Screening, Signalling and Costly Misrepresentation*

Raymond Deneckere[†] Sergei Severinov[‡]

This version: November 2 2018

Abstract

We study mechanism design and signalling in environments where misrepresenting private information is costly. Specifically, a privately informed agent has to take several signalling actions, send several messages or undergo tests in which it is costly for her to misrepresent her type. We establish conditions under which the set of implementable allocation profiles strictly increases in the number of messages while the overall communication cost diminishes. We then derive an optimal screening mechanism in such a setting. A surprising property of this mechanism is the absence of exclusion. Particularly, every consumer type whose valuation for the good exceeds the marginal cost of production is allocated a positive quantity. Reexamining job-market signalling via education in our set-up, we show that welfare losses from unproductive signalling will be small if students signal via a grade point average based on multiple courses rather than via duration of schooling. Thus, our results explain why employers often prefer to screen applicants via multiple interviews rather than via menus of contracts, and why the social losses from signalling activities ("rat race") may not be significant.

JEL Nos: C72, D82, D86

Keywords: costly misrepresentation, multiple signals, screening, signalling, mechanism design.

*We thank Yeon Koo Che, Bart Lipman and seminar participants at University of Arizona, University of Essex, and University of Missouri-Columbia, and participants at numerous conferences for helpful comments and suggestions. Sergei Severinov acknowledges financial support from Social Sciences and Humanities Research Council of Canada through SSHRC grant # 410-2009-1070. The standard disclaimer applies.

†Department of Economics, University of Wisconsin, Madison, 1180 Observatory Drive, Madison, 53706, WI, USA. Email: rjdeneck@wisc.edu.

‡Vancouver School of Economics, University of British Columbia, 6000 Iona Drive, Vancouver, V6T 1L4, BC, Canada and National Research University Higher School of Economics, Shabolovka 26, Moscow, Russia; Email: sseverinov@gmail.com.

1 Introduction

In this paper we study mechanism design and signalling in settings where agents incur some cost of conveying and misrepresenting their private information. These costs may exist for several reasons. First, in order to signal and differentiate themselves agents may have to take costly actions extraneous to the socially productive activities (e.g. excessive education).

Second, misrepresenting the truth may require costly actions, such as acquiring skills and/or technology for manufacturing evidence, and taking effort to conceal one's information or hide evidence that reveals the true state of the world. For example, in order to obtain a supplier contract, or to qualify for a loan, win a grant or a promotion, a firm or an individual may need to be perceived as highly productive, successful and/or creditworthy. This goal may be attained by exhibiting "evidence" exaggerating prior performance and concealing the risk of default or non-performance. Yet, the production of such favorable but inaccurate "evidence" would normally require expending cost and effort.[1]

Third, an individual may find it costly to misrepresent the truth for psychological or ethical reasons. Such individual may experience stress or discomfort from lying. Behavioral psychologists have extensively documented the physical symptoms associated with the emotional discomfort that people experience when lying.

Apart from some mild technical conditions, our model relies on two simple main assumptions:

- The first-best or least cost messages vary across agent types;
- Agents send multiple signals or messages.

We believe that these key two parts of our approach are natural, well motivated, and strongly grounded in theory and empirics.

To highlight this, let us start with the first main assumption, that the first-best or least costly messages are type dependent. Consider one of the prime examples of our paper, Spence's education model. In this set-up, it is compelling to envisage that individuals derive a direct utility from learning, and that this utility is correlated with ability. Indeed, driven by intellectual curiosity, entertainment value or a taste for knowledge and problem-solving, most people like to learn. It is natural to posit that some people like learning more than others, and that their taste for learning increases with ability. As a consequence, each type has a

[1] A prominent example of state falsification efforts concerns college admissions tests. Educators have long been concerned that extensive test preparation and tutoring activities by the privileged skew the measurement of underlying student ability. Indeed, the College Board recently announced plans to redesign the SAT test to address this issue (Richard Perez-Pena, "Then and Now: A Test that Aims to Neutralize Advantages of the Privileged," *New York Times*, March 15, 2014).

different least-cost level of education. Alternatively, education may enhance productivity, and may do so more for higher ability individuals. Indeed, Spence himself explicitly recognized this possibility in his retrospective review on the signaling model (Spence, 2002). It is fair to say that the assumption of type dependent least cost signals is ubiquitous in the signaling literature, spanning a wide range of applications such as limit pricing by a monopolist (Milgrom and Roberts (1982)), advertising (Bagwell (2007)), oligopoly pricing (Mailath (1989)), dividend signaling (Miller and Rock (1985)), electoral competition (Banks (1990)), bequests (Bernheim and Severinov (2003)) and conspicuous consumption (Ireland (1994)).

Furthermore, there are several strands of influential literature on screening that make the same or similar assumption that the least cost message is type dependent. First, there is an extensive literature on certifiable statements or hard evidence whose availability varies with type. This literature, which originates in the seminal contributions of Milgrom (1981) and Grossman (1981), assumes that for any agent-type some messages are costless, while others are infinitely costly, and that the set of costless messages varies with types. The extensive literature that flows from these papers (e.g. Seidmann and Winter (1997), Hagenbach, Koessler and Perez-Richet (2014)) makes the strong assumption that every type can send some set of messages that uniquely identify it.

A related branch of literature on mechanism design and implementation with hard evidence is more permissive in its assumptions, requiring only partial verifiability, but retains the assumption of binary communication cost so that each message is either costless or infinitely costly for a given type. It explicitly allows for the presentation of multiple pieces of evidence, as we do. This literature includes Green and Laffont (1986), Lipman and Seppi (1995), Forges and Koessler (2005), Bull and Watson (2007), Caillaud and Tirole (2007), Deneckere and Severinov (2008), Sher and Vohra (2015), Ben Porath and Lipman (2012), and Hart, Kremer and Perry (2017). A binary communication cost structure also underlies the literature on honesty (Alger and Ma (2003), Alger and Renault (2006, 2007), Kartik (2009), and Severinov and Deneckere (2006)).

In contrast to both of these lines of literature, our paper makes the much milder assumption that all types can send any message available to any other type, but at a cost that is increasing in the magnitude of type misrepresentation. In this assumption we follow the literature on costly state falsification, which originates in the work of Lacker and Weinberg (1989), and includes Maggi and Rodriguez-Clare (1995), Crocker and Morgan (1998), Crocker and Slemrod (2005), Goldman and Slezak (2006), Kartik (2009), and Picard (2013).

The costly state falsification literature spans a wide range of applications, including insurance and tax fraud, financial misreporting, and legal evidence production. In a legal setting, Bull (2008a) studies costly evidence production and disclosure under complete information.

Bull (2009) compares inquisitorial and adversarial litigation systems when evidence may be suppressed at cost. Following reporting scandals such as Enron, managerial misreporting has attracted a lot of attention in the accounting and finance literature. Models involving costly misrepresentation of earnings include Fisher and Verecchia (2000), Goldman and Slezak (2006), Guttman, Kadan and Kandel (2006) and Caskey, Nagar and Petacchi (2010).

Our assumption that misrepresenting type is costly and type-dependent is also motivated by the considerable empirical and experimental evidence showing that agents do not lie as often as would be warranted by maximizing their material payoffs. In the tax literature, Erard and Feinstein (1994) report findings that some tax payers are willing to bear their full tax burden even when presented with financial incentives to underreport their incomes. Survey evidence indicates that while a group of people has no qualms about inflating insurance claims, a greater fraction considers it unacceptable to do so (Tennyson (1997)). A large body of experimental evidence initiated by Gneezy (2005) shows that there are intrinsic costs to misrepresentation, which typically depend upon the "size of the lie." Aversion to misrepresentation or "lying costs" have also been extensively documented in the broader literature on finance, psychology and sociology. Abeler, Nosenzo and Raymond (2016) provide an extensive survey of this literature.

Our second main assumption, that individuals undergo multiple tests or send multiple messages appears to be equally well-motivated. In the signaling literature, the presence of multiple signals to convey private information has long been recognized. For example, in markets for new goods, a firm may use price, advertising, warranties, slotting allowances or brand name to signal the quality of its product to its potential customers. Similarly, the corporate finance literature has considered a variety of signals conveying the future profitability of a firm, such as financial structure (Ross, 1977), dividend policy (Miller and Rock, 1985), stock splits (McNichols and Dravid (1990)), share buybacks. Additionally, in evolutionary biology, scholars have extensively studied the multi-component nature of signals animals use to signal their quality in activities such as courtship and mating behavior or predator deterrence (Johnstone, 1996).

As a final example, consider Spence's celebrated education model. In this context, the important signalling costs are those that are negatively correlated with ability - studying, preparing assignments, and taking tests and exams. The associated costs are necessarily incurred in several tasks, and along different dimensions. For example, a student who majors in mathematics has to take courses in different fields, such as Analysis, Geometry, and Algebra. Each of these fields has a different apparatus and analytical methods, so studying each of them requires some qualitatively new cognitive effort. Therefore it is natural to think of a grade in each of these math courses as a different signal.

The presence of multidimensional signals has also been studied from a theoretical perspective (Cho and Sobel (1990), Ramey (1996)), where the emphasis has been on finding conditions on the primitives and refinements of sequential equilibrium that select separating equilibria.

The literature on screening has so far studied only a few environments where agents send multiple messages. First, multiple pieces of evidence serve as messages in the literature on mechanism design with verifiability and hard evidence cited above. Second, legal scholars have studied the presentation of multiple pieces of evidence and multiple rounds of questioning in court (e.g., Emons and Fluet (2009)).

At the same time, examples in which principals screen agents by eliciting multiple messages are quite common. For example, in the hiring of new professional employees (faculty members, engineers, consultants, accountants) a potential hire goes through multiple interviews with different colleagues and administrators. Each interviewer asks different questions, and evaluates the candidate from a different angle. As another set of examples, messages may correspond to the outcome of an inspection or an audit undertaken by the principal. For example, shareholders or the parent corporation may carry out several accounting and other audits of their subsidiary. The subsidiary's managers would then have to incur the cost of hiding or embellishing the true state of the world and fudging the numbers during each audit. Another example concerns managers who have to defend their financial reports to a variety of different stakeholders and outside experts, including the board of directors, external auditors, financial analysts and credit rating agencies, shareholders, creditors and employees, who each question the reports from a different angle.

The major insight of our paper is that the simple juxtaposition of these two common assumptions, the type-dependence of the least cost messages and the availability of multiple messages, has strong consequences. In the signaling context, we show that when the number of available signals becomes sufficiently large, any sequential equilibrium that satisfies a very mild refinement must be very informative and identify elements in very fine partitions of types. Furthermore, in such equilibria the sender's signalling costs become insignificantly small.

In the screening context, we show that with sufficiently many messages, the principal can come arbitrarily close to implementing any decision rule and any surplus allocation, and at the same time keep misrepresentation costs arbitrarily small.

These conclusions are important, because ever since Spence's (1973) seminal contribution, economists have been concerned with the potential loss of welfare due to signalling. Specifically, Spence demonstrates that job-applicants, engaged in a competitive 'rat race,' will spend too much time and effort on education in order to provide an informative signal about their ability, even though it may not enhance their productivity. Similarly, in the screening context, economists have identified significant inefficiencies resulting from the agents' incentive to earn information rents by manipulating their private information and the principals' desire to limit those information rents. Thus, our paper identifies important conditions under which such welfare losses are relatively small.

Our analysis focuses on four issues. First, we explore the effect of increasing the dimension of the signal space in signaling models in which the least cost message is type-dependent. Second, we study implementation in screening models where the agent can send messages or signals along a large number of dimensions. Third, we characterize the screening mechanism that maximizes the principal's expected profits when the number of signal dimensions is limited. That part of our analysis in related to the contributions by Maggi and Rodriguez-Clare (1995) and Crocker and Morgan (1998), and we comment more on this in Section 4. We establish an important qualitative property of the optimal mechanism - absence of exclusion. Specifically, when costly messages are available, then each agent-type who can generate a positive surplus is assigned a non-zero allocation in the optimal mechanism. Thus we establish that the standard result on the optimality of exclusion in optimal screening is non-robust to the availability of costly signals.[2] Finally, we present a method for choosing the optimal number of messages when the principal also incurs a fixed cost of eliciting or processing each message. Most of our results, with the exception of the characterization of the optimal mechanism, hold for a multidimensional type space.

The intuition for our results is not straightforward. Increasing the number of signal dimensions increases the cost of signaling, and as a consequence the equilibrium signal cost along any given dimension will necessarily fall. However, this does not imply that total signal costs decrease as the number of signal dimensions grows. Indeed, consider Spence's education model, where education is unproductive, and the least cost message of each agent type is the same (zero). In this model, total equilibrium signal cost is invariant to the number of signal dimensions. This is because the reward to mimicking higher ability types, the associated increase in the wage, remains invariant to the number of signal dimensions. So to prevent such mimicking from being profitable, total signal costs must also remain unchanged. A similar conclusion arises in screening models.

The novel contribution of our paper is therefore to show that when the least cost signal differs across agent types, equilibrium message/signaling costs will fall when the dimension of

[2]When costly signals are absent, exclusion is a robust property of optimal screening mechanisms. In particular, a profit-maximizing monopolist will choose not to sell to consumers whose willingness to pay for the good is not sufficiently higher than marginal cost. This holds both under both uniform and non-linear pricing, except for the non-generic case of perfectly inelastic consumer demand at price equal to marginal cost (which requires either that there are no consumers with valuations near marginal cost, or that the density of valuations is infinite at this level) (see Maskin and Riley (1984)). Exclusion also occurs in settings with multidimensional private information (see Armstrong (1996) and Rochet and Choné (1998)).

the signal space increases. To see how this can occur, fix an allocation to be implemented and let there initially be only a single message dimension. If some type θ does not misrepresent itself and sends its least cost message, then some close-by type θ' would have an incentive to mimic type θ since her cost of doing so will be quite small. The principal would then have to pay sufficient information rent to θ' to prevent such imitation. As a consequence, either costly misrepresentation by type θ or providing sufficient surplus to θ' would be necessary to prevent the latter from imitating the former.

Consider now what happens when the agent has to send a second message, along a different dimension. For any pair of types (θ', θ), since their least cost messages are different, it will be more costly for θ' to mimic θ, even if the latter were to send her least cost message in this extra dimension. As a consequence, with two messages less misrepresentation by type θ along the original dimension and/or less surplus to type θ' will be required to prevent θ' from mimicking θ.

This simple intuition explains why global incentive constraints get more relaxed as the number of signal dimensions increases, and the least cost message differs across agent. However, crucially it does not apply to local incentive constraints that involve mimicking by infinitesimally close types. This is so because the marginal cost of misrepresentation is zero at the least cost message, and adding more message dimensions does not change this fact. So simply asking the agents to send their least cost messages will never satisfy local incentive constraints, no matter how many signal dimensions there are. Importantly, the literature on screening and signalling shows that local incentive constraints are the only binding ones in many environments. Connected by the "chains" of local incentive constraints, "higher" types earn informational rents that depend on the allocations given to all "lower" types, even though global incentive constraints are non-binding.

So, in order to satisfy local incentive constraints it is necessary to introduce some degree of misrepresentation, even with a large number of messages. Yet, we demonstrate that by selecting this degree of misrepresentation judiciously the principal can implement her desirable allocation profile with overall expected misrepresentation cost that becomes small as the number of signal dimensions increases. Essentially, the degree of necessary misrepresentation diminishes at a faster rate than the rate at which the number of messages grows.

Our findings have practical relevance. In particular, they can explain why employers in a number of industries prefer to screen and interview job-applicants very thoroughly, rather than to offer self-selecting menus of contracts or strong performance incentives to them. Indeed, the interviewing process in many professional job-markets appears to be consistent with the idea of requesting multiple messages or signals from the candidate, with each signal being somewhat different from the others. For example, in the context of a departmental visit on the academic job-market a prospective candidate meets with faculty members working in different fields. It is plausible that each conversation provides an independent signal of the candidate's ability, because different faculty members, especially if they work in different fields, assess the candidate from different perspectives and inquire about different aspects of the candidate's knowledge and skills. Similar interviewing procedures are used by professional services firms in consulting, law, banking, etc. The interview in such firms involve solving cases, conversations with consultants, managers and partners. Our results imply that if a job candidate had to go through a sufficient number of such interviews, or other tests, then substantially misrepresenting his ability would be too costly. So, the employer will have a quite accurate estimate of the candidate's ability, and would not have to offer a powerful incentive scheme on the job. This is notable since the incentive schemes used in the real world are often not as strong as predicted by the contracting literature.

Our findings also shed some light on the empirical literature on signalling. Ever since Spence's (1973) groundbreaking contributions, researchers have been trying to document the existence and magnitude of signalling cost, but were having trouble doing so. For example, in the educational context, the survey by Page (2010) concludes that "the number of studies that use convincing empirical strategies to test the signaling model (in education) is short, and the evidence is mixed." Several more recent studies, including Chevalier et. al (2004), Clark and Martorell (2014), and Arteaga (2017) indicate that this cost may be quite small. Given the large number of courses taken by modern-day students before graduating from college, our model provides a potential explanation for these findings.

2 Model

We begin with a signaling version of our model. There are two actors in the model, a sender and a receiver. The sender first privately observes the outcome of a random variable θ (which we will refer to as the sender's type) affecting her utility. We assume that $\theta \in \Theta$, where Θ is a compact subset of $\mathbb{R}^l$. Thus, we explicitly allow for multidimensional private information. We assume that θ is distributed according to a commonly known probability distribution function $F(.)$ which possesses a continuous and strictly positive density $f(.)$.

After observing her type θ, the sender sends a vector of n different signals $\mathbf{m}^n = (m_1, ..., m_n)$ to the uninformed receiver. As emphasized in the introduction, each signal is characterized by some specific content, or is sent along a different dimension. We assume that the i-th message m_i belongs to a compact subset $M_i \subseteq \mathbb{R}^l$. Note that we require M_i and Θ to be of the same dimension, so that each message can contain information about all of the aspects or dimensions of the sender's type.[3]

After observing the message vector $\mathbf{m}^n$, the receiver forms posterior beliefs $\mu(\mathbf{m})$ and responds by selecting an action $x \in X$, which affects the utility of both actors. We assume that X is a compact subset of a k-dimensional Euclidean space. The sender's utility function is given by

$$u^n(x, \mathbf{m}^n, \theta)$$

We assume that for each $(\theta, x) \in (\Theta, X)$ there is a message profile $\gamma^n(\theta)$ that uniquely maximizes $u^n(x, \mathbf{m}^n, \theta)$ i.e.,

$$u^n(x, \gamma^n(\theta), \theta) \geq u^n(x, \mathbf{m}^n, \theta),$$

with strict inequality whenever $\mathbf{m}^n \neq \gamma^n(\theta)$. Since sending any message profile other than $\gamma^n(\theta)$ generates strictly less utility for type θ, we will henceforth refer to $\gamma^n(\theta)$ as type θ's costless message profile. The costless message profile $\gamma^n(\theta)$ can be regarded as "truthful" for the agent of type θ, because this is the message profile that the sender would prefer to send if she did not care about affecting the receiver's beliefs and action. In contrast, sending any message $m_i \neq \gamma_i(\theta)$ involves costly and socially inefficient type misrepresentation.

Furthermore, we assume that the utility functions u^n and u^{n+1} for environments with n and $n+1$ messages, respectively, are linked as follows:

$$u^{n+1}(x, m_1, ..., m_n, m_{n+1}, \theta) \leq u^n(x, m_1, ..., m_n, \theta), \text{ for all } (m_1, ..., m_n),\ m_{n+1} \text{ and } \theta$$
$$u^{n+1}(x, m_1, ..., m_n, \gamma_{n+1}(\theta), \theta) = u^n(m_1, ..., m_n, \theta), \text{ for all } (m_1, ..., m_n) \text{ and } \theta.$$

Thus sending more messages is costly for the sender, but the extra cost can always be avoided by sending the truthful message profile $\gamma^n(\theta)$. We denote by $u(x, \theta)$ the sender' utility when she sends the costless message profile $\gamma^n(\theta)$. Thus:

$$u(x, \theta) = u^0(x, \theta) = u^n(x, \gamma^n(\theta), \theta) \text{ for all } n.$$

We assume the costless messages vary with type i.e., there exists a strictly positive constant L such that

$$||\gamma_i(\theta) - \gamma_i(\theta')|| \ \geq L\ ||\theta - \theta'||, \text{ for all } i.$$

Also, there exists $\alpha > 0$ such that

$$u^n(x, \gamma^n(\theta), \theta) - u^n(x, \mathbf{m}^n, \theta) > \alpha\ ||\mathbf{m}^n - \gamma^n(\theta)||^2.$$

The latter assumption says that message profiles further away from the message profile $\gamma^n(\theta)$ are more costly for the sender, and puts a lower bound on these costs.

