \frac{1}{|B|-s} + \frac{1}{|S|-s} \right)$ in expectation.

**Proof of Proposition C.2**

Let $\bar{v}^s \in V^s$ be the valuation for $s$ such that $s$ is always unwilling to trade at any price acceptable to other agents, that is, $\bar{v}^s(\Psi) < -\sum_{j \in I \setminus \{s\}} v^j(\Psi)$ for all $\Psi$ such that $\Psi_s \neq \emptyset$ and $v^I \setminus \{s\} \in V^I \setminus \{s\}$, and consider valuation $v^s \in V^s$ other than $\bar{v}^s$. Let $\bar{v} = (\bar{v}^s, v^I \setminus \{s\})$. It is sufficient to show that

$$u^s(\mathcal{M}^*(v); v^s) - u^s(\mathcal{M}^*(\bar{v}); \bar{v}^s) = V(\mathcal{M}^*(v); v) - V(\mathcal{M}^*(\bar{v}); \bar{v}),$$

as to compare any two arbitrary valuations for $s$ we can simply compare each valuation to $\bar{v}^s$.

**Claim.** $u^s(\mathcal{M}^*(v); v^s) - u^s(\mathcal{M}^*(\bar{v}); \bar{v}^s) \leq V(\mathcal{M}^*(v); v) - V(\mathcal{M}^*(\bar{v}); \bar{v})$.

**Proof.** We suppose, by way of contradiction, that

$$u^s(\mathcal{M}^*(v); v^s) - u^s(\mathcal{M}^*(\bar{v}); \bar{v}^s) > V(\mathcal{M}^*(v); v) - V(\mathcal{M}^*(\bar{v}); \bar{v}). \quad (39)$$

As $\mathcal{M}^*$ is budget-balanced,

$$V(\mathcal{M}^*(v); v) - V(\mathcal{M}^*(\bar{v}); \bar{v}) = \sum_{i \in I} u^i(\mathcal{M}^*(v); v^i) - \sum_{i \in I} u^i(\mathcal{M}^*(\bar{v}); \bar{v}^i);$$

By *non-extremal realization*, we mean a realization that is not strictly greater than all other realizations and not strictly less than all other realizations. Note that $a^b$ can not be an extremal realization for any $b \in B$, as $a^{(|B|+1)} = 0$ and $a^{(0)} = 1$. 

109
hence, (39) implies that
\[ u^s(\mathcal{M}^*(v); v^s) - u^s(\mathcal{M}^*(\bar{v}); \bar{v}^s) > \sum_{i \in I} u^i(\mathcal{M}^*(v); v^i) - \sum_{i \in I} u^i(\mathcal{M}^*(\bar{v}); \bar{v}^i). \] (40)

By rearranging (40), we obtain
\[ \sum_{i \in I \setminus \{s\}} u^i(\mathcal{M}^*(\bar{v}); \bar{v}^i) > \sum_{i \in I \setminus \{s\}} u^i(\mathcal{M}^*(v); v^i). \] (41)

We also have that
\[ \sum_{i \in I \setminus \{s\}} v^i(\mathcal{M}^*(\bar{v}); \bar{v}^i) = \sum_{i \in I \setminus \{s\}} \left( u^i(\mathcal{M}^*(\bar{v}); \bar{v}^i) + r^i(\bar{v}) \right) \\
= \sum_{i \in I \setminus \{s\}} u^i(\mathcal{M}^*(\bar{v}); \bar{v}^i) + \sum_{i \in I \setminus \{s\}} r^i(\bar{v}) \\
= \sum_{i \in I \setminus \{s\}} u^i(\mathcal{M}^*(\bar{v}); \bar{v}^i) + \sum_{i \in I} r^i(\bar{v}) \\
= \sum_{i \in I \setminus \{s\}} u^i(\mathcal{M}^*(\bar{v}); \bar{v}^i) > \sum_{i \in I \setminus \{s\}} u^i(\mathcal{M}^*(v); v^i), \] (42)

where the first line follows from the definition of the utility function, the second line follows upon reorganization of terms, the third line follows from the fact that \( \mathcal{M}^* \) provides a transfer of 0 to \( s \) under valuation \( \bar{v}^s \),\(^{80}\) the fourth line follows from the fact that \( \mathcal{M}^* \) is budget-balanced, and the fifth line follows from (41).

Finally, since \( \mathcal{M}^* \) is efficient, and \( \bar{v}^s \) has been chosen so that no efficient set of trades involves a trade with \( s \), (42) implies that
\[ \max_{\Psi \subseteq \Omega \setminus \{s\}} \left\{ \sum_{i \in I \setminus \{s\}} v^i(\Psi) \right\} > \sum_{i \in I \setminus \{s\}} u^i(\mathcal{M}^*(v); v^i). \] (43)

