Coherent beliefs with costless imitative signalling*

Robert G. King, Yang K. Lu, and Ernesto S. Pasten
Boston University

September 2007; revised November 2007

Abstract

There are many economic situtions where it appears relatively costless for one type of agent to send the *signal* that another type would automatically make. It many of these situations, it is also desirable for such a *costless* signal to be issued, i.e., for one type of agent to *imitate* the signal of the other.

In signalling games, it is well understood that there is a typically a rich set of equilibria, with each member being supported by particular out-of-equilibrium beliefs about the information conveyed by messages. We determine this set of equilibria in a simple game of costless imitative signalling: any signal can be an equilibrium, leading to an indeterminate range of economic activity.

We then show how various refinements (plausibility restrictions on out-of-equilibrium beliefs) reduce the set of equilibria. In constrast to well-known signalling games such as Spence's job market model, in which many different refinements deliver similar and intuitive outcomes, the most standard refinements deliver counter-intuitive conclusions in our costless signalling game. The popular *intuitive criterion* of Cho and Kreps as well as the related refinement of *equilibrium dominance* imply that there is only a zero activity equilibrium. The much discussed alternative strategy of Grossman and Perry, which we describe as requiring *weakly coherent* beliefs and actions, yields non-existence. However, when we use the conceptually related approach of Mailath, Okuno-Fujiwara, and Postelwaite, which we describe as requiring *strongly coherent* beliefs and actions, we find that there is there is a single natural and intuitive equilibrium under costless imitative signalling.

*The authors thank Huberto Ennis, Bart Lipman, Leo Martinez, and Antonio Miralles for useful discussions. However, none of these individuals should be held responsible for any erroneous signals sent in this document.

1 Introduction

In many social and economic contexts, some agents automatically send signals that are mechanically related to their underlying characteristics. At the same time, it can be the case that it is virtually costless for agents with other characteristics to choose to send the same signals. If there are rewards to having the characteristics of the first group, then the second group will thus find it desirable to send the same signal.[1]We describe this situation as one of *costless imitative signalling*; we are interested in learning about equilibria in such contexts.[2]

For this purpose, we construct and analyze a very simple model of costless, imitative signalling. In our setup, there is a firm that will produce one unit of a good and will sell it to the customer who is willing to pay the most for it. The quality of the good to be produced is chosen by the firm: it is costly to produce a higher quality. Signalling comes in to play because there are two *types* of firms. The first type automatically provides a public signal m that is the quality that it will actually produce: it is *committed* to its publically observable quality signal. The second type of firm can voluntarily make a public signal m that exactly replicates the quality signal of the first type, but will not carry through on this quality level. Instead, it will simply produce the cheapest possible product (zero quality) and we call this a *discretionary* type of firm.

The essence of our environment is that the discretionary firm can costlessly imitate its committed counterpart in terms of the signal m. Within this setting, we study a sender-receiver game in which there is uncertainty about the type of the agent issuing a signal: the sender knows his type, but other participants in the game do not, although they can update a prior distribution using the observable signal. More specifically, these models specify that a sender of type τ transmits a message m to a group of other players whose rewards depend on the agent's type τ and, for some types, on the message m. On the basis of the message m, the other players form beliefs μ and take actions p. A Bayesian Perfect Equilibrium involves a strategy $\widehat{m}(\tau)$ that is optimal for the senders, best responses by the receivers $\widehat{p}(m, \mu)$ given beliefs over sender type μ, and beliefs $\widehat{\mu}(\tau|m)$ that utilize Bayes' law in equilibrium, but can depart from from it out of equilibrium (for values of m not equal to $\widehat{m}$).

There has been much study of sender-receiver games in which it is costly to send

[1]For example, many interested and intelligent students appear to mechanically sit in the first few row of seats in a a lecture. At the same time, it is perhaps not too costly for other types of students to sit in the front of the class. So, if the lecturer seeks to reward students that are interested and intelligent, using seat location as a signal, it can be desirable for all students to sit in the front of the lecture.

[2]Our interest in this topic was stimulated by thinking about monetary and fiscal settings in which there is signalling by a policy authority of unknown type. In these settings, the idea of costless imitative signalling is a natural reference point. Further, we think that more completely elaborated policy games will likely feature costless imitative signalling in some macroeconomic circumstances and not in others, so we want to have an appropriate analytical framework for exploring such complex, hybrid signalling equilibria.

messages and where the costs differ importantly across types of senders. A basic problem in these models – first brought to the widespread attention of economists by Spence's analysis of job market signalling – is that there can be many equilibria. These equilibria can differ significantly in their observable properties: they can separate agents of type τ, they can pool agents, or they can involve partial pooling and partial separation. There has accordingly been a great deal of research with the objective of finding additional restrictions on these equilibria.

Our simple model is no exception. While there cannot be separating equilibria or partial pooling equilibria because it is costless to imitate, there is a continuum of pooling equilibria: any message m can be an equilibrium, so long as it involves nonnegative profits for both types of firms.

One important strand of the research on "refinements of equilibria" begins with Grossman and Perry [1986]. Seeing an out-of-equilibrium message $m \neq \widehat{m}$, these authors argue that receivers should ask first themselves: "what set of types would have the incentive to send this message?" Then, given that set of types would be issuing this message, receivers should apply Bayes' rule to determine their beliefs. Of necessity, there is a fixed point element to this calculation, because senders cannot evaluate whether a message is beneficial or not without knowing how beliefs will respond to the deviation, but it can be feasible for these out-of-equilibrium beliefs to be consistent with the incentives of senders and with the Bayesian calculations of receivers. Once such out-of equilibrium beliefs induced by the message m are determined, the equilibrium $\widehat{m}$ can be tested to determine whether the required supporting beliefs $\widehat{\mu}(\tau|m)$ are equal to the previously discussed Bayesian inference for $m \neq \widehat{m}$. If not, then, Grossman and Perry argue for rejecting the equilibrium $\widehat{m}, \widehat{p}, \widehat{\mu}(\tau|m)$. In fact, their approach is so strong that it not infrequently leads to the nonexistence of equilibrium. We view the GP approach as invoking the requirement that beliefs be *weakly coherent* with sending agent incentives, as it involves consideration of all messages that might conceivably be sent and a self-referential measure of the associated benefits. In our simple model, the requirement of weakly coherent beliefs is indeed enough to rule out all equilibria.[3]

While our main focus is on refinements that lead to restrictions on admissable beliefs, it is worth pointing out that our model – which has a simple and intuitive equilibrium conjectured by most economists to whom it is described – also leads to

[3] We use the term "coherent" to describe restrictions on beliefs with the properties suggested by Grossman and Perry [1986] and Mailath, Okuna-Fujiwara and Postelwhaite [1993] because other possible terms cause conflict either with terms in our other work or with other established usage in game theory. In our monetary policy analysis, the idea of credibility plays a key role and we have several model-based definitions that we seek to stress. Hence, we do not follow GP in defining these as credible beliefs. We also considered the use of "consistent", but that terminology is already a key part of the definition of sequential equilibrium provided by Kreps and Wilson [1982]. The Merriam-Webster primary definition of coeherent is (a) logically or aesthetically ordered or integrated (CONSISTENT); and (b) having clarity or intelligibility (UNDERSTANDABLE). We think both versions of this primary definition make the use of this term appropriate in discussion of beliefs.

an unorthodox outcome when we use other standard refinements, such as equilibrium dominance and the intuitive criterion of Cho and Kreps [1987]. Those alternative refinements lead to the conclusion that the only outcome in our model is that all firms produce zero quality, which is not what most observers would predict as the outcome of the model relevant for understanding actual economies.

Yet, when considering the answer to the important question "what type of agent would have the incentive to send the message m?", Grossman and Perry do not pose the complementary question, "if the message m is not sent, then what should be inferred about the type of agent that is present?" When one thinks about it, the answer to the second question can be just as informative from a Bayesian standpoint as the answer to the first. Mailath, Okuno-Fujiwara, and Postelwaite [1993] develop this line of arguement further, arguing that its logical endpoint is to restrict any out-of-equilibrium message m to the set of messages that could be sent in *some* signalling equilibrium and to make the evaluations of "incentive to send" and "incentive not to send" based on incentives within alternative equilibria. We call this the requirement of *strongly coherent* beliefs.

In our setting, since any message can be sent in some Bayesian perfect equilibrium, the range of out-of-equilibrium messages considered is not the central element of the comparison of the two belief-based approaches just discussed. Instead, Grossman and Perry allow for so many coherent out-of-equilibrium beliefs by imposing such weak requirements that there is always some way to break down any equilibrium. By contrast, the stronger coherency requirements of MOP imply that there are fewer coherent out-of-equilibrium messages and beliefs, but these restrictions are sufficient to select a single equilibrium. Further, that equilibrium turns out to the be the single "natural" equilibrium that most observors would predict.

After studying a basic model with two types, we then extend the analysis to think about models in which the distinction between commitment and discretion is not absolute, but involve consideration of costs and benefits from deviating from an announced plan. We study three type and continuous type models, concluding that the requirement of strongly coherent beliefs leads a unique and natural equilibrium in each setting.

2 Setup

Our model is designed to be in the form of the canonical sender-receiver game, but featuring costless, imitative signals. We suppose that there is a firm which will produce one unit of a good whose quality q is variable. The type τ of the firm is unknown to the consumer, who will pay a price based on the expected quality level. Messages m about product quality can affect the price that the customer will pay.

2.1 Firm types, messages, and actions

We initially suppose that there are two types of firms: each will issue a message m and then subsequently take an action q. A committed firm (type $\tau = c$ firm) has the constraint $q = m$. A non-committed firm (type $\tau = d$ firm) has no such constraint and will simply select q so as to maximize its ex post profits, as described further below.

The type $\tau = c$ firm chooses the quality of the good, q, with

$$0 \leq q \leq Q$$

Since it is bounded to set $q = m$, we assume that this is also the range of its messages. It is convenient to write $m \in M$ when denoting messages chosen from this message space.

The type $\tau = d$ firm chooses a message from the same space. It has no restriction on the quality level that it can produce.

Quality is costly to produce, with the cost function $\psi(q)$ being convex in quality, with $\psi(0) = 0$. [4] Given that quality is costly to produce, a firm of type $\tau = d$ will always produce a quality of zero.

The firm type is selected randomly by nature, with probability $\rho(\tau)$ of type τ. Firm type is private information. However, the message issued by the firm, m, can affect the demander's probability $\mu(\tau|m)$ of a type τ firm being in place.

2.2 Consumers

The good will be sold to a consumer who is willing to pay more for higher expected quality,

$$p = \lambda[\mu_c q_c + \mu_d q_d] \tag{1}$$

where μ_τ is short-hand for the consumer's probability that the firm is of type τ and q_τ is the quality level chosen by a firm of type τ. The parameter $\lambda > 0$ indicates the value of the good to the consumer, as discused further below.

As in textbook discussions of the classic Spence [1974] model of job market signalling, where wages are just the expected value of productivities, we do not spell out the details of how this price is established. However, to state matters in more careful game-theoretic terms, we could posit that there are two consumers A and B who play a Nash game in which each offers a price, with the highest price offer getting the good. Under the assumption that the two consumers are risk neutral and that each values the good at λ, then the above pricing relation describes their common equilibrium offer strategies.

[4] In drawing diagrams, we use the specific form $\psi(q) = \frac{1}{2}q^2$: the specific functional form is for convenience: all that matters, we think, is that there is increasing marginal cost.

2.3 Timing

The time line of events in the model economy is as shown in Table 1. At the start, nature draws firm type with a probability $\rho(\tau)$, which we sometimes write as ρ_c or ρ_d for convenience. Next, the firm issues the message m. Consumers form beliefs μ and offer a price p in the middle, as determined according to (1). At the end, firms decide on profit maximizing production, with $q = m$ if $\tau = c$ and $q = 0$ if $\tau = d$.