[3] This involves little loss of generality, as we allow the sender-agent to send multiple messages. For example, if each message could only reflect one dimension of her type, we could always bundle them in groups of l messages.

The receiver's utility function is denoted by $v(x, \theta)$, which we assume to be strictly concave in θ. So for all posterior beliefs $\mu \in \mathcal{P}(\Theta)$, the receiver's best response $BR(\mu) \in X$ is unique, where

$$BR(\mu) = \arg\max_{x \in X} \int v(x, \theta) d\mu(\theta).$$

We use the notation $x^*(\theta)$ to denote the receiver's best response to point beliefs $\delta(\theta)$ at θ i.e.,

$$x^*(\theta) = \arg\max_{x} v(x, \theta)$$

Finally, as standard in many signalling models, we assume that there exists the "best" type $\theta^+ \in \Theta$ such that all sender types wish to be perceived as θ^+. Formally, we have:

$$u^n(x^*(\theta^+), \mathbf{m}^n, \theta) \geq u^n(BR(\mu), \mathbf{m}^n, \theta), \text{ for all } \mu, \mathbf{m}^n, \theta \text{ and } n,$$

with strict inequality whenever $\mu \neq \delta(\theta^+)$.

We focus on pure strategy sequential equilibria of our signalling model. Such an equilibrium is a triple $(\widetilde{\mathbf{m}}^n(\theta), \mu(\mathbf{m}^n), \widetilde{x}^n(\mu))$ where $\widetilde{\mathbf{m}}^n(\theta)$ denotes the equilibrium message strategy of type θ, $\mu(\mathbf{m}^n)$ denotes the beliefs of the receiver after she receives message $\mathbf{m}^n$, and $\widetilde{x}^n(\mu)$ denote the (unique) best response action of the receiver to the beliefs μ, where the beliefs satisfy the standard consistency condition with the sender's equilibrium message strategy, and the sender's message strategy maximizes her expected payoff given the receiver's beliefs.[4] Let $U^n(\theta)$ denote the equilibrium expected utility of type θ. Then we have:

$$U^n(\theta) = u^n(\widetilde{x}^n(\mu(\widetilde{\mathbf{m}}^n(\theta)), \widetilde{\mathbf{m}}^n(\theta), \theta).$$

With a slight abuse of notation, where is does not cause an ambiguity, we will let $\widetilde{x}^n(\theta)$ denote the sender's best response to type θ's equilibrium message profile $\widetilde{\mathbf{m}}^n(\theta)$.

We confine attention to sequential equilibria that survive a very weak refinement termed the dominance criterion (Cho and Kreps, 1987, p. 199). Specifically, fixing some sequential equilibrium of our signalling game let $\mathbf{m}$ be an out-of-equilibrium message profile, and suppose that for type θ sending some message profile $\mathbf{m}'$ dominates sending message profile $\mathbf{m}$ in the following sense:

$$\min_{\mu} u^n(BR(\mu), \mathbf{m}', \theta) > \max_{\mu} u^n(BR(\mu), \mathbf{m}, \theta). \quad (1)$$

Let $J(\mathbf{m})$ denote the set of all such types θ. That is, for every $\theta \in J(\mathbf{m})$ there is a message profile $\mathbf{m}'$ (that may be different across types in $J(\mathbf{m})$) that dominates message $\mathbf{m}$. Then the dominance criterion requires that whenever $J(\mathbf{m})$ is a strict subset of Θ, our equilibrium must survive when the beliefs following the out-of-equilibrium message profile $\mathbf{m}$ are supported on the set $\Theta \setminus J(\mathbf{m})$. Observe that if $J(\mathbf{m}) = \Theta$, then any belief $\mu(\mathbf{m})$ will support our sequential equilibrium. In this case, we may choose $\mu(\mathbf{m}) = \delta(\theta^+)$.

[4]Our restriction to pure strategy sequential equilibria is imposed mainly for brevity. Allowing for mixed strategy equilibria is straightforward, but requires somewhat more cumbersome notation.

3 Signaling with Multiple Messages

It is tempting to argue that our setup trivially implies our results: when the number of messages is large enough, misrepresentation can be easily prevented as the message costs associated with a misrepresentation will be sufficiently large and exceed the benefit thereof. There are two reasons why this logic is fallacious.

First, this reasoning does not imply that under these circumstances the equilibrium signaling cost is small. Indeed, it may very well be that the message costs necessary to prevent misrepresentation are invariant to the dimension of the signal space. This possibility is most easily recognized in Spence's education model, where regardless of the shape of the cost function describing how costly it is to acquire each level of education, the equilibrium signal cost difference must always be the same in order to induce workers to separate. This is illustrated in the following example:

Example 1 *Suppose a worker of productivity θ sends a vector of n signals $(m_1, ..., m_n)$, perhaps reflecting different grades obtained in a variety of different courses, and then receives a wage w determined in a competitive market. The worker's utility is given by $u(w, m_1, ..., m_n, \theta) = w - \sum_{i=1}^{n} c(m_i, \theta)$, where $c(m_i, \theta) = m_i/\theta$. Then since $\sum_{i=1}^{n} c(m_i, \theta) = c(\sum_{i=1}^{n} m_i, \theta)$, any symmetric separating equilibrium with n different signals is also a separating equilibrium for the model with a single signal $m = \sum_{i=1}^{n} m_i$, with cost function $c(m, \theta)$. The equilibrium signal is then given by $m(\theta) = \sum_{i=1}^{n} m_i^n(\theta)$.*

To relate this example to our previous discussion, note that the environment considered in the example is equivalent to an alternative one in which the agent sends a single message, but where the cost function is equals $n\frac{m}{\theta}$. Increasing n, and thereby raising the associated message cost, will induce the agent to send a less costly message, but will result in the same total message cost. What is key in this example is not so much the linearity of the signal cost function, but rather that the costless message is the same across agent types, i.e. $\gamma_i(\theta) = 0$ for all θ.

The basic insight of our paper is that increasing the dimension of the signal space would have an effect if the least cost signal $\gamma_i(\theta)$ varied non-trivially with the agent's type. Indeed, suppose that we fixed the signal cost for all agent types as additional signals are introduced. This would imply that the equilibrium signal profile for any agent type θ gets closer to her truthful or least cost message profile. As our results will ultimately show, the signal profile can be chosen in such a way that, as the number of signals increases, not only will type θ's equilibrium signalling cost decrease, but also the cost of imitating θ will increase for any θ'.

Achieving these two goals simultaneously is non-trivial. As is common, we assume that the marginal cost of misrepresentation is zero at any agent type's costless message.[5] As a consequence, equilibrium requires some misrepresentation. Nevertheless, a judicious choice of message profile ensures that signal costs vanish as the dimension of the signal space increases.

3.1 A Motivating Example

In this section we provide an example highlighting our approach with multiple signals to reexamine the extant results on signalling via education in the job-market, and in particular the seminal contribution of Spence (1973). Consider a worker with privately-known productivity type $\theta \in [0,1]$ whose cost of producing q units of output equals $h(q,\theta) = \frac{q^2}{2}$. The worker's productivity is distributed according to probability distribution $F(.)$. The employer's benefit from q units of output produced by the worker of type θ is equal to $v(q,\theta) = \theta q$. The employer pays the transfer t to the worker. Thus the worker's utility function is given by

$$t - \frac{q^2}{2},$$

and the employer's net benefit is given by

$$\theta q - t$$

As in Spence (1973), we assume that the employment market is competitive, so that the worker gets the full expected surplus $E_\mu(\max_q v(q,\theta) - h(q,\theta))$, where μ describes the employer's beliefs about the worker's type θ. The worker obtains education before entering the job-market. The observable characteristic of education in our model is a profile of grades $(m_1, ..., m_n)$ attained by a worker in the courses she took, with m_i standing for the grade in course i. Education is not productive, but carries a type-dependent communication cost, given by

$$C^n(m_1, ..., m_n, \theta) = \frac{1}{2}\sum_{i=1}^{n}(m_i - \theta)^2$$

Our goal is to characterize a perfect Bayesian equilibrium of this signalling model that satisfies the dominance criterion refinement defined above. For simplicity in exposition, we will focus on separating equilibria. For fixed n, such an equilibrium consists of the equilibrium grade strategy $(\widetilde{m}_1(\theta), ..., \widetilde{m}_n(\theta))$, employer's beliefs $\mu(m)$ mapping grade profiles into probability distributions over types, and a job offer $(\widetilde{q}(\mu), \widetilde{t}(\mu))$ which the worker receives from the employer with beliefs μ.

[5] This assumption follows naturally from differentiability of the cost function u^n, and the fact that u^n is maximized at $m_i = \gamma_i(\theta)$.

In a separating equilibrium, the beliefs on the equilibrium path are degenerate, i.e. $\mu(\widetilde{m}^n(\theta)) = \theta$. We will discuss the off-the-path beliefs supporting this equilibrium below. The competitive markets assumption implies that the quantity allocation in a separating equilibrium must be first-best i.e.,

$$\widetilde{q}(\theta) = \arg\max_q [v(q,\theta) - h(q,\theta)] = \theta$$

and that the worker receive all the surplus from the relation, i.e.

$$t(\theta) = v(\widetilde{q}(\theta), \theta) = \theta^2.$$

Also, because of the symmetry in the grade cost function, all equilibrium signals will be the same, i.e. $\widetilde{m}_i(\theta) = \widetilde{m}(\theta)$ for all $i = 1, ..., n$.

Finally, the following incentive constraints hold for every pair of types (θ, θ'):

$$U^n(\theta) \equiv v(\widetilde{q}(\theta),\theta) - h(\widetilde{q}(\theta),\theta)) - \sum_{i=1,\ldots,n} c(\widetilde{m}_i(\theta),\theta) \geq v(\widetilde{q}(\theta'),\theta') - h(\widetilde{q}(\theta'),\theta) - \sum_{i=1,\ldots,n} c(\widetilde{m}_i(\theta'),\theta),$$

yielding

$$U^n(\theta) \equiv \frac{\theta^2}{2} - \frac{n}{2}(\widetilde{m}(\theta) - \theta)^2 \geq \frac{\theta'^2}{2} - \frac{n}{2}(\widetilde{m}(\theta') - \theta)^2 \tag{2}$$

Thus the right side of (2) is maximized at $\theta' = \theta$, producing the first order conditions:

$$\theta - n(\widetilde{m}(\theta) - \theta)\frac{d\widetilde{m}}{d\theta} = 0$$

This equation admits a solution

$$\widetilde{m}(\theta) - \theta = \alpha\theta, \tag{3}$$

where[6]

$$\alpha = \frac{1}{2}\left(-1 + \sqrt{1 + \frac{4}{n}}\right). \tag{4}$$

It is apparent from (4) that the distortion in each of the worker's equilibrium signals converges to zero as n grows arbitrarily large. Of particular interest are the worker's equilibrium signal cost:

$$\sum_{i=1,\ldots,n} c(\widetilde{m}_i(\theta),\theta) = \frac{n}{2}(\widetilde{m}(\theta) - \theta)^2 = \frac{n}{2}\alpha^2\theta^2 = \frac{\theta^2}{8}\left[\sqrt{n+4} - \sqrt{n}\right]^2 \approx \frac{\theta^2}{2n}$$

Thus equilibrium message costs vanish as the number of signals grows large.

It remains to show that this equilibrium can be supported with off equilibrium beliefs satisfying the dominance criterion. Since the range of equilibrium messages is $[0, 1+\alpha]$, we

[6]There is another solution to (3), with $\alpha = \frac{1}{2}\left(-1 + \sqrt{1 + \frac{4}{n}}\right)$. However, this solution violates the second-order conditions for worker utility maximization.

need only be concerned about messages $m > 1+\alpha$. For any such m, the best possible employer inference for a worker is $\mu(m) = 1$. So by sending an off equilibrium message m, a worker of type θ can get no more than

$$\frac{1}{2} - \frac{n}{2}(m-\theta)^2$$

For any other message m', the worst possible employer inference the worker could face is that $\mu(m') = 0$, in which case he receives a zero transfer and incurs the signal cost $n(m'-\theta)^2$. Thus if any signal m' dominates m for type θ, then $m' = \theta$ also dominates m. Hence m is dominated for type θ whenever

$$1 - n(m-\theta)^2 \leq 0,$$

or equivalently whenever

$$\theta \notin \left[m - \sqrt{\frac{1}{n}}, m + \sqrt{\frac{1}{n}}\right]$$

Since $m \geq 1+\alpha$, and so $m + \sqrt{\frac{1}{n}} > 1$, it follows that we must have $supp(\mu(m)) \subseteq [m - \sqrt{\frac{1}{n}}, 1]$, for $m \leq 1 + \sqrt{\frac{1}{n}}$, and $supp(\mu(m)) = \{1\}$ for $m > 1 + \sqrt{\frac{1}{n}}$. For example, we may take $\mu(m)$ to be the point mass at $\theta = 1$ for all $m \geq 1+\alpha$. It is immediate that these beliefs support our equilibrium.[7]

3.2 Main Results for the Signalling Model

In our more general proof, we do not require that the equilibrium be separating. Indeed, our refinement is too weak to select the signaling equilibrium when n is small. We also allow for arbitrary dimension of the type space, for which few results are known in the extant literature.[8]

[7]To be complete, we should also specify beliefs for message profiles that are not on the diagonal, i.e. for which $m_i \neq m_j$ for some $i \neq j$. Such message profiles m are not dominated for type θ if and only if there exists no message profile m' for which

$$1 - \sum_{i=1}^{n}(m_i - \theta)^2 \geq 0 - \sum_{i=1}^{n}(m_i' - \theta)^2$$

Since the right side of this inequality is maximized at $m_i' = \theta$ for all i, the dominance criterion requires that beliefs following the message profile m be concentrated on the complement of the set $J = \{\theta : \theta \in B(m,1)\}$, whenever it is nonempty. The receiver's beliefs $\mu(m)$ may therefore be specified as a pointmass at

$$\theta = \overline{m} - \frac{1}{n}\sqrt{n + \left(\sum m_i\right)^2 - n\sum m_i^2}$$

whenever $J \neq \phi$, and a point mass at $\theta = 1$, otherwise. It is tedious, though straightforward to check that these beliefs support the equilibrium path.

[8]One notable exception is Lee, Mueller and Vermeulen (2011), who establish existence of a signalling equilibrium in a model with multidimensional types. However, they make use of a strong homogeneity assumption, which effectively reduces the problem to one with single-dimensional type space.

Additionally, our signal space is multidimensional, and we allow the agent to either experience economies or diseconomies in signal cost as the number of available signals increases. Finally, we do not impose any single crossing conditions on the sender's utility function.

Remarkably, we are able to establish that sequential equilibria that satisfy our weak refinement criterion must become arbitrarily close to perfectly revealing as the dimension of the signal space grows without bound, at a signal cost which then becomes arbitrarily small. The intuition for our general result is that for large n, for any type $\theta' \neq \theta$ sending the message profile $\gamma^n(\theta)$ becomes arbitrarily costly relative to sending $\gamma^n(\theta')$. The dominance criterion then guarantees that following the signal profile $\gamma^n(\theta)$, the receiver's beliefs must be concentrated on sender types close to type θ. Thus when n is large, by sending $\gamma^n(\theta)$ type θ can essentially guarantee herself the perfectly revealing payoff, without incurring any message cost.

We now state the main theorem of this section:

Theorem 1 *Consider a sequence of sequential equilibria satisfying the dominance criterion. Then as the number of signals n grows without bounds, equilibrium beliefs following the equilibrium message profile of any type θ converge weakly to point beliefs at θ. So the equilibrium becomes perfectly revealing i.e., $\widetilde{x}^n(\theta) \to x^*(\theta)$ for all θ. Furthermore, the associated equilibrium utility $U^n(\theta) = u^n(\widetilde{x}^n(\theta), \widetilde{\mathbf{m}}^n(\theta), \theta)$ of sender type θ converges to $u(x^*(\theta), \theta)$, for all θ.*

According to Theorem 1, the cost of signalling becomes small as the number of signals n increases. In the job-market signalling application of our model, this implies that each worker-type sends only "almost" costless messages, i.e. gets her "natural" grades" in equilibrium. But the cost to type θ' of imitating type θ becomes large as n grows.

4 Screening with Multiple Messages

In this Section, we recast the model of the previous section in a screening context. Specifically, we consider a principal-agent problem in which the principal controls an allocation $x \in X \subset \mathbb{R}^k$ which contains a vector of production and/or consumption decisions, as well as monetary transfers. The agent privately observes her type $\theta \in \Theta \subset \mathbb{R}^l$. When the agent's type is θ, she obtains utility $u(x, \theta)$ and the principal obtains utility $w(x)$ from allocation x.

To this standard screening environment we add a costly communication process in which the agent sends several messages or signals to the principal.[9] When sending a vector of

[9]We use the term 'signal' both in the mechanism design and in the signalling settings. The term 'signal' is appropriate in the mechanism design and screening settings here because it refers to certain agent's actions or messages which possess only informational value that has to be inferred by the principal.

messages or signals $\mathbf{m}^n = (m_1, ..., m_n) \in M^n = \prod_{i=1}^n M_i$ an agent of type θ incurs the cost $C^n(\mathbf{m}^n, \theta)$. The agent's overall payoff is thus given by:

$$u(x, \theta) - C^n(m_1, ..., m_n, \theta) \tag{5}$$

The cost function C^n is such that for each i message m_i is costless for type θ if and only if $m_i = \gamma_i(\theta)$. Thus we have $\gamma_i(\theta) = \arg\min_{m_i} C^n(m_1, ..., m_i, ..., m_n, \theta)$ for all $(m_1, .., m_{i-1}, m_{i+1}, ..., m_n)$, and $C^n(m_1, ..., m_n, \theta) = 0$ if and only if $m_i = \gamma_i(\theta)$ for all $i = 1, ..., n$.

The agent's cost of misrepresenting her type, i.e. sending message $m_i \neq \gamma_i(\theta)$, generally depends on the other communicated messages. Such dependencies arise naturally. For example, as the agent proceeds with sending more messages to the principal, she may learn how to misrepresent her information more effectively and at a lower cost. Alternatively, additional effort spent on one test could be fatiguing the agent, and therefore raise the cost of undergoing a subsequent test.

Apart from the costly messages $\mathbf{m}^n$, we also allow the agent to send a cheap talk message τ. By the Revelation Principle, we can without loss of generality take the latter to be a type announcement, so that $\tau \in \Theta$ and the agent's message space is equal to $\Theta \times \prod_{i=1}^n M_i \equiv \Theta \times M^n$. In the sequel, we rely on mechanisms that do not use cheap talk messages. However, it is important to show that our results are robust to the addition of such messages and that the latter do not affect the scope of implementation.

A mechanism that the principal offers to the agent can be represented by a mapping $g(.)$ from the agent's message space $\Theta \times M^n$ into the space of allocations X. We will say that an allocation profile $x(\cdot)$ is implementable if there is a mechanism $g(\cdot) : \Theta \times M^n \mapsto X$, an agent's message/signal rule (strategy) $\mu^n : \Theta \to M^n$ and agent's cheap-talk rule (strategy) $\tau : \Theta \mapsto \Theta$ such that for all $\theta \in \Theta$ we have $x(\theta) = g(\tau(\theta), \mu^n(\theta))$ and:

$$u(x(\theta), \theta) - C^n(\mu^n(\theta), \theta) \geq \max_{\theta' \in \Theta, \mathbf{m}^n \in M^n} u(g(\theta', \mathbf{m}^n), \theta) - C^n(\mathbf{m}^n, \theta) \tag{6}$$

When the incentive constraints (6) hold, it is optimal for the agent of type $\theta \in \Theta$ to send an array of messages $(\tau(\theta), \mu^n(\theta))$ in the mechanism $g(.)$, resulting in the allocation $x(\theta)$. In the next section we study how the set of implementable allocations varies with the dimension of the signal space n.