---

\(^{80}\)The transfer from \( s \) must be 0 under \( \mathcal{M}^* \) as \( s \) does not trade at any efficient outcome under the valuation \( \bar{v}^s \) and hence must receive a transfer of 0: see (11).
From (43), we see that \( \mathcal{M}^*(v) \) is not in the core at valuation profile \( v \), as the set of agents \( I \setminus \{s\} \) can do strictly better trading among themselves than under the outcome \( \mathcal{M}^*(v) \). Hence, by Theorems 5 and 9 of Hatfield et al. (2013), \( \mathcal{M}^*(v) \) does not correspond to a competitive equilibrium, contradicting the assumption that \( \mathcal{M}^* \) is the seller-optimal stable mechanism.

\[ \text{Claim. } u^s(\mathcal{M}^*(v); v^s) - u^s(\mathcal{M}^*(\bar{v}); \bar{v}^s) \geq V(\mathcal{M}^*(v); v) - V(\mathcal{M}^*(\bar{v}); \bar{v}). \]

\[ \text{Proof. } \]

We slightly abuse notation by writing

\[ u^i([\Psi; p]; v^i) \equiv u^i \left( \left( \Psi, \left( \sum_{\psi \in \Psi} p_{\psi} - \sum_{\psi \in \Psi_{\rightarrow j}} p_{\psi} \right) \right); v^i \right). \]

Let \( \bar{p} \) be the vector of prices defined in Proposition C.1 and let \( \Psi \) be an efficient set of trades, so that \([\Psi; \bar{p}]\) is a competitive equilibrium. Let \( \Phi \in \arg \max_{\Xi \subseteq \Omega_{I \setminus \{s\}}} \{ \sum_{j \in I \setminus \{s\}} v^j(\Xi) \} \).

We now consider an augmented economy with an additional agent \( b \) and an additional trade \( \varphi \), where \( b(\varphi) = b \) and \( s(\varphi) = s \). The valuation function of \( b \) is given by

\[ v^b(\Xi) = \begin{cases} \sum_{i \in I} v^i(\Psi) - \sum_{i \in I} v^i(\Phi) & \Xi_b = \{\varphi\} \\ 0 & \text{otherwise.} \end{cases} \]

We extend the valuation function \( v^s \) to \( \hat{v}^s \), where

\[ \hat{v}^s(\Xi) = \begin{cases} 0 & \Xi_s = \{\varphi\} \\ -\bar{A} & |\Xi_s| > 1 \text{ or } |\Xi_{\rightarrow s}| > 0 \\ v^s(\Xi) & \text{otherwise,} \end{cases} \]

and where \( \bar{A} \) is sufficiently large that no set of trades \( \Xi \) such that \( |\Xi_s| > 1 \) or \( |\Xi_{\rightarrow s}| > 0 \) is efficient. We keep the valuation function for every other agent unchanged.

---

**Footnote:** An arrangement \([\Psi; p]\) is in the core if, there does not exist another arrangement \([\Phi; q]\) such that, for all \( i \in b(\Phi) \cup s(\Phi) \), \( u^i([\Phi; q]; v^i) > u^i([\Psi; p]; v^i) \).
Note that $\Psi$ is an efficient set of trades in the augmented economy by construction. By Theorem 1 of Hatfield et al. (2013), a competitive equilibrium $[\Xi; q]$ exists in the augmented economy for some price vector $q = (q_\Omega, q_\varphi)$, where $q_\Omega$ is the profile of prices for trades in $\Omega$ and $q_\varphi$ is the price of $\varphi$. Thus, as $\Psi$ is an efficient set of trades in the augmented economy, by Theorem 3 of Hatfield et al. (2013), $[\Psi; q]$ is a competitive equilibrium in the augmented economy. This fact has two implications:

- $q_\varphi \geq \sum_{i \in I} v^i(\Psi) - \sum_{i \in I} v^i(\Phi)$, as otherwise $b$ will strictly demand $\{\varphi\}$ at price $q_\varphi$. This fact implies that $u^s([\Psi; \bar{v}^s]; v^s) \geq \sum_{i \in I} v^i(\Psi) - \sum_{i \in I} v^i(\Phi)$, as otherwise $s$ would strictly demand $\{\varphi\}$.

- the arrangement $[\Psi; q_\Omega]$ is a competitive equilibrium of the original economy. Hence, by Proposition C.1, $q_\Omega \leq \bar{p}$.

Combining the preceding two observations implies that

\[
u^s([\Psi; \bar{p}]; v^s) \geq \sum_{i \in I} v^i(\Psi) - \sum_{i \in I} v^i(\Phi).
\]

Now, we note that:

- $u^s([\Psi; \bar{p}]; v^s) = u^s(\mathcal{M}^*(v); v^s)$, by the definition of the seller-optimal stable mechanism $\mathcal{M}^*$;

- $u^s(\mathcal{M}^*(\bar{v}); \bar{v}^s) = 0$, as $\mathcal{M}^*$ is efficient and $s$ does not trade at any efficient outcome under the valuation $\bar{v}^s$;

- $\sum_{i \in I} v^i(\Psi) = \Gamma(\mathcal{M}^*(v); v)$, as $\Psi$ is efficient at $v$ and $\mathcal{M}^*$ is efficient; and

