Sophisticated Signaling

Penghui Zhou and Erez Yoeli

August 4, 2016

Abstract

Spence (1973)'s model of costly signaling can explain why people conspicuously consume luxury goods. However, in many luxury markets, there is an emphasis on understatement that cannot be explained by costly signaling alone: why incur the expense of sending a signal but not send it as loudly and clearly as possible? We develop a signaling model to explain this phenomenon. Our model diverges from the standard model in the following ways: (1) There are two types of receivers, not one, and (2) senders have the opportunity to send no signal, a common signal, or a 'sophisticated' signal that requires an up-front investment by the receiver to recognize. There is an equilibrium where high-type senders will send the sophisticated signal, low-type senders send no signal, high-type receivers invest in recognizing the signal, and low-type receivers do not invest. This equilibrium exists either when higher quality senders prefer to outright avoid one type of receiver, or when higher quality senders and receivers sufficiently prefer to match with each other. We use evolutionary dynamics to show this equilibrium arises even when players do not choose their strategies deliberately, and apply the model to understand fashion trends, such as 'logo fatigue' in China.

Our preferences for many luxury goods can be simply understood as equilibrium outcomes in a costly signaling game (Veblen 1899, Spence 1973, Bearden and Etzel 1982, Bagwell and Bernheim 1996, Charles, Hurst and Roussanov 2007). That is, we learn to like precisely those goods that also distinguish us as especially wealthy, well-connected, or well educated. Intuitively, the most effective signals are loud and clear: say, well-known logos visibly printed or embroidered on shirts (e.g., Calvin Klein, Faconnable, Nike, and Polo), coats (Canada Goose), and handbags (Louis Vuitton).

But, sometimes, we develop preferences for luxury items that are hard to recognize from afar, with no logos, and no easily identified design elements (Bottega Veneta handbags, Theory shirts, or Loro Piana outerwear) (Berger and Ward 2010, Han, Nunes and Drèze 2010). These preferences for 'understated' luxury goods are not costly signals: if they were, we would maximize the signaling benefit from the expensive item, by, say, adding a logo. Such preferences are also not 'counter-signals' (Feltovich, Harbaugh and To 2001): a successful CEO seeking to counter-signal when attending a meeting where everyone else is sporting an Armani suit and a Rolex would not do so by seeking out obscure, hard-to-recognize brands that are just as expensive as her peers'. Rather, she'd wear clothes that are not expensive at all: sweats and a Timex. Why bother to pay for a luxury good, if one chooses to obscure its value?

There are two answers. The first is quite intuitive: people might sometimes wish to avoid attention from those who are only interested in them for their money. The second is perhaps subtler: sometimes, individuals may be willing to give up on relationships with most other people to signal that they strongly value particular relationships. When will each of these answers apply, if at all?

To answer this question, we propose a model of *sophisticated* signaling. As in standard signaling models like Spence's (Spence 1973), there are senders and receivers. Senders are either high and low quality, but cannot directly communicate their quality to receivers. Instead, they have the opportunity to send a signal, which receivers can observe prior to deciding whether to match with a sender. The model diverges from the standard model in the following ways. First, there are also high and low quality receivers. Second, senders have the opportunity to send an additional, third signal. In particular, the signaler chooses between sending no signal, a common signal, or a 'sophisticated' signal that requires an initial investment by the receiver to recognize.

We derive conditions under which there is an equilibrium where high-type senders will send the sophisticated signal, low-type senders send no signal, high-type receivers invest in recognizing the signal, and low-type receivers do not invest. To gain insight on when sophisticated signaling is expected to explain behavioral phenomena of interest, we compare these conditions to those needed for a standard separating equilibrium, in which high-type senders send a common signal, low-type senders send no signal, and neither type of receiver invests in recognizing the sophisticated signal. We also show that the equilibria of interest arise under evolutionary dynamics with the Wright-Fisher (Wright 1931, Fisher 1930, Fudenberg and Imhof 2006) framework with selection and low mutation rates. We apply this model to shed light on a number of questions, including: why there is a shift of preference from ‘logo-ed’, ostentatious goods to understated ones, and under what conditions this shift occurs.