Since the primary focus of this section is on implementability, we do not generally impose individual rationality. However, we point out below how certain results of this section can be interpreted as full surplus extraction by the principal leaving the agent at her outside option level.

4.1 An Example

A principal contracts with an agent to supply output $q \in \mathbb{R}_+$, for which he compensates the agent with a transfer $t \in \mathbb{R}_+$. Output q generates a profit or surplus $S(q) = q - \frac{q^2}{2}$ for the principal. After the payment t to the agent, the principal obtains the net payoff:

$$S(q) - t = q - \frac{q^2}{2} - t.$$

The agent's constant marginal cost of production θ is randomly drawn according to the distribution function $F(\theta)$ with support $\Theta = [0,1]$, and is the agent's private information. The agent can submit a set of claims $(m_1, ..., m_n) \in \Theta^n$ to the principal regarding her marginal cost of production. The agent can manipulate and submit false claims, but in doing so incurs a cost that is increasing in the distance between the message and the true marginal cost. Specifically, the agent's cost of submitting a vector of claims $(m_1, ..., m_n)$ is given by

$$C^n(m_1, ..., m_n, \theta) = \frac{1}{2}\sum_{i=1}^{n}(m_i - \theta)^2$$

Thus after sending a vector of claims $(m_1, ..., m_n)$, producing output q and getting a payment t from the principal, the agent obtains net utility equal to:

$$t - \theta q - C^n(m_1, ..., m_n, \theta) = t - \theta q - \frac{1}{2}\sum_{i=1}^{n}(m_i - \theta)^2 \tag{7}$$

The production contract which the principal offers to the agent specifies a quantity/transfer pair (q, t) as a function of the vector of claims $(m_1, ..., m_n)$ submitted by the agent. So, after signing the contract, the agent sends a vector of n claims regarding her marginal cost to the principal and the quantity/transfer pair $(q(m_1, ..., m_n), t(m_1, ..., m_n))$ is then implemented. The agent's reservation utility is zero, so her net payoff (7) should be at least zero.

Consider the principal's first-best allocation which she would implement if she knew the agent's marginal cost. The output in this allocation maximizes the social surplus

$$q - \frac{q^2}{2} - \theta q$$

and hence is given by $q^{FB}(\theta) = 1 - \theta$. The principal would extract all the surplus from the relation by ordering the quantity $q^{FB}(\theta)$ and paying to the agent a transfer equal to:

$$t^{FB}(\theta) = \theta.$$

When the agent does not incur a misrepresentation costs, i.e. $C^n(m_1, ..., m_n, \theta) \equiv 0$, the outcome $(q^{FB}(\cdot), t^{FB}(\cdot))$ is not implementable, as the agent would overstate her true marginal cost. In the presence of misrepresentation costs, $(q^{FB}(\cdot), t^{FB}(\cdot))$ is not costlessly implementable. This is because the agent's marginal cost of misrepresentation is equal to zero at a truthful claim i.e., $\frac{\partial C^n(m_1,...,m_n,\theta)}{\partial m_i}|_{m_i=\theta} = (m_i - \theta)|_{m_i=\theta} = 0$, so the agent then still has an incentive to overstate her true marginal cost.

It follows that the optimal mechanism will involve some misrepresentation on the part of the agent. This is costly for the principal, for he then has to compensate the agent for her costs of misrepresentation to induce the desired allocation. As a consequence, the principal will have to raise the transfer above $t^{FB}(\theta)$.

Below, we illustrate how the principal can choose the level of misrepresentation $\mathbf{m}^n(\cdot)$ to implement the outcome $(q^{FB}(\cdot), t^n(\cdot), \mathbf{m}^n(\cdot))$ such that, when the number of claims n becomes large, the misrepresentation cost $C^n(\mathbf{m}^n(\theta), \theta)$ becomes arbitrarily small and at the same time the transfer $t^n(\theta)$ exceeds $t^{FB}(\theta)$ by an arbitrarily small amount for all $\theta \in [0, 1]$.

Let us select the transfer $t^n(\cdot)$ and the claim profile $\mathbf{m}^n(\cdot)$ so that each agent-type's net utility is equal to zero, her reservation value. This condition and incentive compatibility can be combined as follows:

$$0 = U(\theta) = t^n(\theta) - \theta q^{FB}(\theta) - \frac{1}{2}\sum_{i=1}^{n}(m_i(\theta) - \theta)^2 = \max_{\theta'}\left\{t^n(\theta') - \theta q^{FB}(\theta') - \frac{1}{2}\sum_{i=1}^{n}(m_i(\theta') - \theta)^2\right\} \tag{8}$$

Applying the Envelope Theorem to (8) yields:

$$0 = U'(\theta) = -q^{FB}(\theta) + \sum_{i=1}^{n}(m_i(\theta) - \theta) \tag{9}$$

Let us choose the same level of misrepresentation in every claim i.e., $m_i(\theta) - \theta = z$ for all i. Using this and $q^{FB}(\theta) = 1 - \theta$ in (9) yields:

$$m_i(\theta) = \theta + \frac{1 - \theta}{n}$$

Thus as n becomes large, the equilibrium claims converge to the truth θ, uniformly in θ. Furthermore, the associated communication costs equal

$$C^n(m_1(\theta), ..., m_n(\theta), \theta) = \frac{1}{2}\sum_{i=1}^{n}(m_i(\theta) - \theta)^2 = \frac{n}{2}(m_i(\theta) - \theta)^2 = \frac{(1 - \theta)^2}{2n},$$

which converges to zero as n becomes large.

Under our chosen claims profile, the transfer function $t^n(\theta)$ satisfying $U^n(\theta) = 0$ is:

$$t^n(\theta) = \theta q^{FB}(\theta) + \frac{1}{2}\sum_{i=1}^{n}(m_i(\theta) - \theta)^2 = (1 - \theta)\left(\theta + \frac{1 - \theta}{2n}\right)$$

To see that this contract satisfies the incentive constraints built in (8), note that under this transfer function the maximand in the max operator in (8) is given by

$$(1-\theta')\left(\theta' + \frac{1-\theta'}{2n}\right) - \theta(1-\theta') - \frac{n}{2}\left(\theta' + \frac{1-\theta'}{n} - \theta\right)^2$$

This maximand is concave in θ' in this simple example. So $(q^{FB}(\cdot), t^n(\cdot), \mathbf{m}^n(\cdot))$ is implementable if the derivative of the maximand is equal to zero at $\theta' = \theta$. It is easy to check that this condition is indeed satisfied. In our more general proof, such concavity is not guaranteed. Instead, we construct a message profile $\mathbf{m}^n(.)$ so that it has a small cost for each agent-type, We show that large deviations in the message strategy are not profitable for any agent-type, as they are simply too costly. To rule out small deviations from $\mathbf{m}^n(.)$, we show that the payoff of agent-type θ as a function of her message profile $\mathbf{m}^n(\theta')$, has a strictly positive directional derivative in the direction of $(\theta - \theta')$.

Our general result considers multidimensional types, permits general quasilinear payoff functions for the agent and the principal, and allows the agent to have either an economy or diseconomy in communication cost as the number of signals increases. In other words, it allows for lying to become either easier or harder as the number of different claims, signals, messages, or pieces of evidence increases.

4.2 Main Result for the Screening Model

To exhibit our key result in the simplest way, we maintain transferable utility model in this section. Specifically, we partition the outcome $x = (q, t)$ into a production assignment $q \in Q$, where Q is a compact subset of $\mathbb{R}^{k-1}_+$, and a transfer $t \in \mathbb{R}$ from the agent to the principal. The agent's utility function is quasilinear and is given by

$$u(x, \theta) = t - h(q, \theta).$$

We assume that $h(q, \theta)$ and $C^n(\mathbf{m}^n, \theta)$ are twice continuously differentiable functions, and that $\gamma^n(\theta)$ is a twice continuously differentiable vector function. The function $h(q, \theta)$ is naturally interpreted as the agent's cost of production, or, alternatively, as her utility from consumption.

Since the focus of this subsection is on implementability, the principal's utility function, defined as $w(x)$ at the beginning of this section, plays no role here.

For any function $g(\rho, \eta) : \mathbb{R}^j \times \mathbb{R}^k \to \mathbb{R}$, where $j, k > 1$, we let $D^2_{\rho\rho} g(\rho, \eta)$ denote the $j \times j$ matrix of second derivatives of g w.r.t. the vector ρ. Similarly, $D^2_{\rho\eta} g(\rho, \eta)$ is $j \times k$ matrix of cross partial derivatives $\frac{\partial^2 g}{\partial \rho_i \partial \eta_l}$. Also, for any function $d(\eta) : \mathbb{R}^k \to \mathbb{R}^k$, $D_\eta d(\eta)$ denotes a $k \times k$ matrix of first derivatives of $d(\eta)$. We may now state the following Assumption:

Assumption 1 *There exist constants* $\alpha \in [0, \frac{1}{2})$, $\omega_1, \omega_2, \omega_3, \omega_4 > 0$, *and* $L < \infty$ *such that:*
(i) For every $\varepsilon > 0$, *there exists* $\delta > 0$ *such that* $|| \sum_{i=1}^{n} D^2_{\theta m_i} C^n(\mathbf{m}^n, \theta) - \sum_{i=1}^{n} D^2_{\theta m_i} C^n(\gamma^n(\theta), \theta)|| < n^{1-\alpha}\varepsilon$ *whenever* $||\mathbf{m}^n - \gamma^n(\theta)|| < \delta$;
(ii) $|| \sum_{i=1}^{n} D^2_{\theta m_i} C^n(\gamma^n(\theta), \theta) z|| \geq \omega_1 n^{1-\alpha} ||z||$;
(iii) For every $\varepsilon > 0$, *there exists* $\delta > 0$ *such that* $||D_{\theta\theta} C^n(\mathbf{m}^n, \theta') - D_{\theta\theta} C^n(\gamma^n(\theta), \theta)|| < \varepsilon n^{1-\alpha}$ *whenever* $||(\mathbf{m}^n, \theta') - (\gamma^n(\theta), \theta)|| < \delta$;
(iv) $x' D_{\theta\theta} C^n(\gamma^n(\theta), \theta) x \geq \omega_2 n^{1-\alpha} \ ||x||^2$;
(v) $\frac{\omega_3}{n^\alpha} ||\mathbf{m}^n - \gamma^n(\theta)||^2 \leq C^n(\mathbf{m}^n, \theta) \leq \frac{\omega_4}{n^\alpha} ||\mathbf{m}^n - \gamma^n(\theta)||^2$;
(vi) $||\gamma_i(\theta) - \gamma_i(\theta')|| \geq L \ ||\theta - \theta'||$, *for all* i.

The requirements embedded in Assumption 1 are rather mild, and mostly technical in nature. They are easily satisfied in most common specifications of the cost function. For example, the cost function

$$C^n(\mathbf{m}^n, \theta) = \frac{1}{2} \sum_{i=1}^{n} (m_i - \theta)^2$$

satisfies all the requirements.

The only premise of economic significance is Assumption 1(vi), which requires the costless message profile $\gamma^n(\theta)$ to vary non-trivially with θ. Assumption 1(v) is rather innocuous, as it only requires $C^n(\cdot, \theta)$ to be majorized and minorized by quadratic functions centered around $\gamma^n(\theta)$.

Assumption 1(ii) requires $D_{\theta m_i}$ to vary non-trivially with θ at $\mathbf{m}^n = \gamma^n(\theta)$, thereby slightly strengthening Assumptions 1(v) and 1(vi).[10] Assumptions 1(i) and (iii) impose mild continuity requirements on $D^2_{\theta mi} C^n(\mathbf{m}^n, \theta)$ and $D^2_{\theta\theta}$in $\mathbf{m}^n C^n(\mathbf{m}^n, \theta)$ in the neighborhood of truthtelling, and control the asymptotic behaviors of these terms. Assumption 1(iv) requires $C^n(\mathbf{m}^n, \theta)$ to be α-convex in θ, with a parameter $\omega_2 n^{1-\alpha}$ that grows along with n. Local convexity of $C^n(\mathbf{m}^n, \theta)$ in θ at $\mathbf{m}^n = \gamma^n(\theta)$ is already implied by our basic assumptions on C^n, provided that γ^n depends non-trivially on θ.[11] At some cost in complexity of the proof, global convexity in θ can be dispensed with, since large deviations are easily deterred when the number of required

[10] To see this, note that since $C^n(m^n, \theta)$ is globally minimized at $m^n = \gamma^n(\theta)$, we have $D_{m_i} C^n(\gamma^n(\theta), \theta) = 0$. Totally differentiating this identity w.r.t. θ yields $D^2_{m_i \theta} C^n(\gamma^n(\theta), \theta) = D^2_{m_i mi} C^n(\gamma^n(\theta), \theta) D_\theta \gamma^n(\theta)$. Assumption 1(5) implies that $D^2_{m_i mi} C^n(\gamma^n(\theta), \theta)$ is a positive definite matrix. A mild strengthening of 1(vi) would have $D_\theta \gamma^n(\theta)$ be nonsingular. This implies that $||D^2_{\theta m_i} C^n(\gamma^n(\theta), \theta) z|| > 0$.

[11] To see this, note that since $C^n(m^n, \theta)$ is globally minimized at $m^n = \gamma^n(\theta)$, the identities $D_m C^n(\gamma^n(\theta), \theta) = 0$ and $D_\theta(\gamma^n(\theta), \theta) = 0$ hold. Totally differentiating w.r.t. θ the following expressions hold at $(\gamma^n(\theta), \theta)$: $D^2_{mm} C^n D_\theta \gamma^n + D^2_{m\theta} = 0$ and $D^2_{\theta m} C^n D_\theta \gamma^n + D^2_{\theta\theta} C^n = 0$. Thus $D^2_{\theta\theta} C^n = - D^2_{\theta m} C^n D_\theta \gamma^n = (D_\theta \gamma^n)' D^2_{mm} C^n D_\theta \gamma^n$. Since $m^n = \gamma^n(\theta)$ minimizes $C^n(m^n, \theta)$, $D^2_{mm} C^n$ is a positive semidefinite matrix, implying that $D^2_{\theta\theta} C^n$ is positive semidefinite. Slightly strengthening this to positive definitiveness of $D^2_{mm} C^n$ then implies that $D^2_{\theta\theta} C^n$ is positive definite because by Assumption 1(vi) we have $D_\theta \gamma^n \neq 0$.

messages is large. Hence the essential assumption embodied in 1(iii) is that the degree of convexity of C^n in θ becomes increases with the number of messages. This condition in turn follows naturally from a lower bound on the degree of convexity of C^n in each element of the message profile, and a lower bound on the rate at which costless messages vary with type.[12]

We may now state:

Theorem 2 *Suppose Assumption 1 holds. Then for every pair of twice continuously differentiable functions (q, t) and every $\varepsilon > 0$, there exist $N < \infty$ such that whenever the dimension of the message space n exceeds N, we can implement a pair (q, t^n) where the transfer function t^n satisfies $|t^n(\theta) - t(\theta)| < \varepsilon$. Furthermore, the associated communication cost $C^n(\mathbf{m}^n(\theta), \theta)$ is less than ε for every agent-type θ.*

Note that we have not imposed agent individual rationality on our implementation problem. However, an immediate implication of Theorem 2 is that the principal can extract the surplus from the agent, leaving the latter at her outside option value, and implement the first-best quantity allocation $q^{FB}(.)$ with negligible communication costs.

The proof of Theorem 2 is rather subtle. An important part of the mechanism design in Theorem 2 is the construction of an incentive compatible message profile $\mathbf{m}^n(\theta)$, $\mathbf{m}^n \neq \gamma^n(\theta)$, which converges to the profile of costless messages $\gamma^n(\theta)$ as n increases.

The transfer rule t^n is chosen to provide the agent the same net utility as in the allocation (q, t) that we wish to implement. Particularly, the first-order condition for local incentive compatibility is:

$$D_\theta t^n(\theta) = D_q h(q(\theta), \theta) D_\theta q(\theta) + D_{\mathbf{m}^n} C^n(\mathbf{m}^n(\theta), \theta) D_\theta \mathbf{m}^n(\theta) \tag{10}$$

If the agent was to send only costless messages, then the second term in (10) would be zero, as $D_{m_i} C^n(\gamma^n(\theta), \theta) = 0$. Thus in the absence of costly messages the first-order condition imposes a restriction on the set of implementable allocations in the form of a link between $q(\theta)$ and $t^n(\theta)$. Costly signals weaken and eventually eliminate the need for such link, and allows us to implement a larger set of allocation profiles.

When the number of messages n is small, the degree to which the link between $q(\theta)$ and $t^n(\theta)$ is weakened is limited by the magnitude of the required misrepresentation $\mathbf{m}^n(\theta)$, and

[12] To see this, note that since $m_i = \gamma_i(\theta)$ minimizes $C^n(m_{-i}, m_i, \theta)$, we have $D_{m_i} C^n(m_{-i}, \gamma_i(\theta), \theta) = 0$. It follows that $D^2_{m_i m_j} C^n(\gamma^n(\theta), \theta) = 0$ for all $i \neq j$. From the previous footnote we then know that $D^2_{\theta\theta}(\gamma^n(\theta), \theta) = (D_\theta \gamma(\theta)' D^2_{mm} C^n(\gamma^n(\theta), \theta) D_\theta \gamma(\theta) = \sum_{i=1}^n (D_\theta \gamma_i(\theta))' D^2_{m_i m_i} C^n(\gamma^n(\theta), \theta) D_\theta \gamma_i(\theta)$. Suppose now that for each n and each i, $D^2_{m_i m_i} C^n(\gamma^n(\theta), \theta)$ is a positive definite matrix, with eigenvalues bounded below by $\lambda > 0$. Suppose also that for each i, $(D_\theta \gamma_i(\theta))' D_\theta \gamma_i(\theta)$ is a positive definite matrix, with eigenvalues bounded below by $\sigma > 0$. Then we have $\theta' D^2_{\theta\theta}(\gamma^n(\theta), \theta)\theta = \sum_{i=1}^n \theta' (D_\theta \gamma_i(\theta))' D^2_{m_i m_i} C^n(\gamma^n(\theta), \theta) D_\theta \gamma_i(\theta)\theta \geq \lambda \sum_{i=1}^n \theta' (D_\theta \gamma_i(\theta))' D_\theta \gamma_i(\theta)\theta \geq \lambda\sigma n \, ||\theta||^2$, so the function $C^n(m^n, \theta)$ is α-convex in θ, with parameter $\alpha = \lambda\sigma n$.

the associated communication costs, which may have to be fairly large. In contrast, with sufficiently large n, the codependency between $q(\theta)$ and $t^n(\theta)$ is eliminated at a very small cost.

Assumption 1 (iii) is used to show that when the number of messages is sufficiently large, truthtelling is a strict local maximum. It also guarantees that large deviations from truthtelling are not profitable. Assumption (iv) is then invoked to establish that the aggregate message costs vanish as the number of messages increases.

Finally, consider the second-order conditions for implementation. Maggi and Rodriguez-Clare (1995), who characterize the optimal mechanism for $n = 1$, imposed the restriction $q'(\theta) \geq 0$ and $m'(\theta) \geq 0$ to guarantee that their second-order conditions hold. In contrast, a careful inspection of our proof reveals that, with many n, the second-order conditions hold because the agent sends a large number of messages which are close to her costless message $\gamma(\theta)$.

4.3 Optimal Mechanisms

Theorem 2 provides conditions under which almost all decision rules become implementable and the communication cost converges to zero as the number of messages becomes large. However, it is also interesting to consider what happens when the number of available messages remains limited, either exogenously or endogenously.

In this section, we characterize the optimal mechanism for a one-dimensional type space, i.e. $l = 1$, and every possible number of messages n. The agent's type θ is assumed to be randomly drawn from an interval $[\underline{\theta}, \overline{\theta}]$ according to cdf $F(.)$ with continuous density $f(.)$. We maintain the assumption that an allocation x consists of a monetary part t and non-monetary part q. For simplicity, we assume that $q \in \mathbb{R}_+$.