Table 1				
period	start	beginning	middle	end
actor	nature	firm	consumer	firm
action	firm type: τ	signal: $m \in M$	beliefs:μ price: p	$q = m$ if $\tau = c$ $q = 0$ if $\tau = d$

One could add a final stage in which consumers receive payoffs of $\lambda q - p$, but we follow the standard approach in signalling analysis by downplaying that aspect of the economy.

3 A reference analysis

Looking at this economy, we think that it is natural to think that there would be a unique equilibrium of a particularly simple form. In the middle subperiod, consumers recognize that there are good ($q_c > 0$) and bad ($q_d = 0$) producers, so that they offer a price

$$p = \lambda \rho_c q_c$$

where ρ_c is the population probability of committed producers. Relative to a situation where all producers were commited ($\rho_c = 1$), consumers will offer less for the good. Under incomplete information, they are still willing to pay more for higher quality, but recognize that they might just get a worthless product.

Further, given this pricing behavior, committed producers select a profit-maximizing quality

$$q^* = \arg\max_q \{\lambda \rho_c q - \psi(q)\}$$

or, equivalently, they set marginal revenue equal to marginal cost,

$$\lambda \rho_c = \frac{\partial \psi}{\partial q}(q^*)$$

and generate profits

$$\pi_c^* = p^* - \psi(q^*)$$

where $p^* = \lambda\rho_c q^*$. In terms of messages, the commited firm ($\tau = c$) automatically sends the message $m_c = q^*$ as a by-product of its decision to produce.

The discretionary firm earns profits of

$$\pi_d^* = p^* = \lambda\rho_c q^*$$

given that it does not produce costly quality. In terms of its message, it sets $m_d = q^*$ to avoid being distinguished from the committed firm.

By its construction, this equilibrium maximizes the profit of the committed firm, subject to the constraints that (a) it cannot distinguish itself from the discretionary firm; and (b) it must sell its product to consumers that will pay according to their expected quality. We accordingly refer to this as the "optimal equilibrium" in our discussion below.

While this may seem to be a natural equilibrium outcome, our analysis shows that it is not a special one from the perspective of signalling theory. First, it is only one of many Bayesian perfect equilibria: any quality can be an equilibrium if consumers have particular beliefs about firm type. Second, standard refinements – aimed at eliminating some members of that equilibrium set – do not identify this equilibrium. Instead, equilibrium dominance (as described in Green, MasCollel and Winston [1995]) and the intuitive criterion of Cho and Kreps [1990] select $q = 0$. The perfection approach of Grossman and Perry [1986] implies that there are no equilibria at all.

4 The market game and its equilibrium set

The market game unfolds in five steps, as shown in Table 1 above.

First, nature draws a type of firm with the probability of τ being $\rho(\tau)$.

Second, firms issue a message $m \in M$. These messages can potentially affect beliefs. The committed firm's message constrains its actual production behavior, but there is no such constraint for the non-committed firm.

Third, the price is set based on the formula above, so that the price depends on the beliefs that customers have

$$p = \lambda[\mu(\tau = c|m)q(c) + \mu(\tau = d|m)q(d)]$$

where we now return to our more complete notation for beliefs and actions.

Fourth, firms choose a profit-maximizing quality level, given the price. (The quality level that they will choose will depend on the beliefs that customers have, for these will affect the price). For the committed firm in the end of the subperiod, its profits are just

$$\pi(m, p, \tau = c) = p - \psi(m)$$

and there is no decision to be taken. However, for the non-committed firm, it is the case that

$$\pi(q, p, \tau = d) = p - \psi(q)$$

so that the profit-maximizing quality is $q = 0$ and the non-committed firm's profits are just p. This allows us to simply write

$$p = \lambda[\mu(\tau = c|m)q(c)]$$

in the analysis below. Note also that it lets us focus narrowly on the belief that the firm is of a committed type.

4.1 Bayesian Perfect (pooling) equilibria

For this game, there are many Bayesian Perfect Equilibria of a pooling type. Since there is no outside option ($\pi \geq 0$ is the relevant participation constraint), then any $0 \leq \widehat{q} \leq x$ can be BPE with x defined by $\rho_c \lambda x - \psi(x) = 0$. The committed firm would not participate in a pooling equilibrium for higher values of q.

To find the beliefs that can support any chosen equilibrium quality in this range, it is useful to employ diagrams that are drawn in (p, q) or equivalently (p, m) space. There are four ingredients of these diagrams that we will use repeatedly:

- A pricing function applicable if agents believe that the firm is of type $\tau = c$

$$p = \lambda q$$

- A pricing function applicable if agents believe that the firm is of type $\tau = c$ with probability ρ_c.

$$p = \lambda \rho_c q$$

- An isoprofit curve for a firm of type $\tau = c$, giving the price p at production level q which generates the same level of profits as a pooled equilibrium quality e.

$$C(q, e) = \pi(e) + \psi(q) = [\lambda \rho_c e - \psi(e)] + \psi(q)$$

- An isoprofit curve for a firm of type $\tau = d$,

$$D(e) = \pi(e) = \lambda \rho_c e$$

These constructions are displayed in Figure 1: the dotted line is the pricing function, $p = \lambda q$, which is applicable if agents believe that the firm is of type $\tau = c$; the dashed line is the pricing function $p = \lambda \rho_c q$ applicable if agents believe that the firm is of type $\tau = c$ with probability ρ_c; the heavy solid line is the isoprofit curve $C(q, e)$ for a firm of type $\tau = c$ with the point e being the crossing point with the **dashed** line; and the light solid (horizontal) line is the isoprofit line for a firm of type $\tau = d$.

Throughout our analysis, we will make use of the fact that any price, quality pair on or below $p = \lambda q$ can be equivalently viewed as a probability, quality pair since

$$p = \lambda \mu_c q$$

implies that

$$\mu_c = \frac{p}{\lambda q}$$

That is, a particular price reflects a consumer probability of the firm being of the committed type.

We will repeatedly go back and forth between the price and probability interpretation for our problem, as each seems a useful form source of information.

4.2 The optimal equilibrium

Figure 2 displays the optimal pooling equilibrium described in section 2 above. It involves a tangency between an isoprofit function for the type $\tau = c$ firm and the pooled equilibrium price function. The equilibrium price is $p^* = \lambda \rho_c q^*$; the associated profit levels for the two types of firms are

$$\begin{aligned} \pi_c^* &= p^* - \psi(q^*) \\ \pi_d^* &= p^* \end{aligned}$$

Referring to this diagram, we can ask the question: what type of price functions

$$p(q, q^*)$$

would lead the two types of firms to uniquely select $q = q^*$? The answer is a direct one: any price function that lies below both firms' isoprofit functions at all points other than q^*. (In the figure, this is any price function that lies below the dark line). Mathematically, the price function must satisfy: (a) $p(q, q^*) < C(q, q^*)$ for $q < q^*$; (b) $p(q, q^*) = C(q, q^*) = D(q^*)$ for $q = q^*$; and (c) $p(q, q^*) < D(q^*)$ for $q > q^*$. The first portion of this price function, for $q < q^*$, assures that it is not profitable for type $\tau = c$ firms to choose low quality. The third portion assures that it is not profitable for type $\tau = d$ firms to choose high quality. We can thus evidently posit in many price functions which "support" this optimal equilibrium.

Equivalently, there are many off-equilibrium beliefs,

$$\mu(\tau = c|m) = \frac{p(m, q^*)}{\lambda m},$$

that support $m_c = m_d = q^*$ as a Bayesian perfect equilibrium. Note that we have made the switch to using m in this condition, because we are now explicitly thinking about the signalling game, while the above discussion of supporting prices might have described some other market outcome. As discussed further below, the equilibrium q^* is supported as an equilibrium by the specified belief function as follows: both firms think that messages $m \neq m^*$ lead to consumers being sufficiently pessimistic (low $\mu(\tau = c|m)$, low $p(m, q^*)$) about the likelihood of a firm being committed, so that it is profit-maximizing to send the message $m = m^*$.

4.3 Supporting any quality level as an equilibrium

The same argument can be used to support *any* quality level as a Bayesian perfect equilibrium of the pooling type, so long as it is not unprofitable for the committed firm (q is less than x or $m \in M$ in terms of the discussion above).

It is useful to refer to Figure 3 for this purpose, which displays a candidate equilibrium $\widehat{q} < q^*$ in terms of price and quality. Any price function which lies below the heavy solid line (the isoprofit for $\tau = c$) for $q < \widehat{q}$, which has $p = \widehat{p} = \lambda\rho_c\widehat{q} - \psi(\widehat{q})$ at $q = \widehat{q}$ and has $p < \widehat{p}$ for $q > \widehat{q}$ will support this candidate equilibrium.

We can make a similar argument for any candidate equilibrium $\widehat{q}$, including those $\widehat{q} > q^*$.

4.3.1 Messages and beliefs

Equivalently, there are out-of-equilibrium beliefs that will support the candidate equilibrium, which we can construct as follows. Let $p(q; \widehat{q})$ be a price function that supports $\widehat{q}$.

Then, if beliefs are linked to messages by

$$\widehat{\mu}(\tau = c|m) = \frac{p(m; \widehat{q})}{\lambda m}$$

then these will support an equilibrium with

$$\widehat{m}_c = \widehat{m}_d = \widehat{q}$$

as the equilibrium messages and production quality by the committed firm.

4.3.2 Restriction to pooling equilibria

Our discussion does consider separating equilibria, as there are none in this market game. No type $\tau = d$ would signal its type and receive a zero price, when it could alternatively issue the same message as that of a type $\tau = c$ firm without any additional cost and receive a positive price.

4.4 An inverse Bayesian interpretation

Suppose that we consider a particular message m and a particular candidate pooling equilibrium $\widehat{q}$ with an associated price function $p(q, \widehat{q})$. We can reformulate the constraints on prices, $p < C(q, \widehat{q})$ for committed firms and $p < D(q) = \widehat{p}$ for discretionary firms in terms of beliefs, using $p = \lambda\mu_c m$, as

$$\mu(\tau = c|m) < \frac{C(m,\widehat{q})}{\lambda m}$$
$$\mu(\tau = c|m) < \frac{D(\widehat{q})}{\lambda m}$$

Further, writing Bayes' law as

$$prob(\tau = c|m) = \frac{prob(m|\tau = c)}{prob(m|\tau = c)\rho_c + prob(m|\tau = d)\rho_d}\rho_c,$$

we can see that this belief in turn depends on the probabilities that each type of firm would send the specified message.

Return to Figure 3 and look at the candidate equilibrium $\widehat{q} < q^*$ and consider a messsage m that is marginally larger than $\widehat{q}$. It must be that the price at m is lower[5] than $\lambda\rho_c m$ for the firms of type $\tau = c$ to find it optimal to send $\widehat{q}$ rather than m, i.e., for $\widehat{q}$ to be an equilibrium: they would earn higher profits in the pooled equilibrium $m > \widehat{q}$. In turn, this requires $\mu_c < \rho_c$ in terms of beliefs at message m. Finally, looking at Bayes' rule, it must be the case that the message m is viewed as being sent more frequently by type $\tau = d$ firms than by type $\tau = c$ firms: $\mu_c = prob(\tau = c|m) < \rho_c$ implies that $prob(m|\tau = c) < prob(m|\tau = c)\rho_c + prob(m|\tau = d)\rho_d$ and thus that $prob(m|\tau = c) < prob(m|\tau = d)$ since $\rho_c + \rho_d = 1$ That is, from a Bayesian perspective, the specification of off-equilibrium beliefs is equivalent to a specification of off-equilibrium message-sending behavior.

5 Disciplining off-equilibrium beliefs

To sharpen the predictions of the model, one is therefore led to refinements. We find developments initiated by Grossman and Perry [1986] particularly attractive in overall design. Such approaches discipline the structure of off-equilibrium beliefs in our context, relative to BPE, by applying a combination of economic and Bayesian reasoning: it requires that beliefs be coherent with sender incentives and statistical inference. However, it turns out that the Grossman-Perry approach provides too much discipline in our setting: it rules out all equilibria!