1 Model

The game (see Fig. 1; a formal definition of the game is presented in Appendix A) is played by two players: a sender (S) and a receiver (R). It proceeds as follows:

1. The sender and receiver’s types are determined by chance. The sender is of high quality (H) with probability p and low quality (L) with probability $1-p$. The receiver is of high quality with probability q and low quality with probability $1-q$. We label the sender and receiver’s types $t_s, t_r \in \{H, L\}$, respectively.

2. The receiver decides to Invest (I) or Not Invest (N). The sender does not observe whether the receiver invests.

3. The sender chooses to send one of three signals: sophisticated signal (SS), common signal (CS), or no signal (NS). The receiver observes the signal. However, the receiver can only tell apart the signals SS and NS if she has invested.

4. The receiver decides to accept (A) or reject (R) the sender. If the receiver accepts, we say that the players are ‘matched’.

Payoffs are as follows: The sender pays c^s to send a signal, regardless of whether it is sophisticated or common. The receiver pays $c^r_{t_r}$ to invest, a cost dependent on his type.

The benefit from matching depends on a player’s type and their match’s type. The sender’s benefit if there is a match is $b^s_{t_s,t_r}$, and the receiver’s benefit is $b^r_{t_s,t_r}$, where t_s and t_r are the sender’s and receiver’s type, respectively. Both senders and receivers benefit 0 when they do not match.

We assume:

(A1) $0 < p, q < 1$

(A2) Matching with a high type is strictly better than matching with low type:

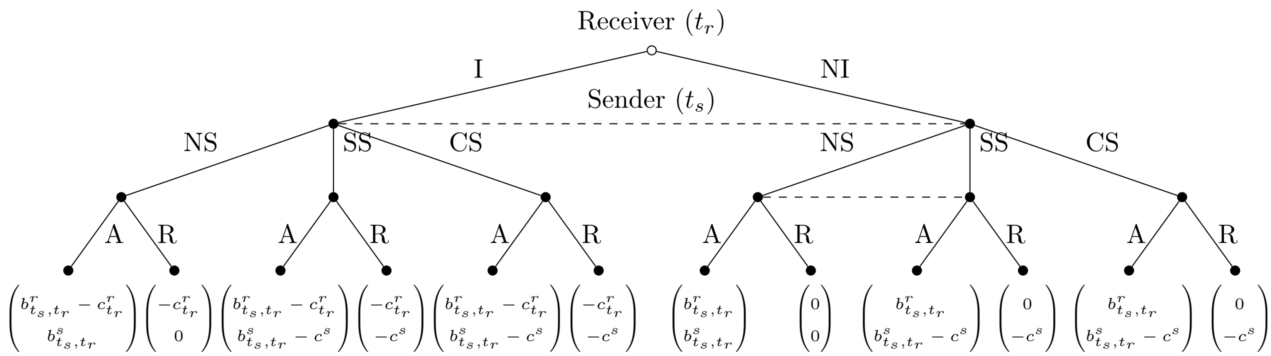
$$b^s_{HH} > b^s_{HL}, b^s_{LH} > b^s_{LL} b^r_{HH} > b^r_{LH}, b^r_{HL} > b^r_{LL}$$

(A3) All types prefer matching with a high type than not to match:

$$b^r_{HH} > 0,\ b^r_{HL} > 0,$$
$$b^s_{HH} > 0,\ b^s_{LH} > 0.$$

(A4) Low types receivers incur strictly higher cost of investing: $c^r_L > c^r_H > 0$,

Figure 1: The Sophisticated Signaling Game



2 Results

We use the standard solution concept for games of this type, the Perfect Bayesian Equilibrium (PBE) and its refinement, the Sequential Equilibrium (SE) (Fudenberg and Tirole 1991). In a PBE, no player can benefit by unilaterally deviating, and players' beliefs must be consistent with equilibrium outcomes in all states that are reached in equilibrium. In an SE, this is true even in states that are not reached in equilibrium. See appendix B for formal definitions.

There are many SE of this game. We categorize them by players' behavior along the equilibrium path. For example, consider the following strategy profiles, which specify players' actions at each decision node, as well as their beliefs about others' types, specified as a tuple (t_H, t_L) where t_H refers to a high type player and t_L refers to a low type player:

1. High type receivers invest, Low type receivers do not. Senders' belief on receiver's type is always $(q, 1-q)$ at all information sets. High type senders send SS, Low type senders send NS. High type receivers' belief upon seeing SS is $(1, 0)$ and accept; belief upon seeing NS is $(0, 1)$ and reject. Low type receivers' belief upon seeing SS is $(0, 1)$ and reject; belief upon seeing NS is $(0, 1)$ and reject.