The principal's and the agent's payoff functions are assumed to be quasilinear and are given by $v(q) - t$ and $t - h(q, \theta)$, respectively. We impose individual rationality, and normalize the agent's reservation utility to 0.

We make standard assumptions that $v(.)$ and $h(.)$ are twice continuously differentiable, $h_\theta \geq 0$, $h_{q\theta} > 0$, $h_{\theta\theta} \geq 0$ and $v_{qq} - h_{qq} < 0$ for all q and $\theta \in [\underline{\theta}, \overline{\theta}]$, and that $v(0) = h(0, \theta) = 0$ for all θ. In addition, we assume that $v_q(0) - h_q(0, \overline{\theta}) = 0$, $v_q(0) - h_q(0, \theta) > 0$ for all $\theta \in [\underline{\theta}, \overline{\theta})$, and that there exists $\bar{q} \in (0, \infty)$ s.t. $v_q(q) - h_q(q, \theta) < 0$ for all $q > \bar{q}$ and all θ. These assumptions guarantee that the solution $q^{FB}(\theta)$ to the following problem:

$$\max_{q \geq 0} \{v(q) - h(q, \theta)\}$$

exists, satisfies $q^{FB}(\theta) > 0$ for all $\theta \in [\underline{\theta}, \overline{\theta})$, $q^{FB}(\overline{\theta}) = 0$, and is decreasing in θ.

Finally, we make the following technical assumptions:

Assumption 2 *(i)* $v(q) - h(q,\theta) + \frac{F(\theta)}{f(\theta)} h(q,\theta)$ *is has strictly decreasing differences in* (q,θ)*;*

(ii) $C^n(\mathbf{m}^n,\theta) + \frac{F(\theta)}{f(\theta)} C_\theta(\mathbf{m}^n,\theta)$ *is submodular in* $\mathbf{m}^n$*, and has strictly decreasing differences in* $(\mathbf{m}^n,\theta)$*;*

(iii) $C^n(\mathbf{m}^n,\theta)$ *is convex in* θ*, submodular in* $\mathbf{m}^n$*, and has strictly decreasing differences in* $(\mathbf{m}^n,\theta)$*.*

Parts (i) and (ii) of Assumption 2 require the cross-partial derivatives of the agent's virtual utility and virtual communication cost to be negative. It is well-known that Assumption 2 (i) holds if, in addition to the above assumptions on v and h, the cdf $F(.)$ possesses the monotone hazard rate property. Part (iii) requires the cost function to be convex in $\mathbf{m}^n$ and in θ, and imposes a single crossing condition.

We now proceed to derive the optimal mechanism. Specifically, the principal selects a "quantity" $q(\cdot)$ and transfer function $t(.)$, and a vector of messages $\mathbf{m}^n(\cdot)$ to solve:

$$\max_{q(\theta),t(\theta),\mathbf{m}^n(\theta)} \int_{\underline{\theta}}^{\overline{\theta}} (v(q(\theta)) - t(\theta)) f(\theta) d\theta$$

subject to incentive constraints:

$$t(\theta) - h(q(\theta),\theta) - C^n(\mathbf{m}^n(\theta),\theta) \geq t(\theta') - h(q(\theta'),\theta) - C^n(\mathbf{m}^n(\theta'),\theta), \text{ for all } \theta \text{ and } \theta' \quad (11)$$

and the individual rationality constraint:

$$U(\theta) \equiv t(\theta) - h(q(\theta),\theta) - C^n(\mathbf{m}^n(\theta),\theta) \geq 0. \quad (12)$$

Let us first consider the benchmark case with no costly messages, i.e. $n = 0$. It is well-known that under Assumption 2 (i) the solution to the principal's problem involves selecting an allocation $q^{SB}(\theta)$ that maximizes the 'virtual' surplus

$$\Gamma(q,\theta) \equiv v(q) - h(q,\theta) - \frac{F(\theta)}{f(\theta)} h_\theta(q,\theta). \quad (13)$$

Let $\theta^* \in (\underline{\theta}, \bar{\theta})$ be the unique solution to $\Gamma_q(0,\theta) = 0$. Such θ^* exists because $\Gamma_q(0,\overline{\theta}) < 0$, $\Gamma_q(0,\underline{\theta}) > 0$ and by Assumption 2 (i) $\Gamma_q(0,\theta)$ is decreasing in θ. Thus, $q^{SB}(\theta) > 0$ if and only if $\theta \in [\underline{\theta}, \theta^*)$. In other words, it is optimal for the principal to exclude all types in $[\theta^*, \overline{\theta}]$ by assigning them zero quantity, which is socially inefficient because all agent types except $\overline{\theta}$ can generate a positive social surplus. Under a procurement interpretation of our model, this means that higher-cost firms are excluded from production. Under an equivalent reinterpretation of our model as that of a monopolistic seller facing consumers with privately known demands parameterized by θ, the seller excludes buyers with low but positive demands from consumption. Note that it follows from Assumption 2(i) that $q^{SB}(\theta)$ is strictly decreasing in θ over the region $[\underline{\theta}, \theta^*)$.

Next, let us consider the problem for $n > 0$. Solving (12) for $t(\theta)$ and substituting in the objective, and replacing the incentive constraints (11) by the envelope condition associated with the agent's utility maximization, yields the following "relaxed" problem:

$$\max_{q(\theta), \mathbf{m}^n(\theta), U(\theta)} \int_{\underline{\theta}}^{\bar{\theta}} \{v(q(\theta)) - h(q(\theta), \theta) - C^n(\mathbf{m}^n(\theta), \theta) - U(\theta)\} f(\theta) d\theta \tag{14}$$

subject to individual rationality constraint (12) and the envelope condition:

$$U'(\theta) = -h_\theta(q(\theta), \theta) - C^n_\theta(\mathbf{m}^n(\theta), \theta). \tag{15}$$

We will verify that the solution to the relaxed problem (14) subject to (12) and (15) satisfies (11) and hence also solves the unrelaxed problem. To solve the relaxed problem, define the Hamiltonian

$$H = \{v(q) - h(q, \theta) - C^n(\mathbf{m}^n, \theta) - U\} f(\theta) - \sigma \left(h_\theta(q, \theta) + C^n_\theta(\mathbf{m}^n, \theta) \right) + \rho U \tag{16}$$

Maximizing (16) w.r.t. q and $\mathbf{m}^n$ yields the first order conditions:

$$\{v_q(q) - h_q(q, \theta)\} f(\theta) - \sigma h_{q\theta}(q, \theta) \quad \leq \quad 0 \ (= 0, \text{ if } q > 0) \tag{17}$$

$$\frac{\partial C^n}{\partial m_i}(\mathbf{m}^n, \theta) f(\theta) + \sigma \frac{\partial^2 C^n}{\partial m_i \partial \theta}(\mathbf{m}^n, \theta) \quad = \quad 0 \tag{18}$$

The costate equation is

$$\sigma'(\theta) = f(\theta) - \rho(\theta), \tag{19}$$

Furthermore, the solution has to satisfy complementary slackness conditions

$$\rho(\theta) U(\theta) = 0, \ \rho(\theta) \geq 0, \text{ and } U(\theta) \geq 0, \tag{20}$$

Also, the following transversality conditions have to hold: $\sigma(\underline{\theta}) U(\underline{\theta}) = 0$, $\sigma(\bar{\theta}) U(\bar{\theta}) = 0$, $\sigma(\underline{\theta}) \leq 0$ and $\sigma(\bar{\theta}) \geq 0$

To describe the solution to the relaxed problem, let $\{\widehat{q}(\theta), \widehat{\mathbf{m}}^n(\theta), \widehat{\sigma}(\theta)\}$ be the solution to the system consisting of (17), (18) and $-h_\theta(q, \theta) - C^n_\theta(\mathbf{m}^n, \theta) = 0$. The latter condition implies that $U'(\theta) = 0$ which holds on any interval where the individual rationality constraint (12) is binding. So $\{\widehat{q}(\theta), \widehat{\mathbf{m}}^n(\theta), \widehat{\sigma}(\theta)\}$ is the solution to the relaxed problem applicable to this case.

Next, let $\widetilde{q}(\theta)$ and $\widetilde{\mathbf{m}}^n(\theta)$ be the solutions to (17) and (18), respectively, when $\widetilde{\sigma}(\theta) = F(\theta)$. It follows that $\widetilde{\mathbf{m}}^n(\theta) = q^{SB}(\theta)$ where $q^{SB}(\theta)$ is the standard second-best quantity, as noted above. Thus $\{q^{SB}(\theta), \widetilde{\mathbf{m}}^n(\theta), \widetilde{\sigma}(\theta) = F(\theta)\}$ is the solution to the relaxed problem on an interval $[\underline{\theta}, \hat{\theta}]$ where the individual rationality constraint is not binding so that $\rho(\theta) = 0$. Note that the costate equation and the transversality condition at $\underline{\theta}$ require that $\widetilde{\sigma}(\theta) = F(\theta)$ in this case.

Our next Theorem shows that the solution to the full (unrelaxed) program combines these two allocation rules, with $(q^{SB}(\theta), \widetilde{\mathbf{m}}^n(\theta))$ applicable on the lower part of the type space and $(\widehat{q}(\theta), \widehat{\mathbf{m}}^n(\theta))$ applicable on the upper part of the type space.

Theorem 3 *Suppose Assumption (2) holds and recall that* $\theta^* \in (\underline{\theta}, \overline{\theta})$ *is the unique solution to* $\Gamma_q(0, \theta) = 0$ *with* $\Gamma(q, \theta)$ *defined in (13).*

The allocation $(q(\theta), \mathbf{m}^n(\theta))$ *in the optimal mechanism is continuous. Furthermore, there exists* $\widehat{\theta} \in (\underline{\theta}, \theta^*)$ *such that:*

$$(q(\theta), \mathbf{m}^n(\theta)) = \begin{cases} (q^{SB}(\theta), \widetilde{\mathbf{m}}^n(\theta)) & \text{if } \theta \in [\underline{\theta}, \widehat{\theta}), \\ (\widehat{q}(\theta), \widehat{\mathbf{m}}^n(\theta)) & \text{if } \theta \in [\widehat{\theta}, \bar{\theta}]. \end{cases} \tag{21}$$

For every $\theta \in [\underline{\theta}, \widehat{\theta})$*,* $\widetilde{\mathbf{m}}^n(\theta) > \gamma^n(\theta)$*,* $U(\theta) > 0$ *and* $U'(\theta) < 0$*. For every* $\theta \in [\widehat{\theta}, \bar{\theta})$*,* $\widehat{q}(\theta) \in (q^{SB}(\theta), q^{FB}(\theta))$*,* $\widehat{\mathbf{m}}^n(\theta) > \gamma^n(\theta)$*, and* $U(\theta) = 0$*.*

Theorem 3 shows that the principal can exploit the agent's misrepresentation costs in the optimal mechanism to improve the efficiency of the allocation and at the same time to reduce the information rent of the agent. Indeed, positive surplus in our mechanism is obtained by agent-types in $[\underline{\theta}, \widehat{\theta}]$ whereas, as noted in the proof, in the standard second-best mechanism some types larger than $\hat{\theta}$ also get positive surplus. The principal of course incurs some additional cost as she ultimately has to compensate the agent for the latter's misrepresentation costs. However, this is compensated by the extra efficiency of the mechanism. Specifically, all agent types in $[\hat{\theta}, 1]$ consume a positive quantity since there is no exclusion in our mechanism, as demonstrated by Theorem 4, below. Yet, those types earn zero surplus, so the benefit of their consumption goes to the principal net of the misrepresentation costs.

Theorem 3 is related to Proposition 1 in Maggi and Rodriguez-Clare (1995) which characterizes the optimal mechanism with a single costly message. In particular, both in their and our models a non-trivial set of types $[\widehat{\theta}, \overline{\theta}]$ are held at the reservation utility level in our optimal mechanism. So this property is robust to the number of messages. However, there are important differences between our results. First, in our model the agent sends multiple signals, and we focus on exploring how the number of signals affects the optimal mechanism (see Theorem 5 below).

Second, our strategy of proof is different from Maggi and Rodriguez-Clare (1995), and as a consequence we are able to establish our Theorem 3 and Theorems 4 and 5 under more general conditions. In particular, unlike Maggi and Rodriguez-Clare (1995), we do not assume that the cost of a signal m depends only on the difference $(m-\theta)$ and do not require the Hamiltonian to be concave.

Another difference between our results and those of Maggi and Rodriguez-Clare (1995) is the next Theorem which shows that the optimal mechanism exhibits no exclusion when misrepresentation is costly. This result is important, because it underscores that the phenomenon of exclusion, present in the standard case without costly signals, is not robust.

Theorem 4 *In the optimal mechanism, $q(\theta)>0$ for all $\theta\in[\underline{\theta},\overline{\theta})$.*

Thus, according to Theorem 4, no type who can generate a positive surplus is excluded in the optimal mechanism as long as misrepresentation costs are positive, no matter how small.

The intuition for the absence of exclusion with costly signals is as follows. Any agent-type that generates a positive surplus in the first-best is potentially profitable to the principal. The reason some agent types are excluded in the second-best is that if the principal were to give them a positive quantity, he would also have to raise the surplus of all agents with lower θ's (lower costs). But when misrepresentation costs are positive, this is no longer necessary: the principal can prevent imitation by requiring agent-types that now receive a positive quantity to send costly signals. Since lower-cost (lower θ) agent types also incur lower signal costs, their incentives to imitate are weakened and exclusion is avoided.

Finally, we characterize the nature of the solution as the number of costly messages, n, increases.

Theorem 5 *Suppose that $C^n(m_1,...,m_n,\theta)=\sum_{i=1}^n c_i(m_i,\theta)$ and there exist $\underline{v},\bar{v}\in(0,\infty)$ such that $\underline{v}\leq\frac{\partial^2 c}{\partial m_i^2}\leq\bar{v}$ and $\underline{v}\leq|\frac{\partial^2 c}{\partial\theta\partial m_i}|\leq\bar{v}$. Then, as $n\to\infty$, in the optimal mechanism $\widehat{\theta}\to\underline{\theta}$. Furthermore $\widehat{q}(\theta)\to q^{FB}(\theta)$, $\mathbf{m}_i^n(\theta)\to\ \gamma_i(\theta)$ for all i, and $C^n(\mathbf{m}^n(\theta),\theta)\to 0$, uniformly in θ.*

Theorem 5 implies that, as the number of messages gets large, the quantity allocation in the optimal mechanism converges to the first-best one while the communication/signalling cost becomes negligibly small. Thus, the principal extracts almost all surplus. The small size of the communication costs and almost-efficiency of the allocation contrasts with typical inefficiency of the second-best solution under adverse selection highlighted in the literature.

4.4 Endogenous Signal Space

One factor that could affect our results in the presence of additional fixed cost of each message. For example, the principal may have to incur such cost for developing and administering a test, and/or processing the information received form the agent. So in this subsection, we study the optimal number of messages n in a mechanism, when the principal incurs a fixed cost G to elicit each message.

To simplify matters, we will treat n as a continuous variable. Let $W(n)$ denote the principal's expected surplus, gross of any fixed costs, in the optimal mechanism when the dimension of the signal space is n. The principal then selects n to maximize $W(n) - nG$. We assume that the agent's signalling costs are additively separable across messages, i.e.

$$C^n(m_1, m..., m_n, \theta) = \sum_{i=1}^{n} c(m_i, \theta),$$

and that

$$\frac{c_{m\theta}}{c_m}(m, \theta) \text{ is increasing in } m.$$

The latter assumption ensures that the solution to (18) is unique, and hence independent of i. Henceforth, we shall therefore omit the subscript of the message m_i. The optimal number of messages n^* is characterized in the following Lemma.

Lemma 1 *Suppose that $c_{m\theta}(m, \theta)^2 - c_{mm\theta}(m, \theta)c_m(m, \theta) > 0$ for all (m, θ). Then $W(n)$ is a strictly concave function, and the marginal benefit of an additional message is given by*

$$\frac{dW(n)}{dn} = \int_{\underline{\theta}}^{\bar{\theta}} \left(\frac{c_\theta(m(\theta), \theta)c_m(m(\theta), \theta)}{c_{m\theta}(m(\theta), \theta)} - c(m(\theta), \theta) \right) f(\theta)d\theta \tag{22}$$

To help us interpret the solution, note that in the optimal mechanism $\frac{dW(n)}{dn} = G$. Further, let us define:

$$\underline{K} = \min_{(m,\theta)} \frac{c_\theta(m, \theta)c_m(m, \theta)}{c_{m\theta}(m, \theta)c(m, \theta)} - 1, \qquad \overline{K} = \max_{(m,\theta)} \frac{c_\theta(m, \theta)c_m(m, \theta)}{c_{m\theta}(m, \theta)c(m, \theta)} - 1$$

Then we have

Lemma 2 *Suppose that $\overline{K} > 0$. Then, in the optimal mechanism,*

$$\underline{K} \int_{\underline{\theta}}^{\bar{\theta}} c(m(\theta), \theta)f(\theta)d\theta \leq G \leq \overline{K} \int_{\underline{\theta}}^{\bar{\theta}} c(m(\theta), \theta)f(\theta)d\theta$$

The significance of Lemma 2 lies in characterizing the relationship between the principal's total cost of setting up the signalling system, nG, and the agent's communication costs in the optimal mechanism $C^n(m, \theta)$. As the Lemma shows, the ratio of the former to the latter is between $\underline{K}$ and $\overline{K}$. For example, if $c(.)$ is quadratic i.e., $c(m, \theta) = (m - \theta)^2$, then $\underline{K} = \overline{K} = 1$. This implies that in the optimal mechanism the principal's fixed cost of setting up the messages nG and the agent's communication cost $C(m, \theta)$ will be equal.

To illustrate this result, consider the example in which welfare and signalling costs are quadratic, and the type distribution is uniform. Specifically, suppose that $v(q) = q - \frac{1}{2}q^2$, $h(q, \theta) = \theta q$, $c(m, \theta) = \frac{1}{2}(m - \theta)^2$, and $F(\theta) = \theta$ for $\theta \in [0, 1]$. Then by Theorem 3, for fixed n the solution to the principal's problem is given by:

$$\begin{aligned}
\widetilde{q}(\theta) &= 1 - 2\theta;\ \widetilde{m}(\theta) = 2\theta \\
\widehat{q}(\theta) &= \frac{n}{n+1}(1-\theta);\ \widehat{m}(\theta) = \theta + \frac{1-\theta}{n+1};\ \widehat{\sigma}(\theta) = \frac{1-\theta}{n+1} \\
\widehat{\theta}(n) &= \frac{1}{n+2}
\end{aligned}$$

Furthermore, the principal's marginal benefit of an additional message is $\frac{dW(n)}{dn} = \frac{1+(n+1)^2}{6(n+1)^3}$.

Recall that every message generates some amount of additional welfare, because the allocation profile gets closer to the first-best when the agent has to send more messages (see Theorem 5). To illustrate the relation between the fixed cost G and the welfare generated by an extra message, let us express the fixed cost as a fraction of the potential surplus gain $\Delta W = W^{FB} - W^{SB}$, where W^{FB} (W^{SB}) is the total welfare under the first-best (second-best) quantity allocation. Since $W^{FB} = \frac{1}{6}$ and $W^{SB} = \frac{1}{8}$, we have

n	1	2	3	4	5	6	7	8	9	10
$G/\Delta W$	74%	62.5%	54.4%	48.2%	43.2%	39%	35.7%	32.8%	30.4%	28.2%

Thus in this example the principal will elicit at least 10 messages if the fixed cost of eliciting an extra message does not exceed 30% of the potential welfare gain. Note that in the process, the agent will incur expected message costs of at least 30% of the potential welfare gain, thereby dissipating a substantial portion of the benefit. It should also be noted that with four signals, the allocation $q(\theta)$ is already close to the first-best, as $\widehat{q}(\theta)$ is within 9% of $q^{FB}(\theta)$.