For a message m and an Bayesian Perfect Equilibrium $\widehat{q}$, GP argue that an *updating function*

$$\mu_c = \mu(\tau = c|m) = \gamma_c(m, \rho_c; \widehat{q})$$

[5] **Since** $\hat{q} < m < q^*$**,** $\lambda\rho_c m - \psi(m) > \lambda\rho_c\hat{q} - \psi(\hat{q})$**. Hence, the condition** $p(m) - \psi(m) < \lambda\rho_c\hat{q} - \psi(\hat{q})$**, which is necessary to support pooling equlibrium** $\hat{q}$**, implies that** $p(m) - \psi(m) < \lambda\rho_c m - \psi(m)$**. Therefore, the necessary condition to support pooling equlibrium** $\hat{q}$ **is** $p(m) < \lambda\rho_c m$**, equivalently**

$$\mu_c < \rho_c$$

should be used to describe beliefs, with the function γ restricted as follows in our simple model.

1. In the equilibrium $m = \widehat{q}$, this function is simply Bayes's law.

2. Out of equilibrium, the function takes on the value: $\mu_c = 1$ if it would be profitable for m to be sent just by type $\tau = c$ firms; $\mu_c = 0$ if it would be profitable for m to be sent by only type $\tau = d$ firms; $\mu_c = \rho_c$ if it would be profitable for both types of firms to send the message.

   In these evaluations of profits, Grossman and Perry assume that firms correctly understand the implications of out-of-equilibrium beliefs for prices: in terms of sending the message, they act with a clear understanding how the message will be interpreted. Equivalently, there is a fixed point in terms of out-of-equilibrium beliefs: firms understand the signalling content of their messages and the inferences that will be drawn from them, so that they evaluate their profits in terms of revised beliefs.

3. This function may not be restricted in form – at a given message and candidate equilibrium – if it is not possible to produce an internally consistent set of beliefs and actions, as illustrated further below. That is, if the updating function is not restricted, then a researcher is free to specify beliefs as in the BPE analysis above.

5.1 Low quality candidate equilibria

We now show that GP restrictions on out-of-equilibrium beliefs rule out low quality equilibria, i.e., $\widehat{q} = l < q^*$ Define h as the largest quality level such that

$$\lambda\rho_c q - \psi(q) \geq \lambda\rho_c l - \psi(l)$$

In words, this is largest quality level such that a type $\tau = c$ firm earns profits that are not smaller than at the candidate equilibrium l.

Consider messages

$$l < m < h$$

If these messages are treated as implying $\mu_c = \rho_c$, then they will be profitable for both firms

$$\begin{aligned} \tau &= c : \lambda\rho_c m - \psi(m) > \lambda\rho_c l - \psi(l) \\ \tau &= d : \lambda\rho_c m > \lambda\rho_c l \end{aligned}$$

with the first line indicating that m is profitable for the committed firm and the second line that m is profitable for the non-committed firm.

Hence, according to GP, it follows that updating has the form

$$\mu_c = \gamma_c(m, \rho_c) = \rho_c \text{ for } l < m < h$$

The key observation, then, is that such a coherent belief is inconsistent with the BPE requirement for the support of l. That is, as explained at the end of the prior section, $m > \widehat{q}$ requires that the probability of a $\tau = c$ given m is less than ρ_c, while the above reasoning indicates that this is inconsistent with coherent beliefs in the sense of GP, i.e., beliefs that are consistent with the incentives that individuals would have to send out-of-equilibrium messages taking into account the information that those messages would convey.

We again use Figure 3 to better understand that such a message must always exist and that it is inconsisent with the BPE off-equilibrium requirement on beliefs. Any quality $m > l$ which leads to a price p that lies above the two isoprofit lines is one that both firms would like: they would both issue message m, $l < m < h$, if they believed that these messages would result in such a price p. Further, receiving such a mesage, a consumer would think that it could reasonably come from either firm so that Bayesian formation of beliefs would set $\mu_c = \rho_c$, thus rationalizing the message. By contrast, the supporting of l as a BPE requires a price lower than $p_c = \lambda\rho_c m$ or a belief that is more pessimistic than a Bayesian one, $\mu_c < \rho_c$. Hence, $l < q^*$ cannot be an equilibrium under the GP refinement because it requires off-equilibrium beliefs that are not coherent with sender incentives.

5.2 High quality candidate equilibria.

Consider next a candidate equilibrium $\widehat{q} = h > q^*$. Consider messages

$$l < m < h$$

where l is the lowest message such that $\lambda\rho_c m - \psi(m) \geq \lambda\rho_c h - \psi(h)$. These messages are ones that it would be rational for a type $\tau = c$ firm to send if they were interpreted as implying $\mu_c = \rho_c$. Looking Figure 4, the candidate equilibrium is now the upper intersection of the type $\tau = c$ firm isoprofit line with the pooled equilibrium price function: any message m above the lower intersection (l) would be desirable for the committed firm.

At the same time, if these messages were interpreted as implying $\mu_c = \rho_c$, then these would be *undesirable* from the standpoint of type $\tau = d$ firms. That is, for messages in the range,

$$\begin{aligned} \tau &= c : \lambda\rho_c m - \psi(m) > \lambda\rho_c h - \psi(h) \\ \tau &= d : \lambda\rho_c m < \lambda\rho_c h \end{aligned}$$

because they lie below the type 0 firm's isoprofit line $p = \lambda\rho_c h$. Hence, they cannot imply $\mu_c = \rho_c$ using the GP logic.

If these messages are interpreted as implying $\mu_c = 1$, then matters are more complicated. Firms of type $\tau = c$ always benefit from such messages.

$$\tau = c : \lambda m - \psi(m) > \lambda \rho_c m - \psi(m) > \lambda \rho_c h - \psi(h)$$

However, firms of type $\tau = d$ benefit from some and not from others. Define δ as that message at which the type $\tau = d$ firm is just indifferent between sending it if he will be treated as a type $\tau = c$ firm (prob$(\tau = c|\delta) = 1$) and not sending it.

$$\lambda \text{prob}(\tau = c|\delta)\delta = \lambda \rho_c h \Rightarrow \delta = \rho_c h$$

Then, for $\delta < m < h$, both firms would have higher profits with m than with h. But this is not consistent with the assumption that $\mu_c = 1$. So, according with condition 3 in restrictions for coherency, beliefs are *not restricted* on this range.

However, with $l < m < \delta$, type $\tau = c$ firms are better off while type $\tau = d$ firms are worse off. Hence, on this range beliefs are restricted to $\mu_c = 1$. This is inconsistent with the off-equilibrium belief requirements of the BPE $\widehat{q} = h$: the price $p = \lambda m$ lies above the indifference curve for the type $\tau = c$ firm or, equivalently, the belief $\mu_c = 1$ is greater than the largest supporting belief

$$\mu_c < \rho_c$$

so that a candidate equilibrum $h > q^*$ is not sustained under the GP refinement.

5.3 Optimal quality candidate equilibria

We have thus seen that there cannot be any other equilibrium besides q^*. The difficulty is that optimal quality equilibria are vulnerable to the same line of argument that was just made above: it is always possible to find a message that will break q^*, so that it also is ruled out.

This is illustrated in Figure 5. Firms of type $\tau = c$ signal that they are committed by announcing a low quality in the indicated range. The beliefs at that signal are (weakly) coherent because only firms of type $\tau = c$ will send the message, since firms of type $\tau = d$ would earn lower profits by doing so (the price $p = \lambda q$ is below $p^* = \lambda \rho_c q^*$). So, there is no equilibrium with coherent beliefs along the lines advocated by Grossman and Perry [1986].

5.4 A digression on other refinements

The weakly coherent beliefs approach of Grossman and Perry [1986] was just one of a number of refinements that were developed in the early 1980s and that are now the subject of textbook material in microeconomic theory courses.[6] Cho and Kreps

[6] See, for example, chapter 3 of Kreps [1990] and chapter 13 of Green, MasCollel, Whinston [1995].

[1987] highlight two of these refinements, equilibrium domination and the intuitive criterion.

As in the GP refinement, these approaches involve developing restrictions on plausible beliefs so as to rule out some members of the large set of perfect Bayesian equilibria (or sequential equilibria).

Equilibrium domination: In this case, a given equilibrium candidate survives this criterion if there is no alternative message (in our illustration, announced quality) which gives higher return under its "best possible equilibrium interpretation" to the "committed" type. That is, the non-committed type must not also issue the message so that sender type can be unambiguously identified. As above, the idea is to restrict beliefs using the rational reaction of consumers who, expecting a signal m, observe a signal $\widehat{m}$. If the answer is that only the committed type would issue $\widehat{m}$, then beliefs that support a pooling equilibrium for signaled quality m are not plausible.

In our quality game, an arbitrary equilibrium m satisfies the equilibrium dominance criterion if there is no $\widehat{m}$ such that

$$\begin{aligned} \tau &= c : \lambda\rho_c m - \psi(m) < \lambda\widehat{m} - \psi(\widehat{m}) \\ \tau &= d : \lambda\rho_c m > \lambda\widehat{m} \end{aligned}$$

Using this criterion, the only pooling equilibrium which is not ruled out by this criterion is $m = 0$. To see why, notice that for any proposed equilibrium $m > 0$ (including the best pooling for the committed type m^*), we can find an alternative message $\widehat{m} \in (l, h)$ which is preferred by the committed type if doing so its type is identified, where l and h are the lower and higher messages satisfying the first condition with equality. In the Figure 6, this condition is satisfied by any message higher than the crossing between the dotted line (price if $\rho_c = 1$) and the isoprofit relevant for the candidate pooling equilibrium and lower than the higher crossing between these two functions **(alter Figure 6 to show this point).**

The second condition – the requirement that it is desirable for the discretionary type not to send the message $\widehat{m}$ rather than m – is satisfied only by messages $\widehat{m} \in (l, m)$, thus we restrict the potential deviations to that range. If any of those messages is observed, consumers realize that the non-committed firm has no incentives to issue that signal, since the price that can obtain is lower than the one obtained in the candidate equilibrium. Thus, such a deviation would be interpreted as issued by the committed type, generating an inconsistency on off-equilibrium beliefs that support the candidate equilibrium. Thus, any candidate equilibrium with $m > 0$ cannot be supported by beliefs, if equilibrium domination is the plausibility requirement.

In the case of the candidate equilibrium $m = 0$, both types have incentives to deviate to any alternative higher message, and therefore off-equilibrium beliefs cannot be restricted. That is, we cannot rule out the off-equilibrium belief that $\mu(\tau = c|m) = 0$ for any $m > 0$ which would support the equilibrium $m = 0$.

The intuitive criterion of Cho and Kreps works using a similar logic. A candidate PBE violates this criterion when (i) there exists an alternative message which has lower return than the proposed equilibrium for the non-committed type (as with $\widehat{m}$ above); and (ii) the minimum equilibrium payoff of that message for the committed type, taking into account the restrictions in off-equilibrium beliefs implied by (i), is higher that the payoff obtained in the candidate equilibrium (also as with $\widehat{m}$ above).

In the case with only two types, these equilibrium domination and intuitive criterion amount to equivalent restrictions.[7]Thus, the equilibrium $m = 0$ is the only one consistent with the intuitive creiterion.

5.5 A formal definition of weakly coherent beliefs

It is useful to provide a formal definition of weakly coherent beliefs and a specification of the test provided by Grossman and Perry [1986], before turning to other related notions.