2. No receiver invests. Senders' belief on receiver's type is always $(q, 1-q)$ at all information sets. High type sender sends CS, Low type sender sends NS. High type receivers' belief upon seeing CS is $(1, 0)$ and accept; belief upon seeing NS is $(0, 1)$ and reject. Low type receivers' belief upon seeing CS is $(1, 0)$ and accept; belief upon seeing NS is $(0, 1)$ and reject.

We refer to the first profile as a Sophisticated Separating Equilibrium (SSE), and to the second as a Common Separating Equilibrium (CSE) (For a complete categorization of equilibria, see appendix B.)

Proposition 1. *Assume*[1]*:*

$$pb^r_{HH} - c^r_H \neq \max\{pb^r_{HH} + (1-p)\,b^r_{LH}, 0\},$$
$$pb^r_{HL} - c^r_L \neq \max\{pb^r_{HL} + (1-p)\,b^r_{LL}, 0\}.$$

Then there is a Sophisticated signaling Equilibrium if and only if any one of the following holds:

(SSE1) Receivers of both type value a match with high type sender sufficiently to invest:

$$pb^r_{HH} - c^r_H > \max\{pb^r_{HH} + (1-p)\,b^r_{LH}, 0\},$$
$$pb^r_{HL} - c^r_L > \max\{pb^r_{HL} + (1-p)\,b^r_{LL}, 0\}.$$

But only high-type senders value matching with a random receiver enough to send a sophisticated signal:

$$qb^s_{HH} + (1-q)\,b^s_{HL} \geq c^s \geq qb^s_{LH} + (1-q)\,b^s_{LL}. \tag{1}$$

(SSE2) High-type receivers value a match with a high-type sender sufficiently to invest, but low-type receivers do not:

$$pb^r_{HH} - c^r_H > \max\{pb^r_{HH} + (1-p)\,b^r_{LH}, 0\},$$
$$pb^r_{HL} - c^r_L < \max\{pb^r_{HL} + (1-p)\,b^r_{LL}, 0\}.$$

Only high-type senders value matching with a high-type receiver enough to send a sophisticated signal:

$$qb^s_{HH} \geq c^s \geq qb^s_{LH}. \tag{2}$$

Proof. See appendix C for all proofs. □

The second class of equilibria we consider are "common separating" equilibria: the sender sends a common signal if he is a high type and no signal if he is a low type; the receiver–regardless of her type–does not invest and accepts the sender only if he sends the common signal.

Proposition 2. *Assume* $b^r_{LH} \neq 0$ *and* $b^r_{LL} \neq 0$[2]*. Then there is a Common Separating Equilibrium if and only if any one of the following holds:*

(CSE1) If both high- and low-type receivers prefer to reject than accept low-type senders (1), and low type senders prefer NS to CS (2):

[1] These conditions are of measure zero

[2] These conditions are of measure zero.

$$b^r_{LH} < 0, b^r_{LL} < 0,$$
$$qb^s_{LH} + (1-q)b^s_{LL} \leq c^s \leq qb^s_{HH} + (1-q)b^s_{HL}.$$

(CSE2) If only high type receivers prefer to reject than accept low-type senders (1), it must be that low type senders prefer NS to CS (2):

$$b^r_{LH} < 0, b^r_{LL} > 0,$$
$$qb^s_{LH} \leq c^s \leq qb^s_{HH}.$$

If $b^r_{LH} > 0$ and $b^r_{LL} > 0$, there is no Common Separating Equilibrium. Both receivers will accept NS and CS; since CS imposes an additional cost, there will be incentive for both types of senders to deviate to NS.

Finally, there are “pooling" equilibrium in which senders do not send either signal.

Proposition 3. *If assumptions A1 - A4 are satisfied, there is always a pooling equilibrium where NS is sent, and where receivers do not invest.*

Fig. 2 represents the regions where CSE or SSE are equilibria, as a function of the parameters of our model. It can be used to illustrate some key results that we will revisit when applying our model to recent fashion trends. First, consider what happens when the cost of the signal, c_s is high: then CSE is an equilibrium. What happens if c_s falls? Eventually, CSE ceases to be an equilibrium, but SSE may still be an equilibrium (Fig. 3).