5 Conclusions

This paper demonstrates that in environments with misrepresentation costs, the ability of the principal to offer mechanisms in which an agent sends several messages significantly expands the set of implementable outcomes.

Our results have a number interesting implications for screening and signaling. In particular, they suggests that the problem of the dissipation of resources and effort in unproductive signalling, the so-called 'rat race,' may not be as significant as previously thought. Our paper also indicates that an optimal method of dealing with the problem of asymmetric information regarding employees' abilities may involve the design of testing and interviewing procedures, rather than on-the-job screening via incentive schemes. This can explain why incentive schemes offered in a variety of industries are not as steep and high-powered as incentive literature may suggest. Indeed, our paper indicates that an employer can obtain a good estimate of a job-candidate's ability and at a low cost, if the tests and interviews can be designed to have the following properties: (i) Each test identifies a candidate's ability accurately if the candidate does not attempt to manipulate the results of the test by expending effort; (ii) A candidate incurs some cost of effort when (s)he attempts to misrepresent her type.

In our setting, the marginal cost of a message/signal can depend on the content and number of other messages/signals sent by the agent. For example, the amount of effort that an agent of ability θ may need to exert in the n-th test to perform at a level corresponding to ability $\theta' \neq \theta$ may depend on how hard she worked to prepare for other tests and how many other tests she has taken. Our results hold when the effect of the true ability θ on the cost of sending signal $m \neq \gamma(\theta)$ does not go to zero "too quickly' in n. Intuitively, the learning process cannot be too fast so that performing at a certain level in a testing procedure involving $n+1$ tests is only slightly more costly and requires a bit more effort than performing at the same level in a testing procedure consisting of n tests.

It is conceivable that there may exist fixed costs incurred either by the principal or the agent in association with each test or interview. The presence of such costs would limit the feasible number of interviews/tests from above and perfect screening may become too costly. Still, our results indicate that multi-test procedures would dominate the ones relying on one test. Furthermore, it is likely that the fixed costs would be associated with a particular test, and not a particular job-candidate. Then test-specific fixed costs will be amortized over all the job-candidates who undergo it, and therefore would create less of an obstacle for increasing the number of tests. In this case, our model predicts that larger firms who interview many applicants will put more emphasis on rigorous testing and evaluation of candidates before hiring, rather than on providing on-the-job incentives. This appears to be broadly consistent with reality.

6 Appendix

Proof of Theorem 1.

Let $\mu(\gamma^n(\theta))$ denote the receiver's posterior following the message profile $\gamma^n(\theta)$. Let us show that $\mu(\gamma^n(\theta))$ must be supported on a sufficiently small neighborhood of θ. Indeed, we will argue that if θ' lies outside such a neighborhood, then for this type sending the message profile $\gamma^n(\theta)$ is dominated by sending the message profile $\gamma^n(\theta')$.

To this end, let $\overline{u}$ and $\underline{u}$ be the maximal and minimal possible sender's utilities (gross of signaling cost), respectively. Formally,

$$\overline{u} = \max\{u(x,\theta) : \theta \in \Theta \text{ and } x = BR(\mu) \text{ for some } \mu \in \Delta(\Theta)\},$$

and

$$\underline{u} = \min\{u(x,\theta) : \theta \in \Theta \text{ and } x = BR(\mu) \text{ for some } \mu \in \Delta(\Theta)\}.$$

The continuity of the function $u(\cdot,\cdot)$ and the compactness of the sets X and Θ imply that $-\infty < \underline{u} < \overline{u} < \infty$.

Now, suppose that

$$||\theta - \theta'|| > \sqrt{\frac{\overline{u} - \underline{u}}{\alpha n L^2}}. \tag{23}$$

Using (23) we obtain:

$$\begin{aligned}
&\min_{\mu} u^n(BR(\mu), \gamma^n(\theta'), \theta') - u^n(x^*(\theta^+), \gamma^n(\theta), \theta') = \\
&\min_{\mu} u^n(BR(\mu), \gamma^n(\theta'), \theta') - u^n(x^*(\theta^+), \gamma^n(\theta'') + u^n(x^*(\theta^+), \gamma^n(\theta'') - u^n(x^*(\theta^+), \gamma^n(\theta), \theta') \\
&\geq \underline{u} - \overline{u} + \alpha||\gamma^n(\theta') - \gamma^n(\theta)||^2 > \underline{u} - \overline{u} + \alpha n L^2 ||\theta' - \theta||^2 > 0,
\end{aligned} \tag{24}$$

So for any type θ' satisfying (23) sending $\gamma^n(\theta)$ is dominated by sending $\gamma^n(\theta')$. This argument applies when $\gamma^n(\theta)$ is an off-equilibrium message profile as well as when it is on the equilibrium path.

Thus, $supp(\mu(\gamma^n(\theta))) \subset B(\theta, \sqrt{\frac{\overline{u}-\underline{u}}{\alpha n L^2}})$ in any equilibrium satisfying the dominance criterion. It follows immediately that $\mu(\gamma^n(\theta))$ converges weakly to $\delta(\theta)$, the point mass at θ.

Since sender type θ can always choose to send the message profile $\gamma^n(\theta)$ we can bound the equilibrium utility of this type from below as follows:

$$U^n(\theta) = u^n(\widetilde{x}^n(\theta), \widetilde{\mathbf{m}}^n(\theta), \theta) \geq u^n(BR(\mu(\gamma^n(\theta)), \gamma^n(\theta), \theta) = u(BR(\mu(\gamma^n(\theta)), \theta) \tag{25}$$

Because $BR(\mu)$ is single-valued and $\mu(\gamma^n(\theta))$ converges weakly to $\delta(\theta)$, we then have $\lim_{n\to\infty} BR(\mu(\gamma^n(\theta)) \to x^*(\theta)$. Hence,

$$\lim_{n\to\infty} \inf U^n(\theta) \geq u(x^*(\theta), \theta). \tag{26}$$

Next, we show that $\mu(\widetilde{\mathbf{m}}^n(\theta))$ converges weakly to $\delta(\theta)$ for all θ. Indeed, suppose to the contrary that for each n there existed $\theta^{'}_n$ such that $\widetilde{\mathbf{m}}^n(\theta'_n) = \widetilde{\mathbf{m}}^n(\theta)$, and such that $\theta^{'}_n \to \theta' \neq \theta$.[13] Then we claim that either $||\widetilde{\mathbf{m}}^n(\theta) - \gamma^n(\theta)|| \to \infty$ or $||\widetilde{\mathbf{m}}^n(\theta) - \gamma^n(\theta'_n)|| \to \infty$

[13] If θ'_n does not converge, we let θ' be the limit of any convergent subsequence. Without loss of generality, let this subsequence be the original sequence.

as $n \to \infty$. Indeed, suppose there were a constant $k < \infty$ such that $||\widetilde{\mathbf{m}}^n(\theta) - \gamma^n(\theta)|| \leq k$ and $||\widetilde{\mathbf{m}}^n(\theta) - \gamma^n(\theta'_n)|| \leq k$, for all n. Then we have:

$$L\sqrt{n}||\theta'_n - \theta|| \ \leq ||\gamma^n(\theta'_n) - \gamma^n(\theta)|| \ = || \left(\widetilde{\mathbf{m}}^n(\theta) - \gamma^n(\theta)\right) - \left(\widetilde{\mathbf{m}}^n(\theta) - \gamma^n(\theta'_n)\right) || \ \leq$$
$$||\widetilde{\mathbf{m}}^n(\theta) - \gamma^n(\theta)|| \ + ||\widetilde{\mathbf{m}}^n(\theta) - \gamma^n(\theta'_n)|| \ \leq 2k.$$

Since $\lim_{n\to\infty} ||\theta'_n - \theta|| = ||\theta' - \theta|| = \varepsilon > 0$, the outer inequalities produce a contradiction.

Suppose first that $||\widetilde{\mathbf{m}}^n(\theta) - \gamma^n(\theta)|| \to \infty$. Then we have:

$$U^n(\theta) = u^n(\widetilde{x}^n(\theta), \widetilde{\mathbf{m}}^n(\theta), \theta) \leq u^n(\widetilde{x}^n(\theta), \gamma^n(\theta), \theta) - \alpha \ ||\widetilde{\mathbf{m}}^n(\theta) - \gamma^n(\theta)||^2 \leq u_{\max} - \alpha \ ||\widetilde{\mathbf{m}}^n(\theta) - \gamma^n(\theta)||^2,$$

where $u_{\max} = \max\{u(x, \theta) : \theta \in \Theta, \ \ x = BR(\mu) \text{ for some } \mu \in \Delta(\Theta)\}$. The above inequality implies that $U^n(\theta) \to -\infty$, contradicting (26). So we cannot have $||\widetilde{\mathbf{m}}^n(\theta) - \gamma^n(\theta)|| \to \infty$.

A parallel argument establishes that we cannot have $||\widetilde{\mathbf{m}}^n(\theta) - \gamma^n(\theta'_n)|| \to \infty$ either. It follows that for every sequence θ'_n such that $\widetilde{\mathbf{m}}^n(\theta'_n) = \widetilde{\mathbf{m}}^n(\theta)$ it must be that $\theta'_n \to \theta$. We conclude that as n grows without bound, $\mu(\widetilde{\mathbf{m}}^n(\theta))$ converges weakly to $\delta(\theta)$ for all θ.

It then follows from the single-valuedness of $BR(\mu)$ that $\widetilde{x}^n(\theta) = BR(\mu(\widetilde{\mathbf{m}}^n(\theta)) \to x^*(\theta)$. Therefore, we have

$$U^n(\theta) = u^n(\widetilde{x}^n(\theta), \widetilde{\mathbf{m}}^n(\theta), \theta) \leq u^n(BR(\mu(\widetilde{\mathbf{m}}^n(\theta))), \gamma^n(\theta), \theta) = u(BR(\mu(\widetilde{\mathbf{m}}^n(\theta))), \theta) \to u(x^*(\theta), \theta) \tag{27}$$

Combining (26) and (27) yields

$$\lim_{n\to\infty} U^n(\theta) = u(x^*(\theta), \theta), \text{ for all } \theta,$$

completing the proof of the theorem. *Q.E.D.*

Proof of Theorem 2. Fix any pair of twice continuously differentiable functions $q : \Theta \to Q$ and $t : \Theta \to \mathbb{R}$, and let $U(\theta)$ be the associated payoff of agent-type θ i.e., $U(\theta) = t(\theta) - h(q(\theta), \theta)$. We will show that there exists $N < \infty$ and a sequence of transfers and messages rules $(t^n, \mathbf{m}^n)$, such that for all $n \geq N$ and all $\theta \in \Theta$, $(q(.), t^n(.), \mathbf{m}^n(.))$ is an incentive compatible mechanism which provides each agent-type with the desired net payoff $U(\theta)$ i.e.,

$$U(\theta) \equiv t^n(\theta) - h(q(\theta), \theta) - C^n(\mathbf{m}^n(\theta), \theta) = \max_{\theta' \in \Theta}\{t^n(\theta') - h(q(\theta'), \theta) - C^n(\mathbf{m}^n(\theta'), \theta)\} \tag{28}$$

Significantly, we will also show that the total message cost $C^n(\mathbf{m}^n(\theta), \theta)$ of every type θ converges to zero as n increases.

The proof goes through several steps. In Step (i) we construct the message rule $\mathbf{m}^n(\theta)$. In Step (ii) we show that show that $||\mathbf{m}^n(\theta) - \ \gamma^n(\theta)|| \to 0$, uniformly in θ. Step 3 establishes that, under out message rule $\mathbf{m}^n(\theta)$, $C^n(\mathbf{m}^n(\theta), \theta)$ goes to zero in n. Finally, in Step (iv) we define the transfer rule $t^n(\theta)$ and establish the incentive compatibility of our mechanism.

Step (i). First, let us construct the message rule $\mathbf{m}^n(\theta)$.

Let $\mathbf{z}^n : \Theta \to \mathbb{R}^l$ and for each i set $\mathbf{m}_i(\theta) = \gamma_i(\theta) + \mathbf{z}^n(\theta)$. We choose $\mathbf{z}^n(\theta)$ as follows. Assuming that incentive constraints in (28) hold, the envelope theorem implies that

$$D_\theta U(\theta) = -D_\theta h(q(\theta), \theta) - D_\theta C^n(\mathbf{m}^n(\theta), \theta) \tag{29}$$

Therefore we select $\mathbf{z}^n(\theta)$ so that (29) holds. For each θ, (29) consists of l equations in the l unknown variables $\mathbf{z}^n(\theta)$.

We claim that for sufficiently large n, such a solution exists for all θ. To establish this, for each n define the function $f^n : \mathbb{R}^l \to \mathbb{R}^l$ by

$$f^n(\mathbf{z}) = n^{-(1-\alpha)} D_\theta C^n(\gamma_1(\theta) + \mathbf{z}, ..., \gamma_n(\theta) + \mathbf{z}, \theta),$$

and let $r^n(\theta) = n^{-(1-\alpha)} \left[-D_\theta h(q(\theta), \theta) - D_\theta U(\theta)\right]$. Then we may rewrite (29) as

$$f^n(\mathbf{z}) = r^n(\theta) \tag{30}$$

We now claim that $f^n(\mathbf{0}) = 0$. Indeed, differentiating the identity $C^n(\gamma^n(\theta), \theta) \equiv 0$, we obtain $D_{\mathbf{m}^n} C^n(\gamma^n(\theta), \theta) D_\theta \gamma^n(\theta) + D_\theta C^n(\gamma^n(\theta), \theta) = 0$. Since $\mathbf{m}^n = \gamma^n(\theta)$ uniquely minimizes $C^n(\cdot, \theta)$, we have $D_{\mathbf{m}^n} C^n(\gamma^n(\theta), \theta) = 0$, establishing that $D_\theta C^n(\gamma^n(\theta), \theta) = 0$, and hence that $f^n(\mathbf{0}) = \mathbf{0}$. Furthermore, we may calculate

$$Df^n(\mathbf{0}) = n^{-(1-\alpha)} \sum_{i=1}^n D^2_{\theta m_i} C^n(\gamma^n(\theta), \theta).$$

By Assumption 1(ii), we have $||Df^n(\mathbf{0})\mathbf{z})|| \geq \omega_1 ||\mathbf{z}||$, implying that $Df^n(\mathbf{0})$ is nonsingular. Thus we have all conditions for applying the inverse function theorem to equation (30). However, some care must be taken, because the function $f^n(\cdot)$ varies with n, and we have to ensure that for all n sufficiently large, $r^n(\theta)$ is sufficiently close to 0.

To this effect, let us define the function $s : \mathbb{R}^l \to \mathbb{R}^l$ by

$$s(\mathbf{z}) = \mathbf{z} - Df^n(\mathbf{0})^{-1}(f^n(\mathbf{z}) - \mathbf{y}).$$

Observe that $\mathbf{z}$ is a fixed point of $s(\cdot)$ if and only if $f^n(\mathbf{z}) = \mathbf{y}$. Observe also that

$$Ds(\mathbf{z}) = I - Df^n(\mathbf{0})^{-1} Df^n(\mathbf{z}),$$

where I is an identity matrix. Now select $\varepsilon = \frac{\omega_1}{2}$, and let δ be as given by Assumption 1(i). Then for $\mathbf{z}$ such that $\sqrt{n}||\mathbf{z}|| < \delta$ we have:

$$\begin{aligned} ||Ds(\mathbf{z})|| &= ||Df^n(\mathbf{0})^{-1} \left[Df^n(\mathbf{0}) - Df^n(\mathbf{z})\right]|| \\ &\leq ||Df^n(\mathbf{0})^{-1}|| || \left[Df^n(\mathbf{0}) - Df^n(\mathbf{z})\right]|| < \varepsilon \; ||Df^n(\mathbf{0})^{-1}|| \; \leq \frac{\varepsilon}{\omega_1} = \frac{1}{2}, \end{aligned}$$

where the penultimate inequality follows from Assumption 1(ii). It follows from the Mean Value inequality that if $\mathbf{z}_1, \mathbf{z}_2 \in B(\mathbf{0}, \frac{\delta}{\sqrt{n}})$, then

$$||s(\mathbf{z}_1) - s(\mathbf{z}_2)|| \le \max_{z \in B} ||Ds(z)|| \, ||\mathbf{z}_1 - \mathbf{z}_2|| \le \frac{1}{2}||\mathbf{z}_1 - \mathbf{z}_2||$$

Thus $s(\cdot)$ is a contraction on the closed ball $\overline{B}(\mathbf{0}, \frac{\delta}{\sqrt{n}})$.

Now let $t = \frac{\delta\omega_1}{4\sqrt{n}}$. Then for any $y \in B(\mathbf{0}, t)$, we have

$$||s(\mathbf{0})|| = ||Df^n(\mathbf{0})^{-1}(f^n(\mathbf{0}) - y)|| \le ||Df^n(\mathbf{0})^{-1}|| \, t \le \frac{\delta}{4\sqrt{n}} < \frac{\delta}{2\sqrt{n}}$$

It follows that

$$||s(\mathbf{z})|| = ||s(\mathbf{z}) - s(\mathbf{0}) + s(\mathbf{0})|| \le ||s(\mathbf{z}) - s(\mathbf{0})|| + ||s(\mathbf{0})|| \le \frac{1}{2}||\mathbf{z}|| + \frac{\delta}{2\sqrt{n}} \le \frac{\delta}{\sqrt{n}}.$$

Hence s maps $\overline{B}(\mathbf{0}, \frac{\delta}{\sqrt{n}})$ into itself. By the contraction mapping theorem, for any $\mathbf{y} \in B(\mathbf{0}, t)$ there exists a unique $\mathbf{z} \in \overline{B}(\mathbf{0}, \frac{\delta}{\sqrt{n}})$ such that $s(\mathbf{z}) = \mathbf{z}$, i.e. such that $\mathbf{y} = f^n(\mathbf{z})$.

It remains to be shown that for sufficiently large n, it is the case that $r^n(\theta) \in B(\mathbf{0}, t)$. Because the function $-D_\theta h(q(\theta), \theta) - D_\theta U(\theta)$ is continuous in θ and Θ is compact, it follows from the Weierstrass Theorem that there exists a constant $\lambda > 0$ s.t. $|| - D_\theta h(q(\theta), \theta) - D_\theta U(\theta)|| \le \lambda$ for all θ. Thus, $||r^n(\theta)|| \le \lambda n^{-(1-\alpha)} \le \frac{\delta\omega_1}{4} n^{-\frac{1}{2}}$ holds whenever $n^{\frac{1}{2}-\alpha} \ge \frac{4\lambda}{\delta\omega_1}$. Let N_1 be the smallest integer greater than $\left(\frac{4\lambda}{\delta\omega_1}\right)^{\frac{2}{1-2\alpha}}$. We may then conclude that whenever $n \ge N_1$, there exists a unique $\mathbf{z}^n(\theta) \in B(\mathbf{0}, \frac{\delta}{\sqrt{n}})$ solving (29).

Step (ii) Let us now show that $||\mathbf{m}^n(\theta) - \gamma^n(\theta)|| \to 0$, uniformly in θ. To this effect, we will tighten the inequality $||\mathbf{z}^n(\theta)|| \le \frac{\delta}{\sqrt{n}}$. We will establish that $||\mathbf{z}^n(\theta)|| \le \frac{2\lambda}{\omega_1} n^{-(1-\alpha)}$, and hence $||\mathbf{m}^n(\theta) - \gamma^n(\theta)|| = \sqrt{n}||\mathbf{z}^n(\theta)|| \le \frac{2\lambda}{\omega_1} n^{-(\frac{1}{2}-\alpha)}$.