> **Definition of weakly coherent beliefs**. For an economy with a set of sender types T and a candidate equilibrium $\widehat{m}, \widehat{p}, \widehat{\mu}(\tau|m)$, an out-of-equilibrium message m gives rise to a weakly coherent out-of-equilibrim belief $\gamma(\tau|m)$ about a subset of types if:
>
> (D.1.1) for all τ, $\widehat{m}(\tau) \neq m$;
>
> (D.1.2) there is a non empty set K such that for all $\tau \in K$,
>
> $$\pi(m, p(m, \gamma(\tau|m)), \tau) \geq \pi(\widehat{m}, \widehat{p}, \tau)$$
>
> and for some $\tau \in K$**, [YANG: is this strict inequality required in GP?]**
>
> $$\pi(m, p(m, \gamma(\tau|m)), \tau) > \pi(\widehat{m}, \widehat{p}, \tau);$$
>
> (D.1.3) for any $\widetilde{\tau} \in K$,
>
> $$\gamma(\widetilde{\tau}|m) = \frac{\rho(\widetilde{\tau})\theta(\widetilde{\tau})}{\sum_{\tau \in T} \rho(\tau)\theta(\tau)}$$
>
> where the function $\theta(\tau)$ specifies the probability that an agent of type τ would make the signal, so that it satifies
>
> $$\begin{aligned} \theta(\tau) &= 1 \text{ for all } \tau \in K \text{ such that } \pi(m, p(m, \gamma(\tau|m), \tau) > \pi(\widehat{m}, \widehat{p}, \tau) \\ \theta(\tau) &= [0,1] \text{ for all } \tau \in K \text{ such that } \pi(m, p(m, \gamma(\tau|m), \tau) \geq \pi(\widehat{m}, \widehat{p}, \tau) \\ \theta(\tau) &= 0 \text{ for all } \tau \notin K \end{aligned}$$

[7] However, where there are more than two types, the condition (i) is modified such that there must exist an alternative message where *at least for one type it is not optimal to follow.* Then, there may be still a "partial pooling" under the case that the alternative message is issued, and thus many options for payoffs depending on off-equilibrium beliefs. The intuitive criterion dictates that the minimum of those payoffs should be used to check if the candidate equilibrium survives this criterion (in our quality game, assigning probability 0 to the committed type), while the equilibrium dominance forms its rule using the maximum of those payoffs (in our quality game, assigning probability 1 to the committed type.

This definition includes the specific structure above, but is compatible with multiple types, some of which can be indifferent: we introduce this generality because we will study such an economy below.

The Grossman and Perry [1986] refinement is then described as follows.

> **Grossman-Perry Refinement**: A Bayesian perfect equilibrium $\widehat{m}, \widehat{p}, \widehat{\mu}(\tau|m)$ is not perfect in the sense of Grossman and Perry [1986], if there is a message m and a type $\widetilde{\tau}$ such that there exists one or more weakly coherent beliefs, $\gamma(\widetilde{\tau}|m)$, and that
>
> $$\widehat{\mu}(\widetilde{\tau}|m) \neq \gamma(\widetilde{\tau}|m).$$
>
> for any such candidate equililbrium beliefs.

Essentially, then, this refinement says that a candidate equilibrium is eliminated if it cannot be supported by a weakly coherent out-of-equilibrium belief, when such a belief exists. Notice that a weakly coherent out-of-equilibrium belief about the type sending a particular message need not exist, in which case no restriction is placed on beliefs.

6 Strongly coherent out-of-equilibrium beliefs

The previous section studied restrictions on beliefs that stem from asking the question: "if the out-of-equilibrium message m is sent, then who would have the incentive to send it?" However, there is an important tension, which can be best understood by thinking about the message deviation that destroyed our candidate equilibrium $\widehat{m} > q^*$. Faced with a suitable designed message $m < \widehat{m}$, consumers thought "Aha!: it is the committed firm because no discretionary firm would send a signal that it wanted to reduce quality". But, if there was a discretionary firm in place and it did not send the message m then it is also plausible for customers to think "Aha!: it is the discretionary firm, so let's not buy the product". And, anticipating that reaction, the discretionary firm should send the message m, thus making the first inference – the one that was so critical above – invalid. One thus enters a wilderness of mirrors, in which there is no evident link between belief and type.[8]

Mailath, Okuno-Fujiwara, and Postelwaite [1993] note that a key property of Bayesian Nash equilibria is that there is a consistent interpretation of messages on the equilibrium path and that systematic application of the sort of inductive reasoning

[8] "Wilderness of mirrors" is a phrase from T.S.Eliot's poem "Gerontion". It was used by by former the CIA counterintelligence chief James Jesus Angleton to describe the world of espionage and, in particular, the process of ferreting out double agents. It is featured in the title of a book on Angleton and the CIA by David Martin.

described above inevitably leads to the view that messages should be evaluated vis-a-vis the rewards in an equilibrium in which they would actually be sent.

In the spirit of MOP, then, we provide a definition of a strongly coherent out-of-equilibrium belief.

Strongly coherent beliefs. For an economy with a set of sender types T, a candidate equilibrium $\widehat{m}, \widehat{p}, \widehat{\mu}(\tau|m)$ and an alternative equilibrium $m', p', \mu'(\tau|m)$, then the message m gives rise to a strongly coherent out-of-equilibrium belief $\gamma(\tau|m)$ about a subset of types if

(D.2.1) for all τ, $\widehat{m}(\tau) \neq m$, there is a non empty set $K = \{\tau \in T|m'(\tau) = m\}$.

(D.2.2) for all $\tau \in K$,

$$\pi(m', p', \tau) \geq \pi(\widehat{m}, \widehat{p}, \tau)$$

and for some $\tau \in K$,

$$\pi(m', p', \tau) > \pi(\widehat{m}, \widehat{p}, \tau);$$

(D.2.3) for any $\widetilde{\tau} \in K$,

$$\gamma(\widetilde{\tau}|m) = \frac{\rho(\widetilde{\tau})\theta(\widetilde{\tau})}{\sum_{\tau \in T} \rho(\tau)\theta(\tau)}$$

where the function $\theta(\tau)$ specifies the probability that an agent of type τ would issue the message, so that it satisfies

$$\begin{aligned}
\theta(\tau) &= 1 \text{ for all } \tau \in K \text{ such that } \pi(m', p', \tau) > \pi(\widehat{m}, \widehat{p}, \tau) \\
\theta(\tau) &= [0, 1] \text{ for all } \tau \in K \text{ such that } \pi(m', p', \tau) \geq \pi(\widehat{m}, \widehat{p}, \tau) \\
\theta(\tau) &= 0 \text{ for all } \tau \notin K
\end{aligned}$$

In words, the requirements of this definition are: (1) that the message m under consideration is one that would not be sent in the candidate equilibrium, but would actually be sent in an alternative equilibrium by some type or types of senders; (2) that some senders of the message m in the alternative equilibrium would benefit from sending it, relative to the candidate equilibrium message and that others would not be harmed by doing so; and (3) that Bayes' law is employed to calculate the out-of-equilibrium belief.

There are several aspects of this definition of strongly coherent out-of-equilibrium beliefs which differ from the prior definition of weakly coherent beliefs. First, a weakly coherent belief can potentially be formed for an arbitrary message m, while a strongly coherent belief can only be formed for a message that is issued in *some alternative* equilibrium. Second, a weakly coherent belief $\gamma(\tau|m)$ requires that the sender weakly benefits from the message m when he evaluates his rewards at this new belief,

$$\pi(m, p(m, \gamma(\tau|m)), \tau) \geq \pi(\widehat{m}, \widehat{p}, \tau)$$

while a strongly coherent belief requires that the sender weakly benefits if he evaulates his rewards at the belief in an alternative equilibrium,

$$\pi(m', p', \tau) \geq \pi(\widehat{m}, \widehat{p}, \tau)$$

With this defintion, we can then state the refinement of Mailath et. al as follows,

> **Mailath, Okuno-Fujiwara, and Postelwaite Refinement**: A Bayesian perfect equilibrium $\widehat{m}, \widehat{p}, \widehat{\mu}(\tau|m)$ is *defeated* by an alternative equilibrium $m', p', \mu'(\tau|m)$ if there exists a type $\widetilde{\tau}$ and one or more strongly coherent belief on message m', $\gamma(\widetilde{\tau}|m')$, and that
>
> $$\widehat{\mu}(\widetilde{\tau}|m') \neq \gamma(\widetilde{\tau}|m').$$
>
> for any such candidate equilibrium belief.

Essentially, then, this refinement says that the candidate equilibrium is eliminated if it cannot be supported by a strongly coherent out-of-equilibrium belief, when such a belief exists.

7 The unique imitative signalling equilibrium

We now show that there is a unique equilibrium when we require that out-of-equilibrium beliefs are strongly coherent or, equivalently, that there is a unique undefeated equilibrium in the terminology of MOP. This unique equilibrium is the natural outcome introduced in section 2 of the paper and denoted with a * in the discussion above. To show that this is the unique equilibrium, we must show that it defeats other candidate equilibria and that it is not defeated by other equilibria. We want to show that m^* is the unique undefeated equilibrium in our setting, which requires demonstrating that:

(a) there is no equilibrium $l < m^*$ such that l is undefeated by m^*
(b) there is no equilibrium $h > m^*$ such that h is undefeated by m^*
(c) there is no equilibrium $l < m^*$ such that l defeats m^*
(d) there is no equilibrium $h < m^*$ such that h defeats m^*

We must structure the disucssion in this manner because the arguments are somewhat different for l and h, so that it is best to discuss (a) and (b) separately. In addition, given the nonexistence under the Grossman-Perry refinement above, it is important to show that the set of undefeated equilibria is not empty, necessitating parts (c) and (d).

7.1 Imposing strong coherency on beliefs

We find it efficient to study restrictions on equilibria in two steps. First, we find the set of price functions that can support these equilibria. Second., we reinterpret these price functions in terms of probabilities and then look for restrictions imposed by strongly coherent beliefs. We combine graphical and analytical techniques to this end.

7.1.1 Low quality equilibria are defeated by m^*

Consider a candidate equilibrium $l < m^*$. For this equilibrium to be supported by a price function $p(m,l)$, it must be the case that

$$p(m,l) \leq \left\{ \begin{array}{c} C(m,l) \text{ for } m \leq l \\ D(m,l) \text{ for } m \geq l \end{array} \right\}$$

and

$$\mu(\tau = c|m) = \frac{p(m,l)}{\lambda m}$$

as the supporting belief. Take m^* as the alternative equilibrium. In this context, the set $K = [c,d]$ since both τ types prefer their payoffs in the alternative equilibrium. That is,

$$\begin{aligned} \pi(m^*, p^*, \tau &= c) = p^* - \psi(m^*) > p(l,l) - \psi(l) = \pi(l, p(l,l), \tau = c) \\ \pi(m^*, p^*, \tau &= d) = p^* > p(l,l) = \pi(l, p(l,l), \tau = d) \end{aligned}$$

Hence, it follows that the Bayesian out-of-equilibrium belief is $\gamma(\tau = c|m^*) = \rho_c$. However, supporting beliefs must satisfy

$$\mu(\tau = c|m^*) = \frac{p(m^*,l)}{\lambda m^*} = \frac{\lambda \rho_c l}{\lambda m^*} = \frac{l}{m^*}\rho_c < \rho_c$$

so that any candidate low equilibrium l is defeated by m^*.

7.1.2 High quality equilibria are defeated by m^*

Consider a candidate equilibrium $h > m^*$ with

$$p(m,h) \leq \left\{ \begin{array}{c} C(m,h) \text{ for } m \leq h \\ D(m,h) \text{ for } m \geq h \end{array} \right\}$$

as the supporting price function and

$$\mu(\tau = c|m) = \frac{p(m,h)}{\lambda m}$$

as the supporting beliefs. Again, take m^* as the alternative equilibrium. In this setting, the set $K = [c]$ since only $\tau = c$ types prefer their payoffs in the alternative equilibrium. That is,

$$\begin{aligned} \pi(m^*, p^*, \tau &= c) = p^* - \psi(m^*) > p(h,h) - \psi(h) = \pi(h, p(h,h), \tau = c) \\ \pi(m^*, p^*, \tau &= d) = p^* < p(h,h) = \pi(h, p(h,h), \tau = d) \end{aligned}$$

Hence, it follows that the Bayesian out of equilibrium belief is $\gamma(\tau = c|m^*) = 1$. However, supporting beliefs must satisfy

$$\mu(\tau = c|m^*) < 1$$

so that any h is defeated by m^*.