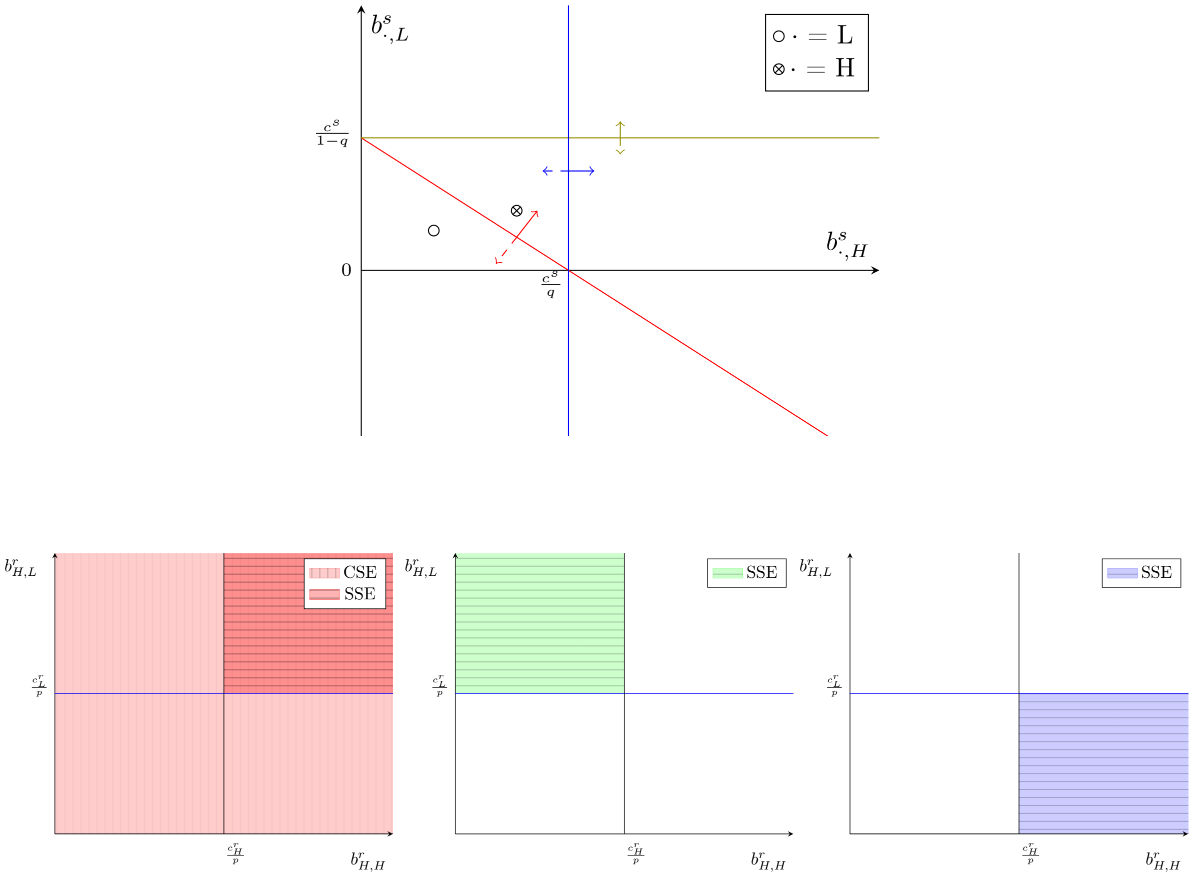


Figure 2: When are there sophisticated signaling and common signaling equilibria?

The diagram is read in two steps.

Top: We start by examining senders' payoffs from matching. On the x-axis, we represent the payoff to matching with a high type, $b^s_{\cdot,H}$. On the y-axis we represent the payoff to matching with a low type, $b^s_{\cdot,L}$. We use a dark circle to represent a high-type sender, and an open circle to represent a low-type sender. For there to be a separating equilibrium of any kind, the two types of senders must have sufficiently different preferences, namely, they must lie on opposite sides of one of three thresholds: the red line, the blue line or the green line. If, for at least one of these thresholds, the dark circle is above the threshold and the open circle is below the threshold, it is possible to have a separating equilibrium, and we continue to second row.