To this end, let $t \in [0, 1]$ and define $\mathbf{m}^n(t)$ to be the vector whose components are $\mathbf{m}_i(t) = \gamma_i(\theta) + t\,\mathbf{z}$. We first claim that whenever $||\mathbf{z}|| < \frac{\delta}{\sqrt{n}}$, we have

$$||D_\theta C^n(\mathbf{m}^n, \theta) - D_\theta C^n(\gamma^n(\theta), \theta)|| \ge \frac{\omega_1}{2} n^{1-\alpha} ||\mathbf{z}||.$$

To establish the claim, define $\rho(t) = D_\theta C^n(\mathbf{m}^n(t), \theta) - D_\theta C^n(\gamma^n(\theta), \theta)$. Then we have

$$\begin{aligned} D_\theta C^n(\mathbf{m}^n, \theta) - D_\theta C^n(\gamma^n(\theta), \theta) &= \rho(1) - \rho(0) = \textstyle\int_0^1 D\rho(t)dt = \int_0^1 \sum_{i=1}^n D^2_{\theta m_i} C^n(\mathbf{m}^n(t), \theta)\mathbf{z}dt \\ &= \textstyle\int_0^1 \sum_{i=1}^n \left\{ D^2_{\theta m_i} C^n(\gamma^n(\theta), \theta) + \left[D^2_{\theta m_i} C^n(\mathbf{m}^n(t), \theta) - D^2_{\theta m_i} C^n(\gamma^n(\theta), \theta) \right] \right\} \mathbf{z}dt \\ &= \textstyle\sum_{i=1}^n D^2_{\theta m_i} C^n(\gamma^n(\theta), \theta)\mathbf{z} + \int_0^1 \left[D^2_{\theta m_i} C^n(\mathbf{m}^n(t), \theta) - D^2_{\theta m_i} C^n(\gamma^n(\theta), \theta) \right] \mathbf{z}dt \end{aligned}$$

It then follows from Assumptions 1(i) and (ii) that

$$\begin{aligned} ||D_\theta C^n(\mathbf{m}^n,\theta) - D_\theta C^n(\gamma^n(\theta),\theta)|| &\geq \omega_1 n^{1-\alpha}||\mathbf{z}|| - \int_0^1 ||D^2_{\theta m_i} C^n(\mathbf{m}^n(t),\theta) - D^2_{\theta m_i} C^n(\gamma^n(\theta),\theta)||||\mathbf{z}||dt \\ &\geq \omega_1 n^{1-\alpha}||\mathbf{z}|| \ - \varepsilon||\mathbf{z}|| \geq \frac{\omega_1}{2} n^{1-\alpha}||\mathbf{z}||. \end{aligned} \tag{31}$$

Now recall from the previous step that $\lambda \geq ||-D_\theta h(q(\theta),\theta) - D_\theta U(\theta)|| = ||D_\theta C^n(\mathbf{m}^n,\theta) - D_\theta C^n(\gamma^n(\theta),\theta)||$. Combining this inequality with inequality (31) and recalling that $||\mathbf{z}^n(\theta)|| \leq \frac{\delta}{\sqrt{n}}$, so it must satisfy (31), we obtain that $||\mathbf{z}^n(\theta)|| \leq \frac{2\lambda}{\omega_1} n^{-(1-\alpha)}$, as to be shown.

Step (iii). Next, let us show that $C^n(\mathbf{m}^n(\theta),\theta) \to 0$.

It follows from Assumption 1(v) and Step (ii) above that:

$$C^n(\mathbf{m}^n(\theta),\theta) \leq \omega_3 n^{1-\alpha}||\mathbf{z}(\theta)||^2 \leq \left(\frac{\lambda}{\omega_1}\right)^2 \omega_4 n^{-(1-\alpha)}$$

Consequently, $C^n(\mathbf{m}^n(\theta),\theta) \to 0$, uniformly in θ. From (28), this also implies that $t^n(\theta) \to t(\theta)$, uniformly in θ, completing the proof of the step.

Step (iv). Next, define the transfer rule $t^n(\theta)$ as follows:

$$t^n(\theta) = h(q(\theta),\theta) + C^n(\mathbf{m}^n(\theta),\theta) + U(\theta) \tag{32}$$

where $\mathbf{m}^n(.)$ is the message rule defined in Step (i).

The mechanism $(q(.), t^n(.), \mathbf{m}^n(.))$ is incentive compatible and gives each type θ the net payoff $U(\theta)$ if and only if for all $\theta',\theta \in \Theta$ we have:

$$U(\theta'^n(\theta) - h(q(\theta),\theta'^n(\mathbf{m}^n(\theta),\theta').$$

Substituting (32) into the above inequality we obtain the following equivalent condition:

$$V(\theta',\theta) \equiv h(q(\theta),\theta) - h(q(\theta),\theta') + U(\theta) - U(\theta') + C^n(\mathbf{m}^n(\theta),\theta) - C^n(\mathbf{m}^n(\theta),\theta') \leq 0 \tag{33}$$

In words, $V(\theta',\theta)$ is the amount by which the payoff of type θ' changes when she imitates θ instead of following her truthful strategy.

In the rest of the proof we will show that $V(\theta',\theta) \leq 0$ for all $\theta',\theta \in \Theta$. The proof consists of two parts that deal with, correspondingly, large deviations when $||\theta - \theta|| \geq \frac{\delta}{2}$ for some $\delta > 0$, and local deviations when $||\theta - \theta|| < \frac{\delta}{2}$.

First, let us show that for sufficiently high n large deviations from truthtelling are not optimal. To this end, pick $\varepsilon \in (0, \min\{1, \omega_1\})$, and consider the corresponding δ as given by Assumption 1(iii). To establish this claim, we will show that there exists N_2 ,$N_1 \leq N_2 < \infty$, such that for all $n \geq N_2$ and all $\theta',\theta \in \Theta$ satisfying $||\theta' - \theta|| \geq \frac{\delta}{2}$, we have $V(\theta',\theta) < 0$.

To this effect, observe that by Assumption 1(iv) we have

$$C^n(\mathbf{m}^n(\theta),\theta') - C^n(\mathbf{m}^n(\theta),\theta) \geq \frac{\omega_3}{n^\alpha}||\mathbf{m}^n(\theta) - \gamma^n(\theta')||^2 - \frac{\omega_4}{n^\alpha}||\mathbf{m}^n(\theta) - \gamma^n(\theta)||^2$$

It follows from Assumption 1(vi) that

$$||\gamma(\theta) - \gamma(\theta')||^2 = \sum_{i=1}^{n} \ ||\gamma_i(\theta) - \gamma_i(\theta)||^2 \ \geq nL^2 \ ||\theta - \theta'||^2$$

Moreover, since $||a+b|| \geq ||a| - ||b||$ for any two vectors a and b, we have

$$||\mathbf{m}^n(\theta) - \gamma^n(\theta')|| \ = ||\gamma^n(\theta) - \gamma^n(\theta') + \mathbf{m}^n(\theta) - \gamma^n(\theta)||$$
$$\geq ||\gamma^n(\theta) - \gamma^n(\theta')|| \ - ||\mathbf{m}^n(\theta) - \gamma^n(\theta)|| \geq \sqrt{n}L \ ||\theta - \theta'|| \ - \ \sqrt{n}||\mathbf{z}^n(\theta)|| \geq \sqrt{n}(\frac{\delta L}{2} - \frac{2\lambda}{\omega_1}n^{-(1-\alpha)})$$

Now select μ to satisfy $\omega_3(1-\mu)^2 - \omega_4\mu^2 \geq \frac{\omega_3}{2}$ and let $N_{21} \geq N_1$ be such that for all $n \geq N_{21}$ we have $\frac{2\lambda}{\omega_1}n^{-(1-\alpha)} \leq \frac{\delta L}{2}\mu$. It follows that $||\mathbf{m}^n(\theta) - \gamma^n(\theta')|| \ \geq \sqrt{n}\frac{\delta L}{2}(1-\mu)$.

By step (ii) we have $\frac{\omega_4}{n^\alpha}||\mathbf{m}^n(\theta) - \gamma^n(\theta)||^2 \ = \frac{\omega_4}{n^\alpha} \ n \ ||\mathbf{z}^n(\theta)||^2 \leq \omega_4 \left(\frac{2\lambda}{\omega_1}\right)^2 n^{-(1-\alpha)}$. It follows that

$$C^n(\mathbf{m}^n(\theta),\theta') - C^n(\mathbf{m}^n(\theta),\theta) \geq \omega_3 n^{(1-\alpha)} \left(\frac{\delta L}{2}(1-\mu)\right)^2 - \omega_4 \left(\frac{2\lambda}{\omega_1}\right)^2 n^{-(1-\alpha)}$$
$$\geq n^{(1-\alpha)} \left(\frac{\delta L}{2}\right)^2 \left(\omega_3(1-\mu)^2 - \omega_4\mu^2\right) \geq n^{(1-\alpha)} \left(\frac{\delta L}{2}\right)^2 \frac{\omega_3}{2}$$

Now define

$$\kappa = | \max_{\theta,\theta'} \left\{h(q(\theta),\theta) - h(q(\theta),\theta') + U(\theta) - U(\theta')\right\} |$$

Next, let $N_{22} \geq N_1$ be such that $n^{(1-\alpha)} \left(\frac{\delta L}{2}\right)^2 \frac{\omega_3}{2} \geq 2\kappa$ for all $n \geq N_{22}$. Then we have $V(\theta',\theta) \leq -\kappa < 0$.

Thus, letting $N_2 = \max\{N_{21}, N_{22}\}$, we conclude that for every $\delta > 0$, $n \geq N_2$, and all θ and all θ' such that $||\theta - \theta'|| \geq \frac{\delta}{2}$, we have $V(\theta',\theta) \leq -\kappa < 0$, as was to be shown.

It remains to establish that $V(\theta',\theta) < 0$ for all $\theta' \neq \theta$ such that $||\theta - \theta'|| < \frac{\delta}{2}$ when n is sufficiently large. We will do so by showing that at all such $\theta' \neq \theta$ the directional derivative of $V(.,\theta)$ at θ' in the direction of θ is positive.

Differentiating (33) and then using (29) we get:

$$D_{\theta'}V(\theta',\theta) = -D_\theta h(q(\theta),\theta') - D_\theta U(\theta') - D_\theta C^n(\mathbf{m}^n(\theta),\theta') =$$
$$-\big(D_\theta h(q(\theta),\theta') - D_\theta h(q(\theta),\theta)\big) - \big(D_\theta U(\theta') - D_\theta U(\theta)\big) - \big(D_\theta C^n(\mathbf{m}^n(\theta),\theta') - D_\theta C^n(\mathbf{m}^n(\theta),\theta)\big) \quad (34)$$

By (34), the directional derivative of $V(\cdot,\theta)$ at θ' in the direction of θ is given by:

$$\begin{aligned} D_{\theta'}V(\theta',\theta)(\theta-\theta') &= -\left(D_\theta h(q(\theta),\theta') - D_\theta h(q(\theta),\theta)\right)(\theta-\theta') - \left(D_\theta U(\theta') - D_\theta U(\theta)\right)(\theta-\theta') \\ &\quad - \left(D_\theta C^n(\mathbf{m}^n(\theta),\theta') - D_\theta C^n(\mathbf{m}^n(\theta),\theta)\right)(\theta-\theta') \end{aligned}$$

Since the closed ball $B(\theta,\frac{\delta}{2})$ is convex and $\theta' \in B(\theta,\frac{\delta}{2})$ it follows that $\theta + t(\theta'-\theta) \in B(\theta,\frac{\delta}{2})$ for all $t \in [0,1]$. Now let $\xi(t) = D_\theta C^n(\mathbf{m}^n(\theta), \theta + t(\theta'-\theta))$. Then we have

$$D_\theta C^n(\mathbf{m}^n(\theta),\theta') - D_\theta C^n(\mathbf{m}^n(\theta),\theta) = \xi(1)-\xi(0) = \int_0^1 D\xi(t)dt = \int_0^1 (\theta'-\theta)' D_{\theta\theta}C^n(\mathbf{m}^n(\theta), \theta+t(\theta'-\theta))dt$$

Now select N_{21} such that for all $n \geq N_{21}$ we have $\frac{\lambda}{\omega_1} n^{\alpha-\frac{1}{2}} < \frac{\delta}{2}$. It follows that for all $n \geq N_{21}$ and all $t \in [0,1]$ we have

$$||(\mathbf{m}^n, \theta+t(\theta'-\theta))) - (\gamma^n(\theta),\theta)|| \leq ||\mathbf{m}^n - \gamma^n(\theta)|| + ||\theta'-\theta|| \leq \sqrt{n}\,||\mathbf{z}^n(\theta)|| + \frac{\delta}{2} \leq \frac{\lambda}{\omega_1} n^{\alpha-\frac{1}{2}} + \frac{\delta}{2} < \delta$$

Next, define matrix A as follows: $A = D_{\theta\theta}C^n(\mathbf{m}^n(\theta), \theta + t(\theta'-\theta)) - D_{\theta\theta}C^n(\gamma^n(\theta),\theta)$. By Assumption 1(iii), $||A|| < \varepsilon n^{1-\alpha}$. Furthermore, A is a symmetric matrix, so that all of its eigenvalues are real. It follows that $x'Ax \leq ||A||||x||^2$. Indeed, letting Λ denote the matrix of eigenvalues of A, and K the orthogonal basis of its eigenvectors, we have $x'Ax = x'K'\Lambda Kx = y'\Lambda y = \sum_{i=1}^n \lambda_i y_i^2 \leq \sum_{i=1}^n |\lambda_i| y_i^2 \leq \max_i |\lambda_i| \sum_{i=1}^n y_i^2 = ||A||^{\frac{1}{2}}\ ||y||^2 = ||A||^{\frac{1}{2}}\ ||x||^2$. Hence we have:

$$\begin{aligned} &(\theta'-\theta)' D_{\theta\theta}C^n(\mathbf{m}^n(\theta), \theta + t(\theta'-\theta))(\theta'-\theta) = (\theta'-\theta)\left[D_{\theta\theta}C^n(\gamma^n(\theta),\theta) + A\right](\theta'-\theta) \\ &\geq \omega_2 n^{1-\alpha}||\theta'-\theta||^2 - ||A||^{\frac{1}{2}}\ ||\theta'-\theta||^2 > (\omega_2 - \varepsilon)n^{1-\alpha}||\theta'-\theta||^2 > \frac{\omega_2}{2} n^{1-\alpha}||\theta'-\theta||^2 \end{aligned}$$

Let $\beta = \max_{\theta',\theta\in\Theta} ||D_{\theta\theta} - h(q(\theta),\theta') - D_{\theta\theta}U(\theta')||$. Then

$$D_{\theta'}V(\theta',\theta)(\theta-\theta') \geq \left(\frac{\omega_2}{2} n^{1-\alpha} - \beta\right) ||\theta'-\theta||^2$$

Let $N_{32} \geq N_2$ be such that $n^{1-\alpha} > \frac{2\beta}{\omega_2}$ for all $n \geq N_{32}$, and let $N_3 = \max\{N_{31}, N_{32}\}$. Then for all $n \geq N_3$, all θ and $\theta^{'} \in B(\theta,\frac{\delta}{2})$ the directional derivative $D_{\theta'}V(\theta',\theta)(\theta-\theta')$ is strictly positive, implying that $\theta' = \theta$ uniquely maximizes $V(\theta',\theta)$ over $B(\theta,\frac{\delta}{2})$. Since $V(\theta,\theta) = 0$, it follows that $V(\theta',\theta) \leq 0$ for $\theta' \in B(\theta,\frac{\delta}{2})$. *Q.E.D.*

Proof of Theorem 3: The proof proceeds through several claims.

In the initial claims, we characterize the solution to the relaxed program (14) subject to (12) and (15). Then to complete the proof, we show that the solution to the relaxed program is globally incentive compatible, and hence solves the full unrelaxed program.

Claim 1. *There exists $\theta_1 \in (\underline{\theta}, \overline{\theta})$ such that $U(\theta) > 0$ and hence $\sigma(\theta) = F(\theta)$ for all $\theta \in [\underline{\theta}, \theta_1)$.*

The proof is by contradiction. So suppose to the contrary that such θ_1 does not exist. As a solution to the optimal control problem, $U(.)$ must be continuous, and so $U(\underline{\theta}) = 0$. By the transversality condition at $\underline{\theta}$, $\sigma(\underline{\theta}) \leq 0$. So, the first-order conditions (17) and (18) imply that $q(\underline{\theta}) \geq q^{FB}(\underline{\theta})$ and $m_i(\underline{\theta}) \leq \gamma_i(\underline{\theta})$. Since $h_{q\theta} > 0$, $C^n_{\theta m_i} < 0$ and $C^n_\theta(\gamma^n(\theta), \theta) = 0$, we have:

$$U'(\underline{\theta}) = -h_\theta(q(\underline{\theta}), \underline{\theta}) - C^n_\theta(\mathbf{m}^n(\underline{\theta}), \underline{\theta}) \leq -h_\theta(q^{FB}(\underline{\theta}), \underline{\theta}) - C^n_\theta(\gamma^n(\underline{\theta}), \underline{\theta}) < 0,$$

The last inequality holds because $q^{FB}(\underline{\theta}) > 0$ and $h_\theta(q, \theta) > 0$ when $q > 0$. So, $U(\underline{\theta} + \epsilon) < 0$ for $\epsilon > 0$ sufficiently small. But this contradicts the individual rationality of the mechanism.

Finally, since $U(\theta) > 0$ on $[\underline{\theta}, \theta_1)$, by the transversality condition we have $\sigma(\underline{\theta}) = 0$, and by complementary slackness condition (20) we gave $\rho(\theta) = 0$. So by the costate equation (19) $\sigma(\theta) = F(\theta)$ on this interval.

Claim 2. $0 < \sigma(\theta)$ *for all* $\theta \in (\underline{\theta}, \overline{\theta}]$, and $\sigma(\theta) \leq F(\theta)$ *for all* $\theta \in [\underline{\theta}, \overline{\theta}]$.

The claim that $\sigma(\theta) \leq F(\theta)$ follows from the transversality condition $\sigma(\underline{\theta}) \leq 0$, the costate equation (19) and the complementary slackness condition $\rho(\theta) \geq 0$ in (20).

Now, let us show that $\sigma(\theta) > 0$ for all $\theta \in (\underline{\theta}, \overline{\theta})$. By claim 1, $\sigma(\underline{\theta}) = 0$, and $\sigma(\theta) = F(\theta) > 0$ for all $\theta \in [\underline{\theta}, \theta_1)$. By the transversality condition, $\sigma(\overline{\theta}) \geq 0$. Also, by the costate equation (19), $\sigma'(\theta) > 0$ for all θ s.t. $U(\theta) > 0$. So to complete the proof it suffices to show that there does not exist θ, $\theta < \overline{\theta}$, such that $U(\theta) = 0$ and $\sigma(\theta) = 0$. The proof is by contradiction. So suppose that such θ does exist. Then $q(\theta) = q^{FB}(\theta) > 0$ by (17), and $\mathbf{m}^n(\theta) = \gamma^n(\theta)$ by (18). Since $C^n_\theta(\gamma^n(\theta), \theta) = 0$, it then follows from (15) that $U'(\theta) = -h_\theta(q^{FB}(\theta), \theta) < 0$. So, $U(\theta + \delta) < 0$ for $\delta > 0$ sufficiently small. But this contradicts the individual rationality of the mechanism, thereby showing that $\sigma(\theta) > 0$ for all $\theta \in (\underline{\theta}, \overline{\theta})$.

Finally, let us show that $\sigma(\overline{\theta}) > 0$. Suppose to the contrary that we had $\sigma(\overline{\theta}) = 0$. Mimicking the proof from the previous paragraph, this implies that $U'(\overline{\theta}) < 0$, so we have $U(\theta) > 0$ in a left neighborhood of $\overline{\theta}$. Since $\sigma'(\theta) > 0$ whenever $U(\theta) > 0$, this implies that $\sigma(\theta) < 0$ over this interval. By the previous result, this then implies that we must have $U(\theta) > 0$ for all $\theta \in (\underline{\theta}, \overline{\theta})$, implying $\sigma(\underline{\theta}) < 0$, contradicting that $\sigma(\underline{\theta}) = 0$. This establishes that $\sigma(\overline{\theta}) > 0$.

Claim 3. *For any $\theta \in [\underline{\theta}, \overline{\theta}]$ and fixed $\sigma \geq 0$ let $q(\sigma, \theta)$ and $\mathbf{m}^n(\sigma, \theta)$ maximize the Hamiltonian H in (16) w.r.t. q and $\mathbf{m}^n$, respectively.*

Then $q(\sigma, \theta)$ is decreasing in σ, strictly so whenever $q(\sigma, \theta) > 0$, and $\mathbf{m}^n(\sigma, \theta)$ is increasing in θ.