- $\sum_{i \in I} v^i(\Phi) = \Gamma(\mathcal{M}^*(\bar{v}); \bar{v})$ since $\Phi$ is efficient at $\bar{v}$—as $\Phi \in \arg\max_{\Xi \subseteq \Omega \setminus \{s\}} \left\{ \sum_{j \in I \setminus \{s\}} v^i(\Xi) \right\}$ and $s$ does not trade at any efficient outcome under the valuation $\bar{v}^s$—and the seller-optimal stable mechanism is efficient.
Combining the preceding four observations with inequality (44), we have that

\[ u^s(M^*(v); v^s) - u^s(M^*(\bar{v}); \bar{v}^s) \geq V(M^*(v); v) - V(M^*(\bar{v}); \bar{v}). \]

Together, the two preceding claims show that

\[ u^s(M^*(v); v^s) - u^s(M^*(\bar{v}); \bar{v}^s) = V(M^*(v); v) - V(M^*(\bar{v}); \bar{v}), \]

i.e., \( M^* \) provides marginal rewards to \( s \).

### E An Approximate Green–Laffont–Holmström Theorem

We say that a mechanism \( M = (\mu, r) \) has a transfer rule \( r^i \) within \( \epsilon \) of a Groves rule if all \( v \in V \) and \( \bar{v}^i \in V^i \),

\[
\left| (r^i(v) - r^i(\bar{v}^i, v^{I \setminus \{i\}})) - \left( \sum_{j \in I \setminus \{i\}} v^j(M(\bar{v}^i, v^{I \setminus \{i\}})) - \sum_{j \in I \setminus \{i\}} v^j(M(v)) \right) \right| \leq \epsilon.
\]

**Theorem E.1** (Approximate Green–Laffont–Holmström Theorem).

1. Any mechanism that is efficient within \( \eta \) and has a transfer rule \( r^i \) within \( \epsilon \) of a Groves rule is strategy-proof within \( \epsilon + \eta \) for \( i \).

2. If \( V^i \) is path-connected, then any mechanism that is efficient within \( \eta \) and strategy-proof within \( \epsilon \) for \( i \) has a transfer rule \( r^i \) within \( |\Omega^i| (\epsilon + \eta) \) of a Groves rule.

**Proof.** Consider the valuation profile \( v \in V \) and some other report \( \bar{v}^i \in V^i \) for \( i \). For ease of notation, we let \( \bar{v} = (\bar{v}^i, v^{I \setminus \{i\}}) \).

To show the first part of the theorem, we first show that any mechanism \( M = (\mu, r) \) that has transfer rule \( r^i \) within \( \epsilon \) of a Groves rule provides marginal rewards within \( \epsilon \) to \( i \). To see
this, we compute that

\[
|u^i(M(v); v^i) - u^i(M(\tilde{v}); \tilde{v}^i)) - (V(M(v); v) - V(M(\tilde{v}); \tilde{v}))|
\]

\[
= |((v^i(M(v)) - r^i(v)) - (\tilde{v}^i(M(\tilde{v}) - r^i(\tilde{v}))) - (V(M(v); v) - V(M(\tilde{v}); \tilde{v}))|
\]

\[
= \left| r^i(v) - r^i(\tilde{v}) - \left( \sum_{j \in I \setminus \{i\}} v^j(M(\tilde{v}, v^{1 \setminus \{i\}})) - \sum_{j \in I \setminus \{i\}} v^j(M(v)) \right) \right|
\]

\[
\leq \epsilon.
\]

Now, as \( M \) provides marginal rewards within \( \epsilon \) to \( i \) and is efficient within \( \eta \), our Theorem 2 implies that \( M \) is strategy-proof within \( \epsilon + \eta \) for \( i \).

To show the second part of the theorem, we first note that Theorem 3 implies that any mechanism \( M = (\mu, r) \) that is efficient within \( \eta \) and strategy-proof within \( \epsilon \) for \( i \) provides marginal rewards within \( |\Omega^i| (\epsilon + \eta) \) to \( i \). Hence,

\[
|\Omega^i| (\epsilon + \eta) \geq |(u^i(M(v); v^i) - u^i(M(\tilde{v}); \tilde{v}^i)) - (V(M(v); v) - V(M(\tilde{v}); \tilde{v}))|
\]

\[
= |((v^i(M(v)) - r^i(v)) - (\tilde{v}^i(M(\tilde{v}) - r^i(\tilde{v}))) - (V(M(v); v) - V(M(\tilde{v}); \tilde{v}))|
\]

\[
= \left| (v^i(M(v)) - r^i(v)) - (\tilde{v}^i(M(\tilde{v}) - r^i(\tilde{v})) - \left( \sum_{j \in I} v^j(M(v)) - \sum_{j \in I} (\tilde{v}^j(M(\tilde{v}))) \right) \right|
\]

\[
= \left| r^i(v) - r^i(\tilde{v}) - \left( \sum_{j \in I \setminus \{i\}} v^j(M(\tilde{v})) - \sum_{j \in I \setminus \{i\}} v^j(M(v)) \right) \right|
\]

Thus, we see that the transfer rule of \( M \) is within \( |\Omega^i| (\epsilon + \eta) \) of a Groves rule. \( \square \)