Second Row: Next we consider the high-type receivers' payoffs from matching, using one or more of the three graphs, depending on what happened in the top diagram. If, in the top diagram, the senders' preferences lay on opposite sides of the red threshold, we use the leftmost graph. If the senders' preferences lay on opposite sides of the green threshold, we use the middle graph. And if the senders' preferences lay on opposite sides of the blue threshold, we use the rightmost graph. For example, in the top diagram, we have presented senders' preferences such that they lie on opposite sides of the red threshold, but on the same side of the yellow and blue threshold. Therefore, in the second row, we consider only the leftmost graph. Then, all three graphs are read the same way. On the x-axis, we represent the high-type receivers' payoffs from matching with a high-type, $b^r_{H,H}$. On the y-axis, we represent the low-type receivers' payoffs from matching with a high type, $b^r_{H,L}$. The parameter regions in which CSE and SSE are equilibria are shaded.

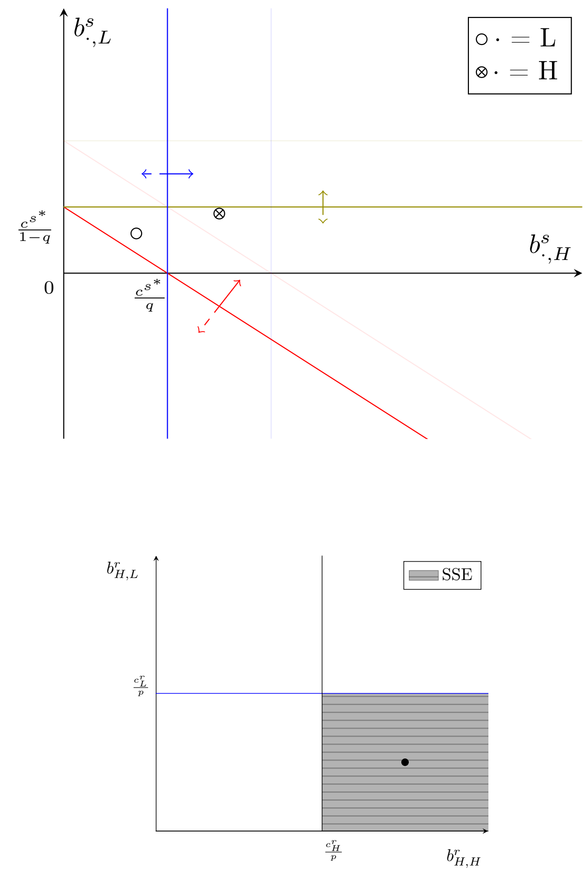


Figure 3: Moving from costly signaling to sophisticated signaling as incomes rise

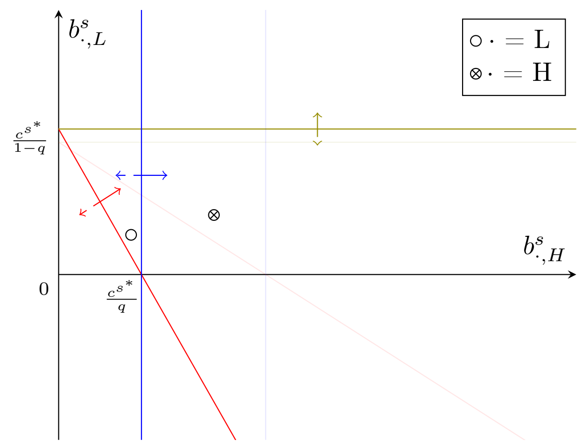
We start, as we did in Fig. 2, in the costly signaling equilibrium: senders' preferences are on opposite sides of the red threshold and we consider the leftmost graph in the second row. As c^s falls, the thresholds shift down and left, so that senders' preferences both lie on the same side of the red threshold, and instead might lie on opposite sides of the yellow or blue thresholds. Then, costly signaling is no longer be an equilibrium, but sophisticated signaling may be.

Next, consider what happens when the proportion of high-type receivers, q, grows. Eventually, CSE ceases to be an equilibrium but SSE may still be one (Fig. 4).

Finally, consider what happens when the benefit to a high type sender from matching with a high type (i.e. homophily) increases. In the model, we represent this by an increase in $B^s_{H,H}$ and/or a decrease in $B^s_{H,L}$. Again, CSE can cease to be an equilibrium, while SSE may be an equilibrium (Fig. 5).