By definition, $q(\sigma, \theta)$ satisfies (17) and is the solution in q to

$$\max_{q \geq 0} \left\{ v(q) - h(q, \theta) - \frac{\sigma}{f(\theta)} h_\theta(q, \theta) \right\}, \tag{35}$$

while $\mathbf{m}^n(\sigma,\theta)$ satisfies (18) and is the solution in $\mathbf{m}^n$ to

$$\min_{\mathbf{m}^n\in\mathbb{R}^n}\left\{C^n(\mathbf{m}^n,\theta)+\frac{\sigma}{f(\theta)}C^n_\theta(\mathbf{m}^n,\theta)\right\}. \tag{36}$$

The existence of $q(\sigma,\theta)$ and $\mathbf{m}^n(\sigma,\theta)$ is guaranteed by the Weierstrass theorem because, respectively: (i) $q(\sigma,\theta)$ belongs to $[0,q^{FB}(\theta)]$; (ii) the value of (36) goes to ∞ as $||\mathbf{m}^n||\to\infty$.[14]

Next, observe that the cross partial of the objective (35) in (q,σ) is equal to $-\frac{1}{f(\theta)}h_{q\theta}(q,\theta)<0$. Hence it has decreasing differences in (q,σ), and so $q(\sigma,\theta)$ must be decreasing in σ by monotone comparative statics. Similarly, $\mathbf{m}^n(\sigma,\theta)$ is increasing in σ. Indeed, the cross-partial of the objective in (36) equals $\frac{1}{f(\theta)}\frac{\partial^2C^n}{\partial m_i\partial\theta}(\mathbf{m}^n,\theta)<0$, so it has decreasing differences in $(\mathbf{m}^n,\sigma)$. Furthermore (36) is submodular in $\mathbf{m}^n$ because $\frac{\partial^2C^n}{\partial m_i\partial m_j}+\frac{\sigma}{f(\theta)}\frac{\partial^3C^n}{\partial m_i\partial m_j\partial\theta}<0$. When $\frac{\partial^3C^n}{\partial m_i\partial m_j\partial\theta}\leq 0$ this inequality follows from Assumption 2 (iii); when $\frac{\partial^3C^n}{\partial m_i\partial m_j\partial\theta}>0$, it follows from Assumption 2 (ii) and the fact that $0\leq\sigma\leq F(\theta)$.

Claim 4. *Let* $U'(\sigma,\theta)=-h_\theta(q(\sigma,\theta),\theta)-C^n_\theta(\mathbf{m}^n(\sigma,\theta),\theta)$.

Then for every $\theta\in[\underline{\theta},\overline{\theta}]$*, there exists a unique* $\widehat{\sigma}(\theta)$*, such that* $U'(\sigma,\theta)=0$ *if* $\sigma=\widehat{\sigma}(\theta)$*,* $U'(\sigma,\theta)<0$ *if* $\sigma<\widehat{\sigma}(\theta)$ *and* $U'(\sigma,\theta)>0$ *if* $\sigma>\widehat{\sigma}(\theta)$*.*

To prove this Claim, first observe that from the definition of $U'(\sigma,\theta)$ it follows that

$$\frac{\partial U'(\sigma,\theta)}{\partial\sigma}=-h_{q\theta}\frac{\partial q(\sigma,\theta)}{\partial\sigma}-\sum_{i=1}^n\frac{\partial^2C^n}{\partial m_i\partial\theta}\frac{\partial m_i(\sigma,\theta)}{\partial\sigma}$$

Since $\frac{\partial q(\sigma,\theta)}{\partial\sigma}\leq 0$ and $\frac{\partial m_i(\sigma,\theta)}{\partial\sigma}>0$, it follows that $U'(\sigma,\theta)$ is strictly increasing in σ, and hence there exists at most one value σ such that $U'(\sigma,\theta)=0$. To show that such a value exists, let us show that $U'(\sigma,\theta)<0$ at $\sigma=0$, and $U'(\sigma,\theta)>0$ when σ is sufficiently large.

First, consider $\sigma=0$. It follows from (35) and (36) that for all θ, $q(0,\theta)=q^{FB}(\theta)$ and $\mathbf{m}^n(0,\theta)=\gamma^n(\theta)$, respectively. Since $C^n_\theta(\gamma(\theta),\theta)=0$, we have $U'(0,\theta)=-h_\theta(q^{FB}(\theta),\theta)<0$.

Next, let $\alpha(\theta)=\min_{q\in[0,q^{FB}(\theta)]}h_{q\theta}(q,\theta)$ and let $\overline{\sigma}=\frac{v'(0)-h_q(0,\theta)}{\alpha(\theta)}f(\theta)$. We now claim that $q(\sigma,\theta)=0$ for all $\sigma\geq\overline{\sigma}$. Suppose to the contrary that $q(\sigma,\theta)>0$ for some $\sigma\geq\overline{\sigma}$. Note that $q(\sigma,\theta)\leq q^{FB}(\theta)$ because $\sigma>0$. Moreover, $q(\sigma,\theta)$ satisfies (17) from which we have:

$$v_q(q(\sigma,\theta))-h_q(q(\sigma,\theta),\theta)=\frac{\sigma}{f(\theta)}h_{q\theta}(q(\sigma,\theta),\theta)\geq\frac{\overline{\sigma}}{f(\theta)}h_{q\theta}(q(\sigma,\theta),\theta)\geq v_q(0)-h_q(0,\theta) \tag{37}$$

But (37) contradicts the assumption that $v_q(q)-h_q(q,\theta)$ is strictly decreasing in q.

Now fix some σ s.t. $\sigma\geq\overline{\sigma}$. Then $q(\sigma,\theta)=0$. Since $h(0,\theta)=0$ for all θ, we have $h_\theta(0,\theta)=0$. Furthermore, $\mathbf{m}^n(\sigma,\theta)>\gamma^n(\theta)$ since $\sigma>0$, $\mathbf{m}^n(0,\theta)=\gamma^n(\theta)$, and $\mathbf{m}^n(\sigma,\theta)$ is

[14] Henceforth, we will also assume that $q(\sigma,\theta)$ and $\mathbf{m}^n(\sigma,\theta)$ are unique. This can always be guaranteed by assuming that the objective function in (35) is quasiconcave in q, and that the objective function in (36) is quasiconvex in $\mathbf{m}^n$.

increasing in σ. Because $C^n_\theta(\gamma^n(\theta),\theta)=0$, and $\frac{\partial^2 C^n}{\partial\theta\partial m_i}<0$, we have $C^n_\theta(\mathbf{m}^n(\sigma,\theta),\theta)<0$. So in this case $U'(\sigma,\theta)=-C^n_\theta(\mathbf{m}^n(\sigma,\theta),\theta)>0$ for $\sigma\geq\overline{\sigma}$.

Claim 5. *Let* $\sigma(\theta)=F(\theta)$ *and* $U'(F(\theta),\theta)=-h_\theta(q(F(\theta),\theta),\theta)-C^n_\theta(\mathbf{m}^n(F(\theta),\theta),\theta)$. *Define* $\widehat{\theta}=\inf\{\theta|\theta\leq\overline{\theta},U''(F(\theta),\theta)>0\}$. *Then* $0<\widehat{\theta}\leq\theta^*$.

Combining the arguments in Claim 1 and Claim 4 we obtain that $\widehat{\sigma}(\underline{\theta})>0=F(\underline{\theta})$. So, $U'(0,\underline{\theta})<0$ and therefore $\widehat{\theta}>0$.

Now, suppose that contrary to the Claim, $\widehat{\theta}>\theta^*$. From (35) and (36), or alternatively, the first-order conditions (17) and (18). it follows that for all $\theta\in[\theta^*,\widehat{\theta})$, $q(F(\theta),\theta)=q^{SB}(\theta)=0$ and $\widetilde{m}_i(F(\theta),\theta)>\gamma_i(\theta)$ for all i. The former implies that $h_\theta(q(F(\theta),\theta),\theta)=h_\theta(0,\theta)=0$ and the latter implies that $C^n_\theta(\widetilde{\mathbf{m}}^n(F(\theta),\theta),\theta)<0$. So for any $\theta\in(\theta^*,\widehat{\theta})$, we have $U'(F(\theta),\theta)>0$, a contradiction. to the definition of $\widehat{\theta}$.

Claim 6. *Type* $\widehat{\theta}$ *satisfies* $\widehat{\theta}=\inf\{\theta:\widehat{\sigma}(\theta)\leq F(\theta)\}$.

Combining Claim 4 with the definition of $\widehat{\theta}$ in Claim 5 yields this Claim.

Claim 7. *In the solution to the relaxed program* $q(\theta)=\widetilde{q}(\theta)=q^{SB}(\theta)$, $\mathbf{m}^n(\theta)=\widetilde{\mathbf{m}}^n(\theta)$ *and* $U'(\theta)\leq 0$ *for all* $\theta\in[\underline{\theta},\widehat{\theta})$.

By Claim 6, $\widehat{\sigma}(\theta)\geq F(\theta)$ for all $\theta\in[\underline{\theta},\widehat{\theta})$. By Claim 2, $\sigma(\theta)\leq F(\theta)$. So to prove that $\sigma(\theta)=F(\theta)$ on this interval, we need to rule out $\sigma(\theta)<F(\theta)$ for some $\theta\in[\underline{\theta},\widehat{\theta})$. Suppose to the contrary that suppose there exists $\theta_2\in[\underline{\theta},\widehat{\theta})$ such that $\sigma(\theta_2)<F(\theta_2)$. By Claim 1, we have $\theta_2>\theta_1$. By the costate equation (19) there then exists an open set $(\theta_3,\theta_4)\subseteq(\theta_1,\theta_2)$ such that $\rho(\theta)>0$ and so $\sigma(\theta)<F(\theta)\leq\widehat{\sigma}(\theta)$ for all $\theta\in(\theta_3,\theta_4)$. But then for all $\theta\in(\theta_3,\theta_4)$ we have both that $U(\theta)=0$ and, by claim 4, $U'(\theta)<0$. This contradiction establishes that $\sigma(\theta)=F(\theta)$, and so $q(\theta)=\widetilde{q}(\theta)=q^{SB}(\theta)$, $\mathbf{m}^n(\theta)=\widetilde{\mathbf{m}}^n(\theta)$ for all $\theta\in[\underline{\theta},\widehat{\theta}]$.

Claim 8. *In the optimal mechanism* $q(\theta)=\widehat{q}(\theta)$, $\mathbf{m}^n(\theta)=\widehat{\mathbf{m}}^n(\theta)$ *and* $U(\theta)=0$ *for all* $\theta\in[\widehat{\theta},\overline{\theta}]$.

Let us show that in the solution to the relaxed program $U(\theta)=0$ for θ in some interval $[\widehat{\theta},\widehat{\theta}+\epsilon]$. Suppose to the contrary that the individual rationality constraint is non-binding on an interval $[\widehat{\theta},\widehat{\theta}+\delta)$, for some $\delta>0$. Then by complementary slackness, we have $\rho(\theta)=0$ on this interval. So from $\sigma(\widehat{\theta})=F(\widehat{\theta})$ and the costate equation (19) it follows that $\sigma(\theta)=F(\theta)$ and hence we have $q(\theta)=q^{SB}(\theta),\mathbf{m}^n(\theta)=\widetilde{\mathbf{m}}^n(\theta)$ on $[\widehat{\theta},\widehat{\theta}+\delta)$.

So on the interval $[\widehat{\theta},\widehat{\theta}+\epsilon)$ for some $\epsilon\in(0,\delta)$ we have $U'(\theta)=U'(F(\theta),\theta)>0$ by the definition of $\widehat{\theta}$ in Claim 5, and $q^{FB}(\theta)>\widehat{q}(\theta)>q^{SB}(\theta)=q(F(\theta),\theta)$, $\gamma(\theta)<\widehat{\mathbf{m}}^n(\theta)<\widetilde{\mathbf{m}}^n(\theta)$ by Claim 3. But then the value of the relaxed program can be strictly increased by using the solution $(\widehat{q}(\theta),\widehat{\mathbf{m}}^n(\theta),\widehat{\sigma}(\theta))$ and setting $U(\theta)=0$ on the interval $[\widehat{\theta},\widehat{\theta}+\epsilon)$. This is so because

$$v(\widehat{q}(\theta),\theta)-h(\widehat{q}(\theta),\theta)-C^n(\widehat{\mathbf{m}}^n(\theta),\theta)>v(q^{SB}(\theta),\theta)-h(q^{SB}(\theta),\theta)-C^n(\widetilde{\mathbf{m}}^n(\theta),\theta)$$

This contradiction establishes that we must have $U(\theta)=0$ for $\theta\in[\widehat{\theta},\widehat{\theta}+\delta)$. This can easily done by choosing $U(\underline{\theta})$ appropriately.

A similar argument establishes that we cannot have $U'(\theta) > 0$ in the solution of the relaxed program at any other θ, $\theta > \widehat{\theta}$.

Finally, we cannot have $U'(\theta) < 0$ in the solution of the relaxed program at any θ, $\theta > \widehat{\theta}$, because it would violate individual rationality.

Claim 9. *Global incentive compatibility of the solution to the relaxed program.*

It remains to show that this solution satisfies incentive constraints (11) i.e., for any pair of types (θ, θ') we have:

$$U(\theta) - U(\theta') + h(q(\theta'), \theta) + C^n(\mathbf{m}^n(\theta'), \theta) - h(q(\theta'), \theta') - C^n(\mathbf{m}^n(\theta'), \theta') \geq 0 \tag{38}$$

First, suppose that $\theta' \in [\hat{\theta}, \overline{\theta}]$ i.e. $U(\theta') = 0$. We will consider the case $\theta' > \theta$. The proof for the case $\theta' < \theta$ is similar. Then we can rewrite the left-hand side of inequality (38) as follows.

$$\begin{aligned} &U(\theta) - U(\theta') - \int_{\theta}^{\theta'} h_\theta(q(\theta'), s) + C^n_\theta(\mathbf{m}^n(\theta'), s) ds \geq \\ &\int_{\theta}^{\theta'} h_\theta(q(\theta'), \theta') + C^n_\theta(\mathbf{m}^n(\theta'), \theta') - h_\theta(q(\theta'), s) - C^n_\theta(\mathbf{m}^n(\theta'), s) ds \\ &= \int_{\theta}^{\theta'} \int_{s}^{\theta'} h_{\theta\theta}(q(\theta'), t) + C^n_{\theta\theta}(\mathbf{m}^n(\theta'), t) dt ds > 0 \end{aligned} \tag{39}$$

The first inequality holds because $U(\theta) \geq 0$, and because $\theta' \in [\widehat{\theta}, \overline{\theta}]$ implies $U(\theta') = 0$ as well as $U'(\theta') = -h_\theta(q(\theta'), \theta') - C^n_\theta(\mathbf{m}^n(\theta'), \theta') = 0$. The second inequality holds because the first integral is non-positive as $h_{\theta\theta} \geq 0$ and $C^n_{\theta\theta} \geq 0$, establishing the incentive compatibility of our mechanism for this case.

Next, suppose that $\theta, \theta' \in [\underline{\theta}, \widehat{\theta}]$. Over this region, the solution is described by $\{q^{SB}(\theta), \widetilde{\mathbf{m}}^n(\theta), F(\theta)\}$, and incentive compatibility holds if $q^{SB}(\theta)$ is decreasing in θ, and $\widetilde{\mathbf{m}}^n(\theta)$ is increasing in θ (Guesnerie and Laffont, 1984, Theorem 2). That $q^{SB}(\theta)$ is decreasing in θ follows from Assumption 2(i). Next, as a maximizer of the Hamiltonian H, $\widetilde{\mathbf{m}}^n(\theta)$ maximizes:

$$-C^n(\mathbf{m}^n, \theta) - \frac{F(\theta)}{f(\theta)} C^n_\theta(\mathbf{m}^n, \theta)$$

By Assumption 2(ii), this objective has strictly increasing differences in $(\mathbf{m}^n, \theta)$, and is supermodular in $\mathbf{m}^n$. Therefore, $\widetilde{\mathbf{m}}^n(\theta)$ is increasing in θ, establishing incentive compatibility for this case.

Finally, let us show that incentive constraints hold for any pair (θ, θ') such that $\theta \in (\hat{\theta}, 1]$ and $\theta' \in [\underline{\theta}, \hat{\theta}]$. As shown above, incentive constraints hold between any $\theta \in (\hat{\theta}, 1]$ and $\hat{\theta}$, and also between $\hat{\theta}$ and any $\theta' \in [\underline{\theta}, \hat{\theta}]$ i.e,

$$\begin{aligned} &U(\theta) - U(\hat{\theta}) + h(q(\hat{\theta}), \theta) + C^n(\widetilde{\mathbf{m}}^n(\hat{\theta}), \theta) - h(q(\hat{\theta}), \hat{\theta}) - C^n(\widetilde{\mathbf{m}}^n(\hat{\theta}), \hat{\theta}) \geq 0 \\ &U(\hat{\theta}) - U(\theta') + h(q(\theta'), \hat{\theta}) + C^n(\widetilde{\mathbf{m}}^n(\theta'), \hat{\theta}) - h(q(\theta'), \theta') \geq 0 \end{aligned}$$

Adding the above inequalities, we get:

$$\begin{aligned} &U(\theta) - U(\theta') + h(q(\theta'), \hat{\theta}) + C^n(\widetilde{\mathbf{m}}^n(\theta'), \hat{\theta}) - h(q(\theta'), \theta') \\ &+ h(q(\hat{\theta}), \theta) + C^n(\widetilde{\mathbf{m}}^n(\hat{\theta}), \theta) - h(q(\hat{\theta}), \hat{\theta}) - C^n(\widetilde{\mathbf{m}}^n(\hat{\theta}), \hat{\theta}) \geq 0 \end{aligned} \tag{40}$$

Comparing (40) with (38) we conclude that the incentive constraint between θ and θ' holds if

$$\begin{aligned} &h(q(\theta', \theta) + C^n(\widetilde{\mathbf{m}}^n(\theta'), \theta) - h(q(\theta'), \hat{\theta}) - C^n(\widetilde{\mathbf{m}}^n(\theta'), \hat{\theta}) \geq \\ &h(q(\hat{\theta}), \theta) + C^n(\widetilde{\mathbf{m}}^n(\hat{\theta}), \theta) - h(q(\hat{\theta}), \hat{\theta}) - C^n(\widetilde{\mathbf{m}}^n(\hat{\theta}), \hat{\theta}) \end{aligned} \tag{41}$$

Finally, observe that inequality (41) holds because $h_{q\theta} > 0$, $q(\theta') > q(\hat{\theta})$, $C^n_{\theta m_i} \leq 0$, $\widetilde{\mathbf{m}}^n(\theta') < \widetilde{\mathbf{m}}^n(\hat{\theta})$ and $\theta > \hat{\theta}$. *Q.E.D.*

Proof of Theorem 4: First, by Theorem 3, $U(\theta) > 0$ and hence $\widetilde{q}(\theta) = q^{SB}(\theta) > 0$ for all $\theta \in [\underline{\theta}, \hat{\theta})$.

Now let us demonstrate that in the optimal mechanism, $\widehat{q}(\theta) > 0$ for all $\theta \in [\hat{\theta}, \overline{\theta})$. Suppose, instead, that $\widehat{q}(\theta) = 0$ for some $\theta \in [\hat{\theta}, \overline{\theta})$. Let us show that this implies $\widehat{\mathbf{m}}^n(\theta) = \gamma^n(\theta)$ and $\widehat{\sigma}(\theta) = 0$.

Note that $U'(\theta) = 0$ because $\theta \geq \hat{\theta}$. Since $h_\theta(0, \theta) = 0$, equation (15) therefore yields $C^n_\theta(\widehat{\mathbf{m}}^n(\theta), \theta) = 0$. Let us show that $\widehat{\mathbf{m}}^n(\theta) = \gamma^n(\theta)$ is the unique solution to this equation.