7.1.3 No low quality equilibrium defeats m^*

Consider a candidate equilibrium m^* and take any $l < m^*$ as the alternative equilibrium (the price in the alternative equilibrium is $p = \lambda\rho_c l$). In this setting, the set K is empty, since neither τ type prefers his payoffs in the alternative equilibrium. Since the set of types motivated to send the signal is empty, there are no restriction on the beliefs that can be used to support m^*. Hence, it is legitimate to use any

$$\mu(\tau = c|m) < \frac{C(m, m^*)}{\lambda m}$$

to support m^* including $\mu(\tau = c|m^*) = \rho_c$.

7.1.4 No high quality equilibria defeats m^*

Consider a candidate equilibrium m^* and take any $h > m^*$ as the alternative equilibrium (with a price $p = \lambda\rho_c h$). In this setting, the set $K = [d]$ since only $\tau = d$ types prefer their payoffs in the alternative equilibrium. That is,

$$\begin{aligned} \pi(m^*, p^*, \tau &= c) = p^* - \psi(m^*) > p - \psi(h) = \pi(h, p, \tau = c) \\ \pi(m^*, p^*, \tau &= d) = p^* < p = \pi(h, p, \tau = d) \end{aligned}$$

Hence, it follows that $\mu(\tau = d|h) = 1$ and, symmetrically, $\mu(\tau = c|h) = 0$. This belief is consistent with the support of m^*:

$$\mu(\tau = c|m) < \frac{p^*}{\lambda m} = \rho_c$$

That is, there is no h that defeats m^*.

7.2 Supporting beliefs for m*

Given the above, we know that m^* is the unique undefeated equilbrium or, equivalently, the unique equilibrium with strongly coherent out-of-equilibrium beliefs. There is still substantial latitude in the specification of these out-of-equilibrium beliefs, as consideration of Figure 7 makes clear.

If a message $m < m^*$ were observed, then the agents in the model would be genuinely puzzled, for no one would have any incentive to issue that message relative to m^*. Formally, there is no restriction admissable out-of-equilibrium beliefs. For that reason, it is coherent for them to believe that it could have been sent by either type, so that the associated price would be $p = \lambda\rho_c m$ as in Figure 7 (solid dark line for $m < m^*$). More generally, any belief function that supports m^* would be admissable as well (any price function below the committed type's isoprofit curve).

If the message $m > m^*$ were observed, then it would be interpreted as only being sent by a discretionary type, so that the belief function would be $\mu(= c|m > m^*) = 0$ and the related price would also be zero, as shown in the figure.

7.3 Comparison with weakly coherent beliefs

It is useful to make a comparison with the analysis of section 5 above, in which we studied the Grossman and Perry [1986] refinement that involved weakly coherent beliefs. The central difference is whether a low quality out-of-equilibrium message can be constructed that is inconsistent with the beliefs necessary to support the optimal equilibrium q^*. When we require only that a belief is weakly coherent, then the rewards to sending a message in the range $m < m^*$ involves the viewpoint on the committed firm's part that he can really separate himself from the discretionary firm and therefore earn a high price $p = \lambda m$. As a result of the fact that messages m can be constructed that are in the committed firm's interest and not the discretionary firm's interest, there is an out-of-equilibrium Bayesian belief $\mu_c = 1$ that is inconsistent with the support belief of m^*. By contrast, when we require that the belief is strongly coherent, the committed firm thinks that the message $m < m^*$ will be interpreted as being from an alternative equilibrium in which the price will be $p = \lambda\rho_c m < \lambda\rho_c m^*$. Neither the committed or discretionary firm benefits from such a price, so that such an out-of-equilibrium signal would never be purposefully sent: no inconsistency arises with the belief function that would support m^*.

8 Extension to multiple sender types

We have seen that the imposition of strong coherence on out-of-equilibrium beliefs, which is the requirement that beliefs about sender type based on an out-of-equilibrium message should be consistent with Bayesian reasoning based on an equilibrium in which the message might actually be sent, provides a powerful restriction on the set of imitative signalling outcomes, limiting the equilibria to the single optimal equilibrium that we suggested in section 2.

It is of interest to think about how this approach works in a modification of our environment where there are multiple sender types, where these types are indexed by the degree of their commitment. Let the set of sender types be T and let the mass of agents of a set of types $t \subset T$ be $\rho(t)$. In particular, though, we suppose that at least one member is a fully committed decision-maker and another is a fully discretion-maker, in ways that are explained further below.

For other members of type $\tau \in T$, we assume that the firm can produce a quality $q \neq m$ but only if it pays a cost $\xi(\tau)$. That is, suppose that the price is p and the announced quality level is m. Then, the profits of the firm are

$$p - \psi(m)$$

if it produces the announced quality level and

$$p - \xi(\tau) - \psi(q)$$

if it pays the deviation cost and produces quality q. If it pays the cost, the optimal action is $q = 0$ as above, so that a firm will pay the cost only if

$$\psi(m) > \xi(\tau)$$

That is: some types will find it to behave in a discretionary manner in some contexts. From this perspective, we can define the fully committed decision-maker as one that has a sufficiently high cost Ξ such that $\max_{m \in M} \psi(m) < \Xi$ and the fully discretionary decision-maker as one that has a zero cost.

8.1 A continuum of types

Suppose now the set of sender types is continuous. For example, we can let the type index τ run from 0 to 1 with agents of type τ having cost $\xi(\tau)$ and assume that this cost is an increasing function of type. Then, the set of agents that will behave in a discretionary manner is simply

$$T_d(m) : \{\tau : \xi(\tau) < \psi(m)\}$$

That is, all agents with costs $\xi(\tau) < \psi(m)$ will pay the cost and produce zero quality.[9] We denote the corresponding set of committed types as $T_c(m)$. Notice the division of types depends on the message m since it determines the benefit of not producing the costly quality.

[9] To deal with the tie issue, we assume whenever the type is indifferent between being committed or not, he chooses to be committed. So $\xi(\tau) \geqslant \psi(m)$ for all types $\tau \in T_c$.

Accordingly, the pooled equilibrium price is

$$p(m) = \lambda\varphi(m)\, m$$

where

$$\varphi(m) = \text{prob}(\tau \in T_c | m)$$

One example of the pooled equilibrium locus is illustrated by the dark dotted line in Figure 8. In this example, the distribution of costs has support $[0, 1]$ with a mass 0.25 on the zero cost point and the costs are otherwise distributed smoothly.[10]. Notice that the pooled equilibrium locus is not strictly increasing, since the price is a product of m, which is increasing, and $\varphi(m)$, which is decreasing because higher levels of m raise production cost $\psi(m)$ and make it desirable for more firm types to behave in a discretionary manner

The payoff of each type of firm is:

$$\begin{aligned}
\pi(m, p, \tau \in T_c(m)) &= \lambda\varphi(m)\, m - \psi(m) \\
\pi(m, p, \tau \in T_d(m)) &= \lambda\varphi(m)\, m - \xi(\tau)
\end{aligned}$$

Define m^* to be the unique message that maximizes the payoff of the most committed type of agent, $\tau = 1$, for which $\xi(\tau) = \Xi$. That is, the message maximizes this type's payoff across all pooled equilibria, taking into account the effect of the message on beliefs about sender type

$$m^* = \arg\max_m \lambda\varphi(m)\, m - \psi(m)$$

We restrict attention to the case in which there is a single maximizer, which imposes some weak restrictions on $\varphi(\cdot)$ and $\psi(\cdot)$ to assure the uniqueness of m^*.

8.2 Imposing strong coherency on beliefs

A similar argument to section 7 can be used to show m^* is the unique undefeated equilibrium in our setting. This requires demonstrating the following four facts for the same reasons as above

(a) that there is no equilibrium $l < m^*$ such that l is undefeated by m^*

(b) that there is no equilibrium $h > m^*$ such that h is undefeated by m^*

(c) that there is no equilibrium $l < m^*$ such that l defeats m^*

(d) that there is no equilibrium $h < m^*$ such that h defeats m^*

However, in this continuum of types case, monotonicity of $p(m)$ does not necessarily hold, which complicates the argument relative to the one in section 7. While

[10] Specifically, the distribution takes the form

$$F(\xi) = .25 + .75 * \beta(\xi; 2, 2)$$

where $\beta(\xi, 2, 2)$ is the cdf of a beta distribution with parameters 2 and 2.

we do not have monotonicity of $p(m)$, though, we do have the fact that the isoprofit curves of committed types are monotonically increasing at any profit level, i.e., that $p = \pi - \psi(m)$ is increasing (and convex) in m for any fixed π.

As in our prior analysis, we use the approach of first thinking about supporting prices and then thinking about supporting beliefs.

8.2.1 Low quality equilibria are defeated by m^*

Consider a candidate equilibrium $l < m^*$ with $p(l) = \lambda\varphi(l)\,l$ as the equilibrium price and $\mu(\tau|m)$ as supporting beliefs about type, conditional on a mesage m being sent. For l to be an equilibrium, it must be supported by a price function that has the form

$$p(m; l)$$

where $p(m; l)$ indicates the price would obtain if out-of-equilibrium message m is sent relative to the equilibrium l.

It must thus be the case that

$$p(m^*; l) < p(l)$$

since any discretionary actor at l would otherwise prefer to be a discretionary actor at m^*.

However, in the equilibrium m^*, it must be the case that

$$p(m^*) > p(l)$$

since m^* is profit-maximizing for the committed type. That is, $p(m^*) - \psi(m^*) > p(l) - \psi(l)$ implies that

$$p(m^*) > p(l) + [\psi(m) - \psi(l)] > p(l)$$

From the price perspective, this is the central difficulty for a low quality equilibrium: the supporting price $p(m^*; l)$ is inconsistent with the price $p(m^*)$.

Now, an out-of-equilibrium belief that supports l is

$$\mu(\tau \in T_c|m^*; l) = \frac{p(m^*, l)}{\lambda m^*} < \frac{p(l)}{\lambda m^*} < \frac{p(m^*)}{\lambda m^*} = \varphi(m^*)$$

In words, l is an equilibrium only if out-of-equilibrium beliefs are that the firm is less likely to behave in a committed manner at m^* than would be the case if m^* were actually an equilibrium.

To investigate the coherency of this belief, we consider the types of firms τ that would be better off in equilibrium m^*. The answer is all $\tau \in T$ because $p(m^*)$ leads to higher profits than $p(l)$ when one is either discretionary or committed. Hence, coherent out-of-equilibrium beliefs imply $\gamma(\tau \in T_c|m^*) = \varphi(m^*)$ which is not consistent with the supporting belief above. So, the m^* equilibrium defeats the l equilibrium.

8.2.2 High quality equilibria are defeated by m^*

Matters are somewhat more complicated when we consider a candidate equilibrium $h > m^*$ with $\mu(\tau|m;h)$ as the supporting beliefs. However, one important feature is the same,

$$\mu(\tau \in T_c|m^*;h) < \varphi(m^*).$$

That is, the signal m^* must mean that the firm is less likely to behave in a committed manner at m^* than would be the case if m^* were actually an equilibrium. If this were not the case, then the implied price would be at least $p(m^*) = \lambda\varphi(m^*)\, m^*$ and h could not be an equilibrium because it would not be individually rational for a committed type firm to choose it.

Taking m^* as the alternative equilibrium to h, we now study three cases.

High prices ($p(h) > p(m^*)$). For the types $\{\tau : \xi(\tau) < \psi(m^*)\} = T_d(m^*)$ and $\{\tau : \xi(\tau) \geqslant \psi(h)\} = T_c(h)$, there is no change in behavior since $T_d(m^*) \subset T_d(h)$ and $T_c(h) \subset T_c(m^*)$. That is: if one is discretionary at a low quality level m^* then one will be discretionary at a higher quality level h and if one is committed at a high quality level h then one will also be committed at a lower quality level m^*. For these types, the payoff comparisons are the following:

$$\begin{aligned}\pi(m^*, p^*, T_c(m^*) &= p(m^*) - \psi(m^*) > p(h) - \psi(h) = \pi(h, p, \tau \in T_c(h))\\ \pi(m^*, p^*, \tau &\in T_d(m^*)) = p(m^*) - \xi(\tau) < p(h) - \xi(\tau) = \pi(h, p, \tau \in T_d(h))\end{aligned}$$

So only very committed types, $T_c(h) = \{\tau : \xi(\tau) \geqslant \psi(h)\}$ types prefer their payoffs in the alternative equilibrium m^*.