3 Dynamics

We next explore the evolutionary dynamics of the sophisticated signaling game. There are two justifications for this analysis. First, we typically do not deliberately choose which fashions to adopt. Rather, we simply find certain clothes or accessories to be beautiful, or feel like they make us look good, perhaps



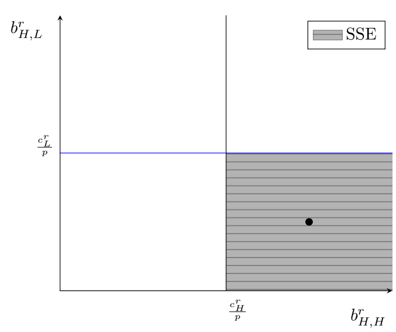


Figure 4: Moving from costly signaling to sophisticated signaling as the proportion of high type receivers grows

As q increases, the yellow threshold rises, the blue thresholds moves left, and the red threshold becomes steeper. This makes it less likely that costly signaling will be an equilibrium: in the second row, the leftmost graph ceases to relevant, and we consider instead the rightmost graph, where SSE may be an equilibrium.

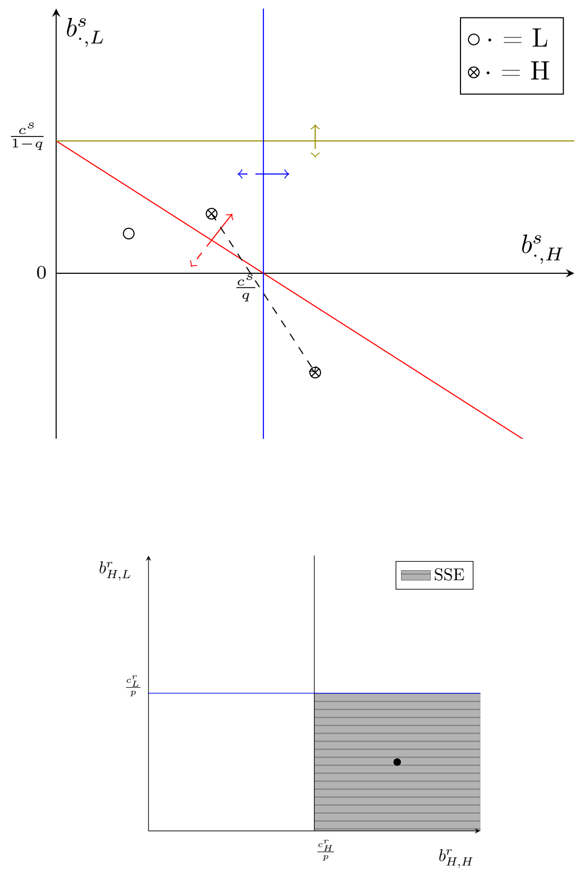


Figure 5: Moving from costly signaling to sophisticated signaling as the benefits to homophily increase.

What happens when the relative benefit for a high type to matching with another high type (i.e. homophily) is growing? In the model, we represent this by an increase in $b^s_{H,H}$ and/or a decrease in $b^s_{H,L}$. Then, the dark circle would move right and down, so that, again, costly signaling is no longer an equilibrium, but sophisticated signaling might be.

developing these tastes by admiring a popular classmate or movie star, or by internalizing compliments and jeers. Evolutionary dynamics are used to describe such settings because they are based on a single, believable criteria: strategies that are successful become more common, via biological evolution, learning, or imitation. Second, since pooling is always an equilibrium, and costly and sophisticated signaling sometimes overlap, it possible that sophisticated signaling is an equilibrium that never or rarely arises, rendering our analysis moot. We can use evolutionary dynamics to explore whether sophisticated signaling arises often in the parameter region where it is an equilibrium.

We employ the Wright-Fisher dynamic, a work-horses of the dynamics literature (Fisher 1930, Wright 1931, Nowak 2006)[3]. Roughly, the Wright-Fisher dynamic works as follows. We have with four populations, each of finite size N: high-type senders, low-type senders, high-type receivers, and low-type receivers. Senders choose from one of three strategies: not to signal, to send either a costly signal, or to send a sophisticated signal, $\{NS, CS, SS\}$. Receivers choose from one of twelve strategies: first choose whether to invest, $\{N, I\}$, and then whether to accept, $\{A, R\}$, but can condition this choice on the signal they observe. That is, receivers choose from the following strategies: N and always R, N and always A, N and R iff observe CS, N and A iff observe CS, I and always R, I and always A, I and A iff observe NS, I and A iff observe CS, I and A iff observe SS, I and A iff observe $\{NS, CS\}$, *etc.* We begin each simulation by randomly seeding each population with a mix of strategies. Then, at each 'generation':

1. Each player in a population plays the sophisticated signaling game against randomly chosen players from each of other populations. We calculate the player's expected payoff given her strategy and the frequencies of strategies in each of the other populations.