First, let us show that $\widehat{\mathbf{m}}^n(\theta) = \gamma^n(\theta)$ is a solution to $C^n_\theta(\widehat{\mathbf{m}}^n(\theta), \theta) = 0$. Indeed, the identity $C^n(\gamma^n(\theta), \theta) \equiv 0$ implies that $\sum_{i=1}^n C^n_{m_i}(\gamma^n(\theta), \theta)\gamma_i'(\theta) + C^n_\theta(\gamma^n(\theta), \theta) = 0$. Since $C^n(\mathbf{m}^n, \theta)$ attains a global minimum at $\mathbf{m}^n = \gamma^n(\theta)$, it follows that $C^n_{m_i}(\gamma^n(\theta), \theta) = 0$ for all i, and hence $C^n_\theta(\gamma^n(\theta), \theta) = 0$.

Further, to establish that $\widehat{\mathbf{m}}^n(\theta) = \gamma^n(\theta)$ is the unique solution to $C^n_\theta(\widehat{\mathbf{m}}^n(\theta), \theta) = 0$, consider equation (18). Since $C^n_{m_i\theta} < 0$ for all i, we have: (i) If $\sigma > 0$, then $C^n_{m_i}(\widehat{\mathbf{m}}^n(\theta), \theta) > 0$, $m_i(\theta) < \gamma_i(\theta)$ and hence $C^n_\theta(\widehat{\mathbf{m}}^n(\theta), \theta) < 0$; (ii) If $\sigma < 0$, then $C^n_{m_i}(\widehat{\mathbf{m}}^n(\theta), \theta) < 0$, $m_i(\theta) > \gamma_i(\theta)$ and hence $C^n_\theta(\widehat{\mathbf{m}}^n(\theta), \theta) > 0$. So, if $C^n_\theta(\widehat{\mathbf{m}}^n(\theta), \theta) = 0$, then we must have $\widehat{\sigma}(\theta) = 0$. So from (18) it follows that $C^n_{m_i\theta}(\widehat{\mathbf{m}}^n(\theta), \theta) = 0$ for all i, from which it follows that $\widehat{\mathbf{m}}^n(\theta) = \gamma^n(\theta)$.

But substituting $\widehat{\sigma}(\theta) = 0$ and $\widehat{q}(\theta) = 0$ into equation (17) then yields $v_q(0) - h_q(0, \theta) \leq 0$, contradicting the assumption that $v_q(0) - h_q(0, \theta) > 0$. So $\widehat{q}(\theta) > 0$ for all $\theta \in [\underline{\theta}, \overline{\theta})$. *Q.E.D.*

Proof of Theorem 5: First, we claim that $\sigma(\theta, n)$ converges to 0 as $n \to \infty$, uniformly in θ. By Claim 2 of Theorem 3, $\sigma(\theta, n) \geq 0$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$ and all n. Suppose now that contrary to the claim, there existed $\delta > 0$ such that for every N there exists $n \geq N$ and $\theta_n \in [\underline{\theta}, \overline{\theta}]$ such that $\sigma(\theta_n, n) > \delta$. Let σ_∞ be the limit of some convergent subsequence of $\sigma(\theta_n, n)\}_{n=1}^\infty$. By renumbering the indices of the subsequence, we may without loss of generality assume that

$\sigma(\theta_n, n) \to \sigma_\infty \geq \delta > 0$. We will show that this leads to a contradiction, thereby establishing the claim.

By the mean value Theorem, $\frac{\partial C^n}{\partial m_i}(m_i, \mathbf{m}^n_{-i}, \theta) - \frac{\partial C^n}{\partial m_i}(\gamma_i(\theta), \mathbf{m}^n_{-i}, \theta) = \frac{\partial^2 C^n}{\partial m_i^2}(\bar{m}_i, \mathbf{m}^n_{-i}, \theta)(\widehat{m}_i - \gamma_i(\theta))$, for some $\bar{m}_i \in (m_i, \gamma_i)$. Since $\frac{\partial C^n}{\partial m_i}(\gamma_i(\theta), \mathbf{m}^n_{-i}, \theta) = 0$, it follows from (18) that

$$(m_i(\theta) - \gamma_i(\theta)) = -\sigma(\theta, n) \frac{\frac{\partial^2 C^n}{\partial\theta\partial m_i}(\mathbf{m}^n(\theta), \theta)}{\frac{\partial^2 C^n}{\partial m_i^2}(\bar{m}_i, \mathbf{m}^n_{-i}(\theta), \theta)} \tag{42}$$

Using the fact that $C^n_\theta(\gamma^n(\theta), \theta) = 0$, and applying the mean value Theorem once more yields:

$$C^n_\theta(\mathbf{m}^n(\theta), \theta) = \sum_{i=1}^{n} \frac{\partial^2 C^n}{\partial\theta\partial m_i}(\bar{\mathbf{m}}^n, \theta)(m_i(\theta) - \gamma_i(\theta)), \tag{43}$$

where $\bar{\mathbf{m}}^n = \gamma^n(\theta) + \varepsilon(\theta)(\widehat{\mathbf{m}}^n - \gamma^n(\theta))$, for some $\varepsilon(\theta) \in (0, 1)$. Using the assumption that $0 \leq \frac{\partial^2 C^n}{\partial m_i^2} \leq \bar{v}$, $\left|\frac{\partial^2 C^n}{\partial\theta\partial m_i}\right| \geq \underline{v} > 0$, (42) and (43) yields:

$$C^n_\theta(\mathbf{m}^n(\theta), \theta) \leq -n\sigma(\theta, n)\frac{\underline{v}^2}{\bar{v}} \tag{44}$$

Since $\sigma(\theta_n, n) \to \sigma_\infty$, we therefore have $C^n_\theta(\mathbf{m}^n(\theta_n), \theta_n) \to -\infty$. But because $h_\theta(q, \theta)$ is bounded, (15) implies that $U'(\theta_n) > 0$ when n is large enough, contradicting that $U'(\theta) \leq 0$ on the interval $[\underline{\theta}, \overline{\theta}]$, and establishing the claim.

Because $\sigma(\theta, n) \to 0$, the first-order conditions (17) and (18) imply that $q(\theta, n) \to q^{FB}(\theta)$ and $m_i(\theta, n) \to \gamma_i(\theta)$, uniformly in θ.

Next, we argue that $\lim_{n\to\infty} \hat{\theta}(n) = \underline{\theta}$. Since $\widehat{\theta}(n)$ is the solution to the equation $\widehat{q}(\theta, n) = \widetilde{q}(\theta) = q^{SB}(\theta)$, it follows that $q(\hat{\theta}(n)) = q^{SB}(\hat{\theta}(n))$ for all n. Suppose now that contrary to the desired result, we had $\lim_{n\to\infty} \widehat{\theta}(n) = \theta$ for some $\theta \in (\underline{\theta}, \overline{\theta}]$. Then since $q^{SB}(\cdot)$ is continuous, we would have $q(\hat{\theta}(n)) \to q^{SB}(\theta) < q^{FB}(\theta)$, contradicting that $q(\cdot, \theta)$ converges uniformly to $q^{FB}(\cdot)$. This contradiction establishes that $\hat{\theta}(n) \to \underline{\theta}$.

It remains to show that $C^n(\mathbf{m}^n(\theta), \theta) \to 0$, uniformly in θ. Since $C^n(\gamma^n(\theta), \theta) = 0$ and $C^n_\theta(\gamma^n(\theta), \theta) = 0$, it follows from Taylor's Theorem that

$$C^n(\mathbf{m}^n(\theta), \theta) = \sum_{i=1}^{n} \frac{(m_i - \gamma_i(\theta))^2}{2} \frac{\partial^2 C^n}{\partial m_i^2}(\underline{\mathbf{m}}^n, \theta) \tag{45}$$

for some $\underline{\mathbf{m}}^n \in (\mathbf{m}^n(\theta), \gamma^n(\theta))$.

By (42), $|m_i - \gamma_i(\theta)| \leq |\sigma(\theta, n)|\frac{\bar{v}}{\underline{v}}$, so (45) implies $C^n(\mathbf{m}^n(\theta), \theta) \leq n\sigma^2(\theta, n)\frac{\bar{v}^3}{\underline{v}^2}$. By (44) and $n|\sigma(\theta, n)| \frac{\underline{v}^2}{\bar{v}} \leq |C^n_\theta(\mathbf{m}^n(\theta), \theta)| \leq \max_{(q,\theta)} h_\theta \equiv k$. Thus $C^n(\mathbf{m}^n(\theta), \theta) \leq k\ |\sigma(\theta, n)| \left(\frac{\bar{v}}{\underline{v}}\right)^2 \to 0$, uniformly in θ. *Q.E.D.*

Proof of Lemma 1: Since the solution to the principal's problem is unique, the value function W is continuously differentiable and

$$\frac{dW(n)}{dn} = \int_{\underline{\theta}}^{\bar{\theta}} \frac{\partial H}{\partial n}(q, m, U, \sigma, n, \theta) d\theta$$

(Seierstad and Sydsaeter (1999), p. 217). Using $\frac{\partial H}{\partial n} = -c(m(\theta), \theta) f(\theta) - \sigma(\theta) c_\theta(m(\theta), \theta)$ and substituting for $\sigma(\theta)$ from the first-order condition (18) yields $\frac{\partial H}{\partial n} = f\left(-c + \frac{c_\theta c_m}{c_{m\theta}}\right)$, which establishes (22). Furthermore,

$$\frac{d}{dn}\frac{-cc_{m\theta} + c_\theta c_m}{c_{m\theta}} = \frac{\partial}{\partial m}\frac{-cc_{m\theta} + c_\theta c_m}{c_{m\theta}}\frac{\partial m}{\partial n} = \frac{c_\theta\left(c_{m\theta}^2 - c_{mm\theta}c_m\right)}{c_{m\theta}^2}\frac{\partial m}{\partial n}$$

As shown in Theorem 5, $\frac{\partial m}{\partial n} > 0$ on $[\underline{\theta}, \widehat{\theta}(n))$. Since $c_\theta < 0$ it follows from the assumption $c_{m\theta}^2 - c_{mm\theta}c_m > 0$ that $\frac{\partial H}{\partial n} > 0$ on $[\underline{\theta}, \widehat{\theta}(n))$. Furthermore, on $(\widehat{\theta}(n), \bar{\theta}]$ we have $\frac{\partial m}{\partial n} = 0$, and hence $\frac{\partial H}{\partial n} = 0$. We conclude that $W'(n)$ is strictly decreasing in n i.e., $W(n)$ is strictly concave. *Q.E.D.*

Proof of Lemma 2: By the definition of $\underline{K}$ and $\overline{K}$ we have $\underline{K}c \le \frac{c_m c_\theta}{c_{m\theta}} - c \le \overline{K}c$ and so,

$$0 < \underline{K}\int_{\underline{\theta}}^{\bar{\theta}} c(m(\theta), \theta) f(\theta) d\theta \le \frac{dW(n)}{dn} = G \le \overline{K}\int_{\underline{\theta}}^{\bar{\theta}} c(m(\theta), \theta) f(\theta) d\theta.$$

Q.E.D.

References

Abeler, J., D. Nosenzo and C. Raymond (2016), "Preferences for Truth-Telling," *Centre for Decision Research and Experimental Economics*, Discussion Papers No. 2016-13, School of Economics, University of Nottingham.

Alger, I. and C.A. Ma (2003), "Moral Hazard, Insurance and Some Collusion," *Journal of Economic Behavior and Organization*, 50, pp. 225-247.

Alger, I. and R. Renault (2006), "Screening Ethics when Honest Agents Care about Fairness," *International Economic Review*, 47(1), 59–85.

Alger, I. and R. Renault (2007), "Screening Ethics when Honest Agents Keep their Word," *Economic Theory*, 30(2), pp. 291-311.

Armstrong, M. (1996), "Multiproduct Nonlinear Pricing," *Econometrica*, 64, pp. 51-75.

Arteaga, C. (2017), "The Effect of Human Capital on Earnings: Evidence from a Reform at Colombia's Top University," *Journal of Public Economics*, 157(1), pp. 212-225.

Banks, J. (1990), "A Model of Electoral Competition with Incomplete Information," *Journal of Economic Theory*, 50(2), pp. 309-325.

Ben-Porath, E. and B. Lipman (2012), "Implementation and Partial Provability," *Journal of Economic Theory*, 147(5), pp. 1689-1724.

Bull, J. (2008a), "Costly Evidence Production and the Limits of Verifiability," *The B.E. Journal of Theoretical Economics,* 8(1) (Topics), Article 18.

Bull, J. (2009), "Costly Evidence and Systems of Fact-Finding," *Bulletin of Economic Research,* 61, 103-125.

Bull, J., and Watson, J. (2004), "Evidence Disclosure and Verifiability," *Journal of Economic Theory,* 118, 1-31.

Bull, J., and Watson, J. (2007), "Hard Evidence and Mechanism Design," *Games and Economic Behavior,* 58, 75-93.

Bernheim, B. D., and S. Severinov (2003), "Bequests as Signals: An Explanation for the Equal Division Puzzle," *Journal of Political Economy,* 111(4), 2003, 733-764.

Caillaud, B. and J. Tirole, (2007) "Consensus Building: How to Persuade a Group," *American Economic Review*, 97(5), pp. 1877-1900.

Caskey, J., V. Nagar and P. Petacchi (2010), "Reporting Bias with an Audit Committee," *The Accounting Review*, 85(2), pp. 447-481.

Chevalier, A., C. Harmon, I. Walker and Y. Zhu (2004), "Does Education Raise Productivity, or Just Reflect it?" *Economic Journal,* 114(499), pp. 499-517.

Cho, I.-K. and J. Sobel (1990), "Strategic Stability and Uniqueness in Signaling Games," *Journal of Economic Theory*, 50(2), pp. 381-413.

Clark, D and P. Martorell (2014), "The Signalling Value of a High School Diploma," *Journal of Political Economy,* 122(2), pp. 282-318.

Crocker, K. and J. Morgan (1998), "Is Honesty the Best Policy? Curtailing Insurance Fraud through Optimal Incentive Contracts," *Journal of Political Economy*, 106, 1998, pp. 355-375.

Crocker, K. and J. Slemrod (2005), "Corporate Tax Evasion with Agency Costs," *Journal of Public Economics*, 89(9-10), pp. 1593-1610.

Deneckere, R. and S. Severinov (2008), "Mechanism Design with Partial State Verifiability," *Games and Economic Behavior*, 64, 487-513

Emons, W. and C. Fluet (2009), "Accuracy versus Falsification Costs: The Optimal Amount of Evidence under Different Procedures," *The Journal of Law, Economics, & Organization*, 25(1), pp. 134-156.

Engers, M. (1987), "Signaling with Many Signals," *Econometrica*, 55, pp. 663-674.

Erard, J. and Feinstein, B. (1994), "Honesty and Evasion in the Tax Compliance Game," *Rand Journal of Economics*, 25(1), pp. 1-19.

Fisher, P. and R. Verecchia (2000), "Reporting Bias," *The Accounting Review*, 75(2), pp. 229-245.

Forges, F. and F. Koessler (2005), "Communication Equilibria with Partially Verifiable Types," *Journal of Mathematical Economics*, 41(7), pp. 793–811.

Gneezy, U. (2005), "Deception: The Role of Consequences," *American Economic Review*, 95 (1), pp. 384-394.

Goldman, E. and S. Slezak (2006), "An Equilibrium Model of Incentive Contracts in the Presence of Information Manipulation," *Journal of Financial Economics*, 80, pp. 603-626.

Green, J. and J.-J. Laffont (1986), "Partially Verifiable Information and Mechanism Design," *Review of Economic Studies*, 53, pp. 447-456.

Grossman, S. (1981) "The Informational Role of Warranties and Private Disclosure about Product Quality," *Journal of Law and Economics*, 79(7), pp. 755-759.

Guttman, I., O. Kadan and E. Kandel (2006), "A Rational Expectations Theory of Kinks in Financial Reporting," *The Accounting Review*, 81(4), pp. 811-848.

Hagenbach, J., F. Koessler and E. Perez-Richet (2014), "Certifiable Pre-Play Communication: Full Disclosure," *Econometrica*, 82 (3), pp. 1093-1131.

Hart, S., I. Kremer and M. Perry (2017), "Evidence Games: Truth and Commitment," *American Economic Review*, 107(3), pp. 690-713.

Ireland, N. (1994), "On Limiting the Market for Status Signals," *Journal of Public Economics*, 53(1), pp. 91-110.

Johnstone, R. (1996), "Multiple Displays in Animal Communication: Backup Signals and Multiple Messages," *Philoshical Transactions of the Royal Society B*, 351(1337), pp. 329–338.

Kartik, N., M. Ottaviani and F. Squintani (2007), "Credulity, lies, and costly talk," *Journal of Economic Theory*, 134, 93–116

Kartik, N. (2009), "Strategic Communication with Lying Costs ," Review of Economic Studies, 76, 1359-1395.

Lacker, J.M. and J.A. Weinberg (1989), "Optimal Contracts Under Costly State Falsification," *Journal of Political Economy*, 97(6), pp. 1345-1363.

Lipman, B. and D. Seppi (1995), "Robust Inference in Communication Games with Partial Provability," *Journal of Economic Theory*, 66(2), pp. 370-405.

Maggi, G. and Rodriguez-Clare, A. (1995), "Costly Distortion of Information in Agency Problems," *Rand Journal of Economics*, 26, pp. 675-689.

Mailath, G. (1989), "Simultaneous Signaling in an Oligopoly Model," *The Quarterly Journal of Economics*, 104(2), pp. 417-427.

Maskin, E. and Riley, J. (1984), "Monopoly with Incomplete Information," *Rand Journal of Economics*, 15(2), pp. 171-196.

McNichols, M. and A. Dravid (1990), "Stock Dividends, Stock Splits, and Signaling," *The Journal of Finance*, 45(3), pp. 857-879.

Milgrom, P. (1981) "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12(2), pp. 380-391.

Milgrom, P. and J. Roberts (1992), "Limit Pricing and Entry under Incomplete Information: An Equilibrium Analysis," *Econometrica*, 50(2), pp. 443-459.

Miller, M. and K. Rock (1985), "Dividend Policy Under Asymmetric Information," *Journal of Finance*, 40(4), pp. 1031-51.

Picard, P. (2013), "Economic Analysis of Insurance Fraud," Handbook of Insurance: Second Edition, Georges Dionne (Ed.), Kluwer Academic Publishers, Boston, pp. 349-395.

Page, M. (2010), "Signaling in the Labor Market," in *International Encyclopedia of Education*, Third Edition, Eva Baker, Penelope Peterson and Barry McGaw, Editors, Elsevier Limited, Oxford.

Ramey, G. (1996), "D1 Signaling Equilibria with Multiple Signals and a Continuum of Types," *Journal of Economic Theory*, 69(2), pp. 508-531.

Rochet, J.C. and M. Quinzii (1985), "Multidimensional Signaling," *Journal of Mathematical Economics*, 14, pp. 261-284.

Rochet, J.-C. and P. Choné (1998), "Ironing, Sweeping, and Multidimensional Screening," *Econometrica*, 66, pp. 783-826.

Ross, S. (1977), "The Determination of Financial Structure: The Incentive-Signalling Approach," *Bell Journal of Economics*, 8(1), pp. 23-40.

Seidmann, H. S. and E. Winter (1997), "Strategic Information with Verifiable Messages," *Econometrica*, 65, pp. 163-169.

Seierstad, A. and K. Sydsaeter (1999), "Optimal Control Theory with Economic Applications," North Holland.

Severinov, S. and R. Deneckere (2006), "Screening when not All Agents are Strategic: Does a Monopolist Need to Exclude?" *Rand Journal of Economics*, 37(4), pp. 816-841.

Sher, I. and R. Vohra. (2015), "Price Discrimination Through Communication," *Theoretical Economics*, 10, 597-648.

Spence, A.M. (1973), "Job Market Signaling," *Quarterly Journal of Economics*, 87(3), pp. 355-374.

Spence, A.M. (2002), "Signaling in Retrospect and the Informational Structure of Markets," *American Economic Review*, 92(3), pp. 434-459.

Tennyson, S. (1997), "Economic Institutions and Individual Ethics: Consumer Attitudes Towards Insurance Fraud," *Journal of Economic Behavior and Organization*, 32(2), pp. 247–265.