For the types with cost $\xi(\tau) \in [\psi(m^*), \psi(h))$, $\tau \in T_c(m^*)$ and $\tau \in T_d(h)$: they switch from discretion to committment when moving from h to m^*. If the equilibrium is h, their payoff is

$$\pi(h, p, \tau \in T_d(h)) = p(h) - \xi(\tau)$$

In equilibrium m^*, their payoff is

$$\pi(m^*, p^*, \tau \in T_c(m^*)) = p(m^*) - \psi(m^*)$$

So only the subset of types with cost $\xi(\tau) > \psi(m^*) + [p(h) - p(m^*)]$ will prefer the alternative equilibrium m^*. This range of types is not empty since $\psi(h) > \psi(m^*) + [p(h) - p(m^*)]$. Therefore, the set K which prefers the alternative equilibrium m^* is $\{\tau : \xi(\tau) > \psi(m^*) + [p(h) - p(m^*)]\}$. Hence, it follows that $\gamma(\tau \in T_c|m^*) = 1$ because all $\tau \in K$ have cost $\xi(\tau) > \psi(m^*)$. This strongly coherent belief is not consistent with the supporting belief.

Low prices ($p(h) < p(m^*)$). In this case, the set $K = T$ since types in both sets $T_d(m^*)$ and $T_c(h)$, as defined above, prefer their payoffs in the alternative equilibrium m^*. That is,

$$\begin{aligned}\pi(m^*, p^*, \tau &\in T_c(h)) = p(m^*) - \psi(m^*) > p(h) - \psi(h) = \pi(h, p, \tau \in T_c(h))\\ \pi(m^*, p^*, \tau &\in T_d(m^*)) = p(m^*) - \xi(\tau) > p(h) - \xi(\tau) = \pi(h, p, \tau \in T_d(m^*))\end{aligned}$$

For the types with cost $\xi(\tau) \in [\psi(m^*), \psi(h))$, they also prefer m^* because

$$\pi(m^*, p^*, \tau \in T_c(m^*)) = p\left(m^*\right) - \psi(m^*) > p\left(h\right) - \xi(\tau) = \pi(h, p, \tau \in T_d(h))$$

Hence, it follows that $\gamma(\tau \in T_c(m^*)|m^*;) = \varphi\left(m^*\right)$: all firm types are better off by sending the signal, which is not consistent with the supporting belief.

Equal prices $p\left(h\right) = p\left(m^*\right)$. In this case, the set $K = T$ but group $\{\tau : \xi(\tau) \leqslant \psi(m^*)\}$ is indifferent between h and m^*. Hence, it follows that $\mu(\tau \in T_c(m^*)|m^*; h) \geqslant \varphi\left(m^*\right)$, which is not consistent with the supporting belief.

Taking these cases together, we see that any h is defeated by m^*.

8.2.3 No low quality equilibrium defeats m^*

Consider a candidate equilibrium m^* with $\pi^* = \lambda\varphi\left(m^*\right) m^* - \psi\left(m^*\right)$ as the payoff and $\mu(\tau|m; m^*)$ as supporting beliefs. The sufficient condition for the belief function to support equilibrium m^* is such that:

$$\begin{aligned} \mu(\tau &\in T_c(m)|m \neq m^*) < \frac{\pi^* + \psi(m)}{\lambda m} \\ \mu(\tau &\in T_c(m^*)|m^*) = \varphi\left(m^*\right) \end{aligned}$$

That is: any $m \neq m^*$ leads to a pessimistic enough belief so that there is no incentive for any type to deviate from m^*.

Now, take $l < m^*$ as the alternative equilibrium (the price in the alternative equilibrium is $p = \lambda\varphi\left(l\right) l$). We have $p\left(l\right) < p\left(m^*\right)$ for the same reason discussed above.

In this case, the set K of signalling types is empty, since no τ type prefers his payoffs in the alternative equilibrium. Since the set of types motivated to send the signal is empty, there is no restriction on beliefs that can be used to support m^*. Hence, it is legitimate to use any belief function satisfying the sufficient condition above.

8.2.4 No high quality equilibria defeats m^*

Consider a candidate equilibrium m^* with $\pi^* = \lambda\varphi\left(m^*\right) m^* - \psi\left(m^*\right)$ as the payoff and $\mu(\tau|m; m^*)$ as supporting beliefs. The sufficient condition for the belief function to support equilibrium m^* is again that:

$$\begin{aligned} \mu(\tau &\in T_c(m)|m \neq m^*; m^*) < \frac{\pi^* + \psi(m)}{\lambda m} \\ \mu(\tau &\in T_c(m^*)|m^*; m^*) = \varphi\left(m^*\right) \end{aligned}$$

Now, take $h > m^*$ as the alternative equilibrium (with a price $p\left(h\right) = \lambda\varphi\left(h\right) h$). We also have three cases:

High prices $(p(h) > p(m^*))$. In this case, the set of agents with a signalling incentive is $K = \{\tau : \xi(\tau) < \psi(m^*) + [p(h) - p(m^*)]\}$ by same sort of calculations as in the high price case above. Hence, it follows that the Bayesian beliefs are $\gamma(\tau \in T_c(h)|h) = 0$ because all $\tau \in K$ have a cost $\xi(\tau) < \psi(m^*) + [p(h) - p(m^*)] < \psi(h)$. This strongly coherent belief is consistent with the sufficient condition above. So, m^* is not defeated by h that gives $p(h) > p(m^*)$.

Low prices $(p(h) < p(m^*))$. In this case, the set K is empty: no type likes the alternative equilibrium h. There is thus no restriction on off-equilibrium belief at h. So it is legitimate to use any belief function which is consistent with the supporting belief of m^*. So, m^* is not defeated by h that gives $p(h) < p(m^*)$.

Equal prices $p(h) = p(m^*)$. In this case, only the group $\{\tau : \xi(\tau) \leqslant \psi(m^*)\}$ is indifferent between sending message h and m. All the other types are strictly worse off by sending message h given the argument in the equal price case above. Therefore, the set K which prefers alternative equilibrium h does not include any type with cost $\xi(\tau) > \psi(h)$. It then follows that $\gamma(\tau \in T_c(h)) = 0$, which satisfies the sufficient condition of the supporting beliefs of m^*. So, m^* is not defeated by h that gives $p(h) = p(m^*)$.

8.3 Three types

Now we turn to the case that the set of sender types is discrete. Suppose that there are just three types, $T = \{c, w, d\}$, with a weak type supplementing the prior discretionary type ($\tau = d$ has a zero cost ($\xi(d) = 0$)) and committed type ($\tau = c$ has an infinite cost). To economize on notation, let the intermediate "weak" type $\tau = w$ have a cost ξ. Let $\overline{q}$ be the quality level such that $\psi(\overline{q}) = \xi$. For all $m < \overline{q}$, the weak type will optimally set $q = m$ and for all $m > \overline{q}$, a zero quality will be produced. Weak firms will be indifferent between producing m or not if $m = \overline{q}$. Hence, the pooled equilibrium price locus will take the form

$$E(m) = \left\{ \begin{array}{ll} \lambda(\rho_c + \rho_w)m & \text{for } m < \overline{q} \\ \lambda(\rho_c + \theta\rho_w)m & \text{for } m = \overline{q} \text{ and any } 0 \leq \theta \leq 1 \\ \lambda\rho_c m & \text{for } m > \overline{q} \end{array} \right\}$$

This locus is shown by the dotted line in Figure 9. Even though the locus is not smooth with a break at $\overline{q}$, it is not more conceptually difficult than the case of continuous types. Using the notation from the previous section, the three type case concentrates the distribution of types $\rho(\tau)$ on three points $\{c, w, d\}$ with mass $\{\rho_c, \rho_w, 1 - \rho_c - \rho_w\}$. Moreover, the type $\tau \in \{c, w, d\}$ has cost $\{+\infty, \xi, 0\}$ respectively. Therefore, we can apply the arguments like those above to show that there can be only three potential equilibria with coherent beliefs: the first is one at point "a" in the figure, the second is at a point like "b", and the third is at a point like "c". In cases "a" and "c", the committed decision-maker's optimum comes with a tangency of his isoprofit function to the relevant portion of the line, while there is a simple intersection in case "b". Further, for any particular parameterization, only one of these equilibria can have strongly coherent beliefs: the other two must be defeated by it.

9 Conclusions

We are interested in signalling games with one large agent who holds private information (for instance, about its preferences) as sender and other receiver agents that observe the message and rationally form expectations. Relative to prior analysis of such models, we focus on a case where types of senders have preferences about their most desired message, but there is not an explicit cost associated to issuing the message. We call a setting of "costless imitative signalling".

As illustration, we have constructed a simple one period game where quality of a product is advertised in advance by a firm, which may be of a "committed" type (forced by assumption to fulfil its announcement) or to a "discretionary" type (without the previous restriction). Consumers receive the message, form expectations about the quality, and offer a price reflecting those expectations. A key feature in this game is that if a firm is identified as "discretionary", then the offered price is zero. This feature generates a discontinuity in payoffs, which together with the costless imitative signalling characteristic, imply that the set of PBE is composed by a continuum of pooling equilibria. Basically, if consumers believe that the "committed" firm should issue message m, then it would be optimal for both types of senders to do so, resulting in a pooling equilibrium.

This case is interesting because it is a suitable simple representation of many economic contexts, where off-equilibrium beliefs substantially undermine any prediction power. We argue that this result is obtained because off-equilibrium beliefs are undisciplined. We examine what restrictions defining coherency of beliefs recover an outcome which we think intuitive: the committed type's message is mimicked by the discretionary type, consumers realize that messages do not distinguish between type and – knowing this – the commtted type becomes "leader", so that it is the best pooling for the committed type which should be predicted.

Exploring different refinement criteria in the literature, the *weakly coherent beliefs* of Grossman and Perry [1986] and the *intuitive criterion* of Cho and Kreps [1987], we find the surprising result that they do not help. The former criterion leads eliminates all equilibria, the latter selects an equilibrium where zero quality is announced and zero price is offered. In both of these refinements from the perspective of a candidate PBE, we can always find incentives for some deviation by a committed type but not for the discretionary type, ruling out that equilibrium.

However, we find a more appealing result after applying the approach suggested by Mailath, Okuno-Fujiwara, and Postelwaite [1993], which restricts beliefs to satisfy a criterion of *strong coherency.* Relative to the previously discussed criteria, this alternative approach involves the idea that consumers agents recognize that if the committed type has incentives to deviate from a candidate equilibrium, given that deviation, the discretionary has also incentives to deviate. Thus, the only candidate equilibrium where the committed type does not deviate from is the best pooling from the standpoint of the committed type. This natural result turn to be robust to the introduction of both with a finite and with an infinite number of types of senders.

References

[1] Cho and Kreps (1987), "Signaling Games and Stable Equilibria", *The Quarterly Journal of Economics*, Vol. 102, No. 2

[2] Grossman and Perry (1986), "Perfect Sequential Equilibrium", *Journal of Economic Theory*, XXXIX

[3] Kreps (1990), *Game Theory and Economic Modelling*, Oxford University Press

[4] Kreps and Wilson (1982), "Sequential Equilibria", *Econometrica*, L

[5] Mailath, Okuno-Fujiwara and Postelwhaite (1993), *Journal of Economic Theory*, 60

[6] MasCollel, Whinston and Green (1995), *Microeconomic Theory*, Oxford University Press

[7] Spence (1974), *Market Signalling*, Cambridge, MA: Harvard University Press

Notation table

Symbol	concept	variants
m	message	$m_c, m_d, m(\tau)$
τ	sender type	
c	committed type ($\tau = c$)	
d	discretionary type ($\tau = d$)	
p	receiver response; price	
ρ	prior distribution of type	$\rho_c, \rho_d, \rho(\tau)$
μ	posterior distribution of type	$\mu_c, \mu_d, \mu(\tau)$
q, Q	quality, maximum quality	
λ	value for consumer (per unit of quality)	
M	set of messages $m \in M$	$0 \leq m \leq Q$
ψ	cost of producing quality	$\psi(q)$
$*$	superscript: optimal equilibrium	
π	profit	$\pi = p - \psi(q)$
$\hat{}$	superscript: candidate equilibrium	
$'$	superscript: alternative equilibrium	
$\tilde{}$	superscript: various uses	
l, h	various message levels (low, high)	
δ	break even message for discretionary type	
C		
D		
R		

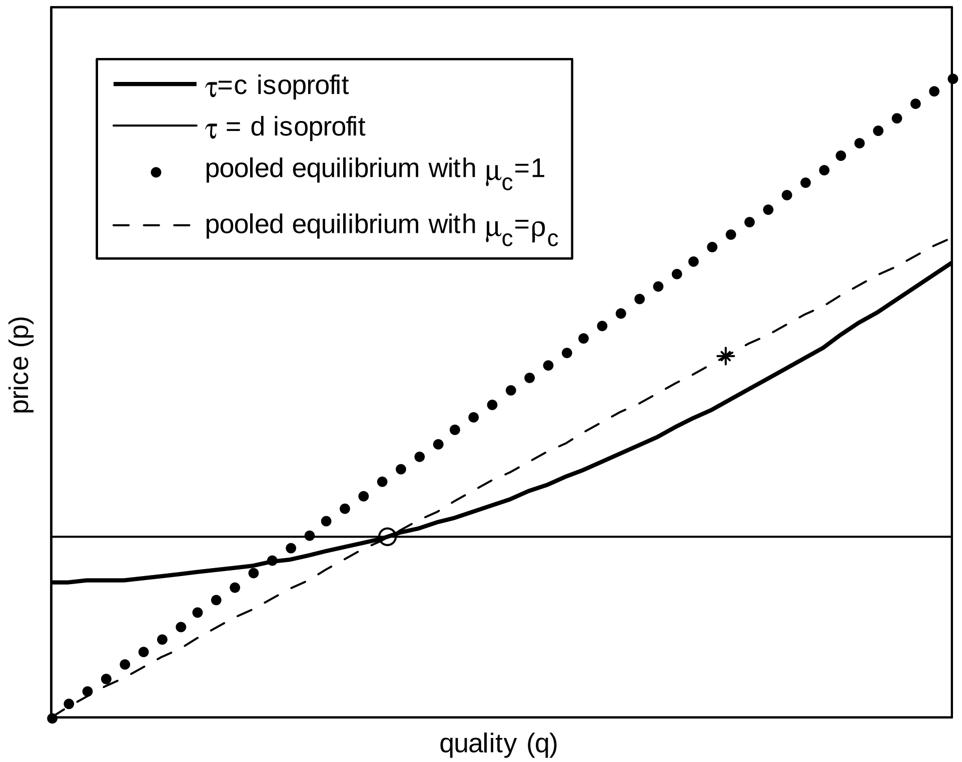

Figure 1: (Standard devices) The dotted line is the equilibrium relationship ($p = \lambda q$) between price and quality relationshp with known $\tau = c$; the dashed line is the pooled equilibrium relationship ($p = \lambda q \rho_c$) between price and quality relationshp with unknown type in population fractions; the horizontal light solid line is the isoprofit locus for the discretionary firm ($\tau = d$); and the heavy solid curve is the isoprofit locus for the committed firm, consistent with the profits earned in the candidate equlibrium $\widehat{q}$, indicated by the circle.

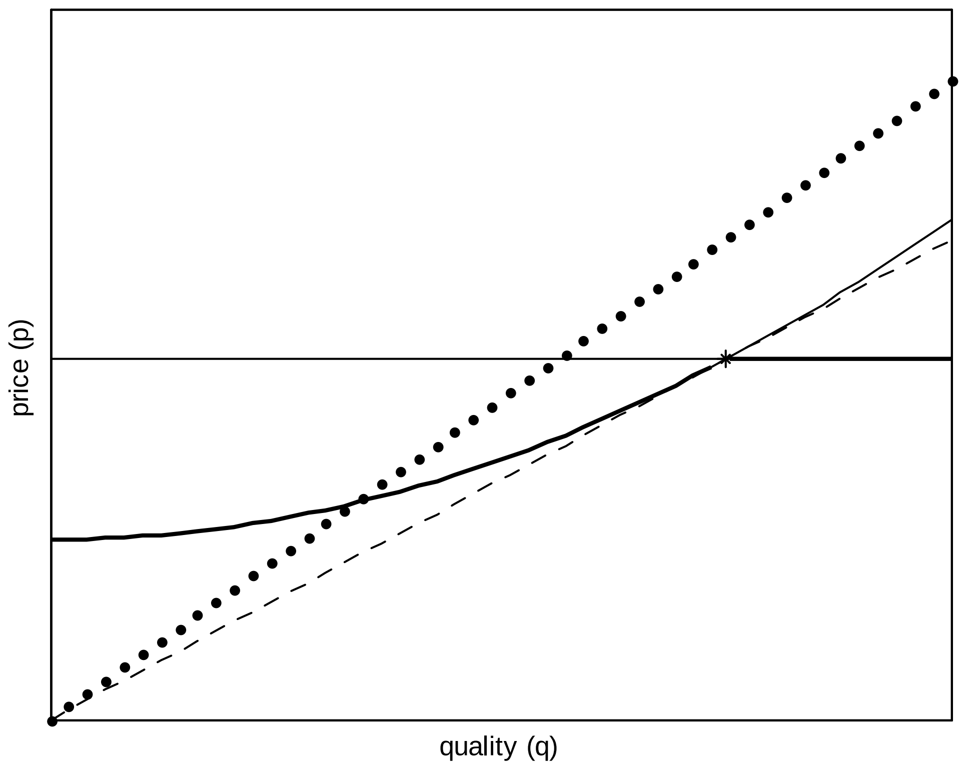

Figure 2: An optimal equilibrium. The supporting price function $p(m, q^*)$ must lie below the heavy black line except at $m = q^*$

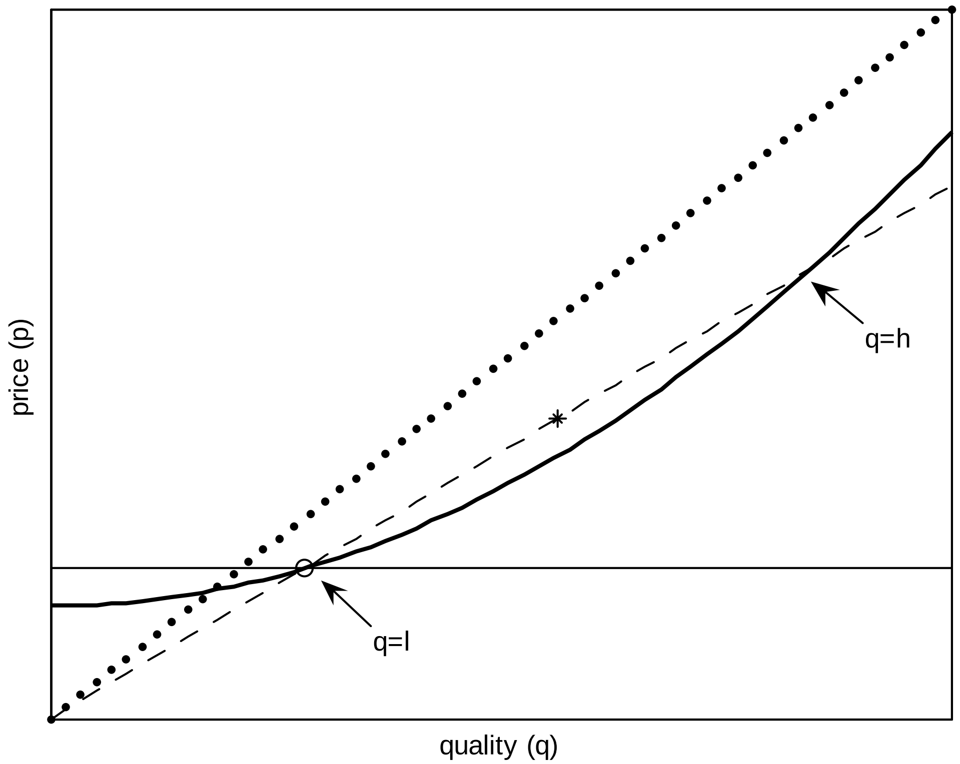

Figure 3: A low quality Bayesian perfect equilibrium $q = l$ occurs at the circled point. The supporting price function must lie below the heavy solid line for $q < l$ and below the light solid line for $q > l$.

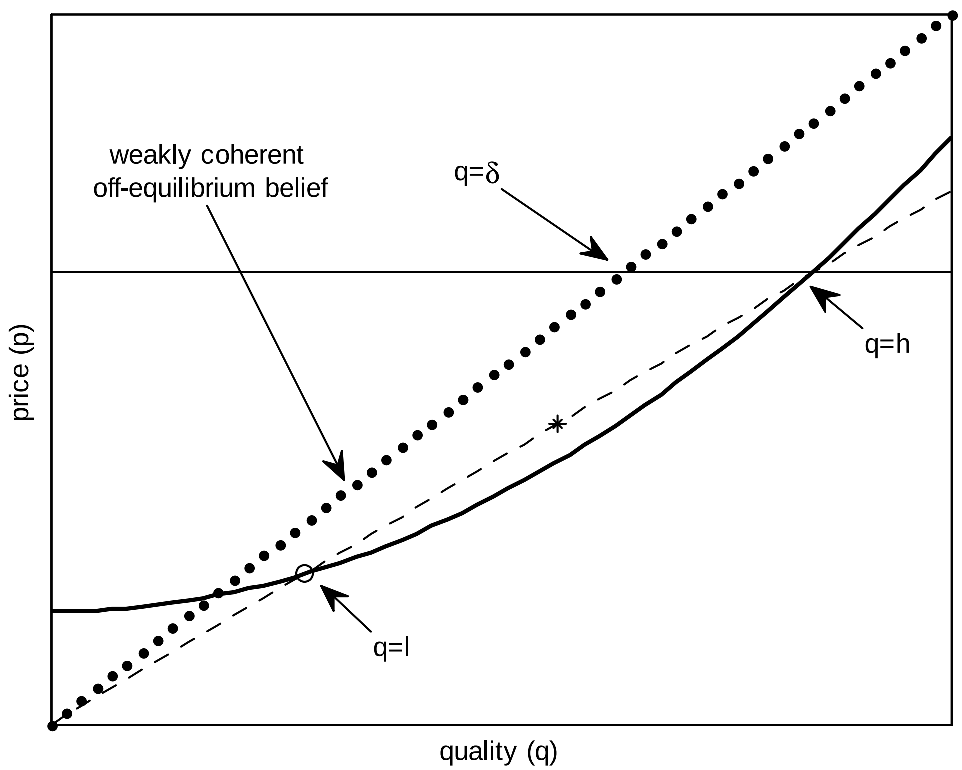

Figure 4: The breakdown of the high quality equilibrium $q = h$, under weakly coherent beliefs. The indicated point $l < m < \delta$ corresponds to a message that would be sent only by the $\tau = c$ type and thus would be inconsistent with the beliefs necessary to support the $q = h$ equilibrium.

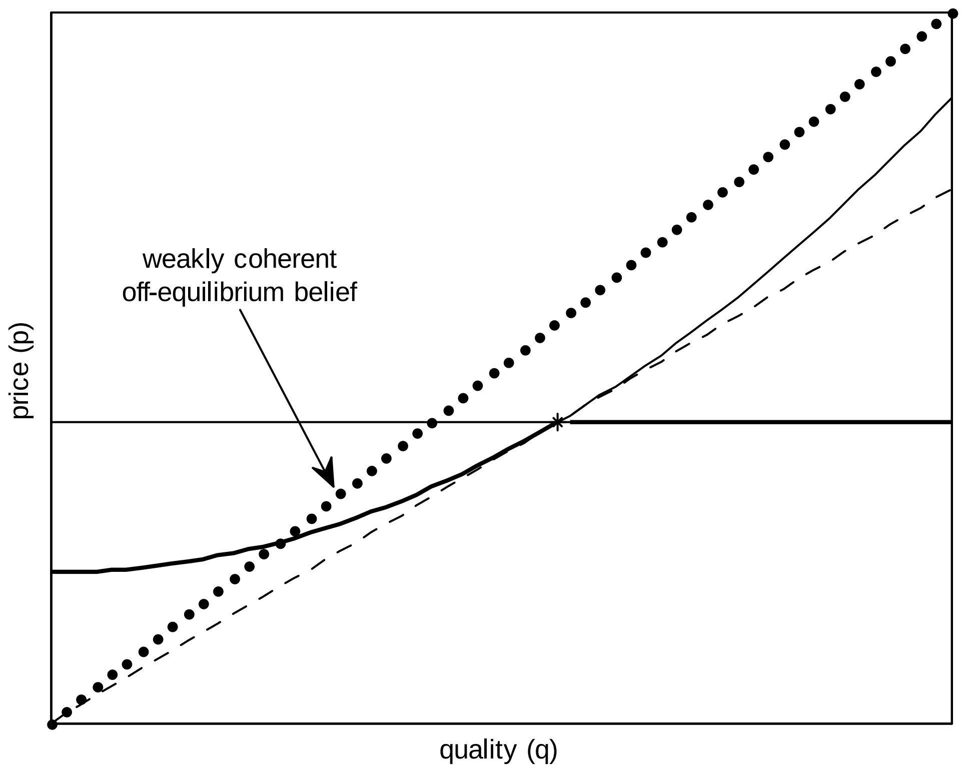

Figure 5: The optimal equilibrium is not perfect, in the sense of Grossman and Perry, because of a weakly coherent off equilibrium belief, in which a committed firm signals its type by reducing quality.

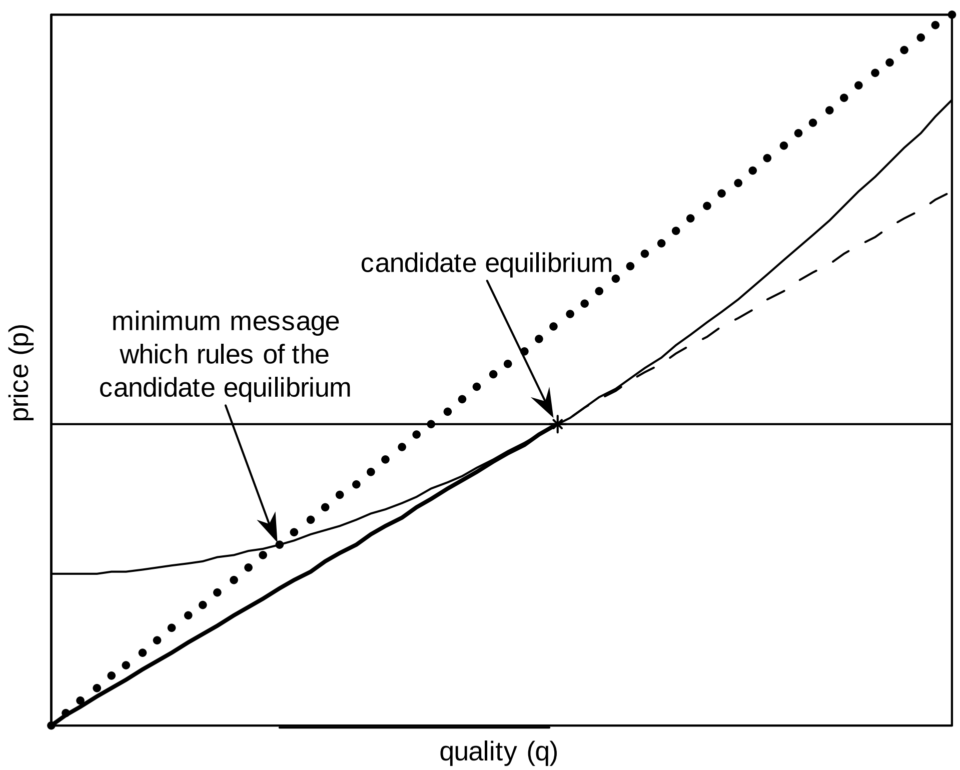

Figure 6: intuitive criterion figure

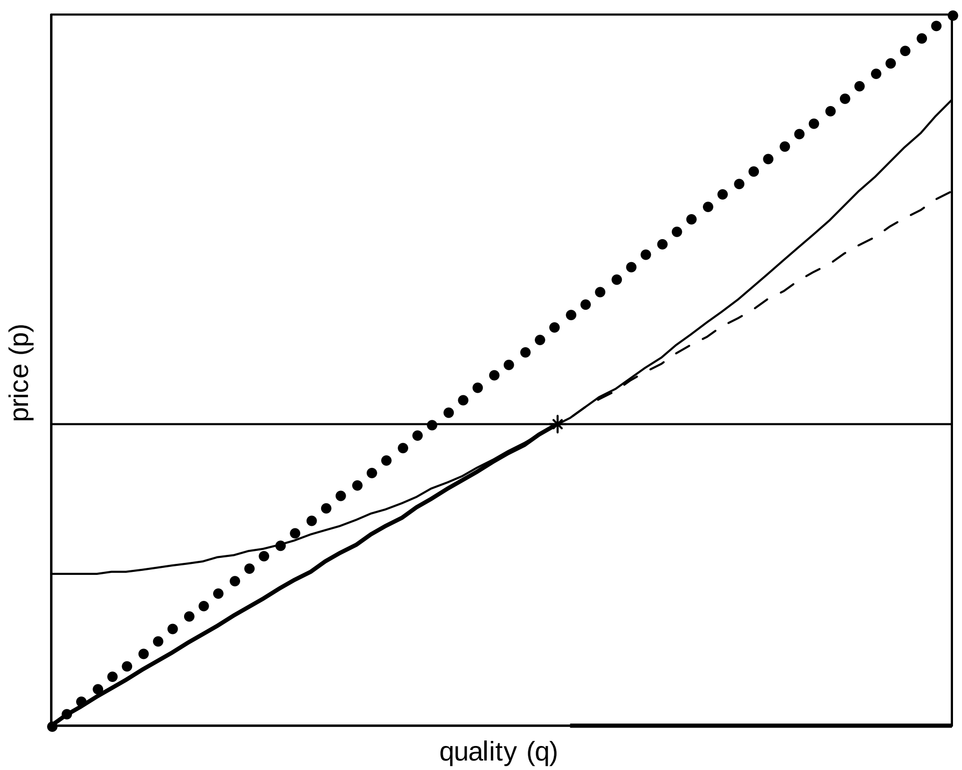

Figure 7: The optimal equilibrium is supported by a strongly coherent belief function

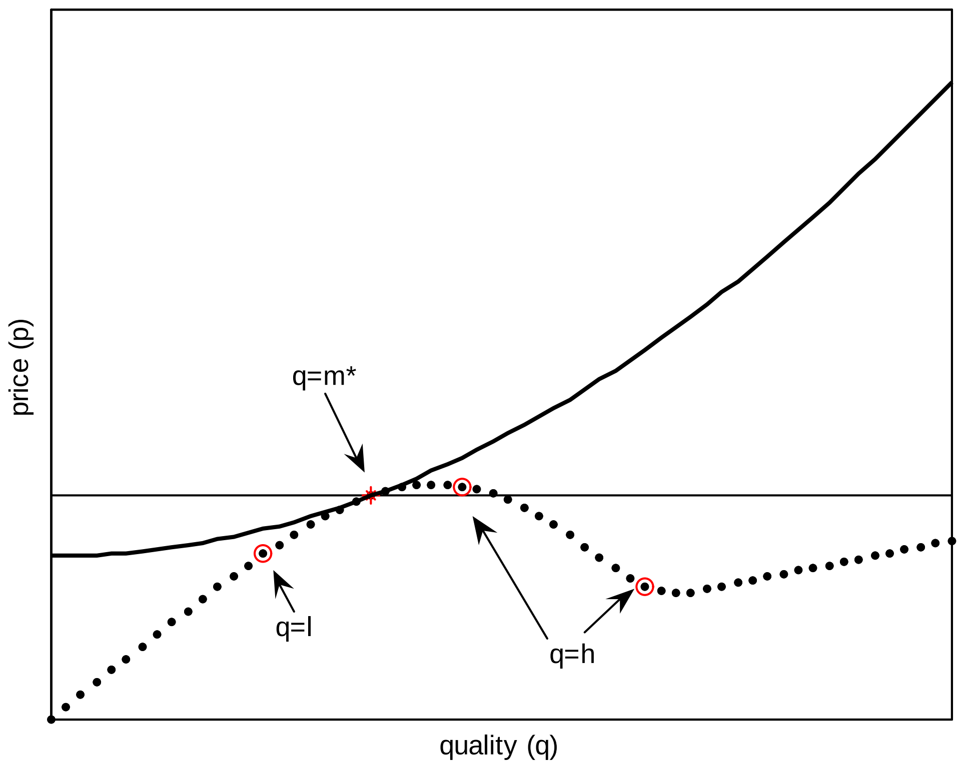

Figure 8: The optimal equilibrium with a continuum of types. The dotted line is the price function. The dark solid line is the isoprofit curve for the committed group. The light solid line is the isoprofit cruve for the discretionary group.

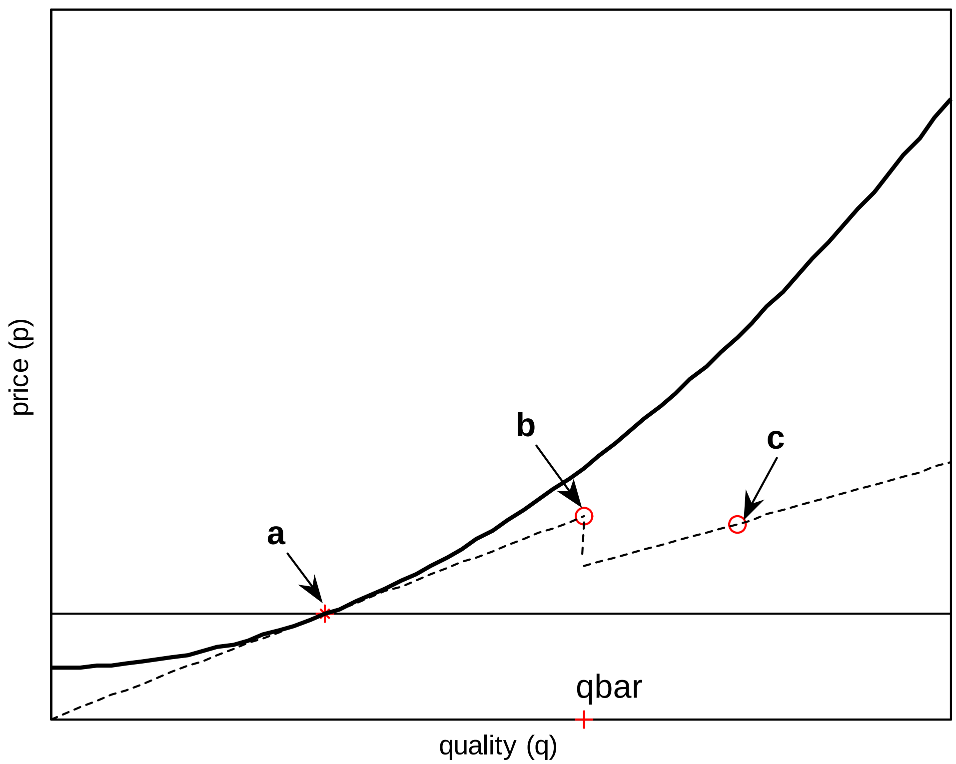

Figure 9: The optimal equilibrium with 3 types. In this parameterization, point a is the optimal equilibrium. However, point b or point c can be optimal under other parameterizations.