2. We assign strategies to the next 'generation' of players proportionally to payoffs (technically, according to the function $f(u) = e^{\eta u}$ where η is the selection strength and u is the expected payoff from a strategy).

3. A small proportion of individuals are randomly selected to mutate to a randomly selected strategy.

We repeat these steps many times, each time recording whether the composition of a particular set of strategies corresponds to any of the equilibria of the sophisticated signaling game (these are represented in the simplexes in Fig. 6 and 7; further details on classification of equilibria are available in appendix B):

1. SSE if all high-type senders send SS, all low-type senders send NS, high-type receivers invest (I) and accept SS (and sometimes also CS), low-type receivers accept CS

2. CSE if all high-type senders send CS, all low-type senders send NS, and both types of receivers do not invest (N) and accept CS

3. Pooling if all senders send NS and all receivers do not invest (N) and reject all and/or accept CS.

[3] Simulations were completed using the Python library DyPyFerdowsian (2016).

In the first simulation (Fig. 6) we vary the cost of sending the signal, c_s. CSE emerges frequently and SSE almost never when CSE is an equilibrium. Then, when c_s falls to the point where CSE ceases to be an equilibrium, SSE emerges frequently and CSE never.

In the second simulation, we select the same initial parameters, but instead vary $b^s_{H,L}$ from the region in which high-type senders prefer matching with low-type receivers over not matching, $b^s_{H,L} > 0$, and and both CSE and SSE are equilibria, to the region where high-type senders have an aversion to matching with low-type receivers, $b^s_{H,L} < 0$, and only SSE is an equilibrium. CSE emerges often and SSE almost not at all when CSE is an equilibrium. When CSE is not an equilibrium, it does not emerge. SSE does, but only once $b^s_{H,L}$ is sufficiently large. In the intermediate region (highlighted), only pooling emerges often.

4 Discussion

Our model provides a post-hoc explanation for understated luxury goods. One assumption of the model is that receivers can, with some investment, identify a sophisticated signal. We thus expect sophisticated signals to have identifying features which, while not as obvious as a "Gucci" logo, can be recognized by savvy observers. Consistent with this, many understated luxury goods do sport subtle distinguishing features. Prominent examples from recent years include: Burberry's Novacheck, Etro's colorful stripes, 7 For All Mankind's pocket stitching, and Bottega Veneta's basketweave.

Our model can also be used to understand why the trend has historically been from ostentatious costly signals to more subdued ones. For example, consider the shift from 'logo-mania' to 'logo-aversion' in China (Atsmon, Dixit and Wu 2011, Anaya 2013, Tse et al. 2013) In recent years, China, long a bastion of large logos and the best known brands, has experience a shift towards subtler designs and boutique brands, reminiscent of a shift that happened in the West some decades earlier. "Our groups are moving toward fewer logos, more discreet luxury... a luxury which is more subtle, more sophisticated..." the chairman of the group that owns Gucci said of the trend. "Luxury in China is now about being 'in the the know' versus 'in the show'," said another industry insider (Anaya 2013). The trend is perhaps exemplified by a 2013 campaign by luxury handbag maker Loius Vuitton, whose handbags and accessories typically feature large logos, often printed in repeating patterns over the entire item (see Fig. 8a). But in this campaign, featuring Chinese celebrities such as the actress Fan Bingbing pictured the company's handbags, there are no logos anywhere–not even on the bag–nor even any text to identify the ad as Louis Vuitton's (see Fig. 8b). The message of the ad echoes the second quote above: only those who are in the know can buy the product.

We propose that China has shifted from a costly signaling equilibrium to a sophisticated signaling equilibrium. Why the shift? One simple possibility is the remarkable rise in income. China's GDP per capita (current USD) grew from \$4,514 in 2010 to 7,593 in 2015. In the early-to-mid 1990s, the cost–

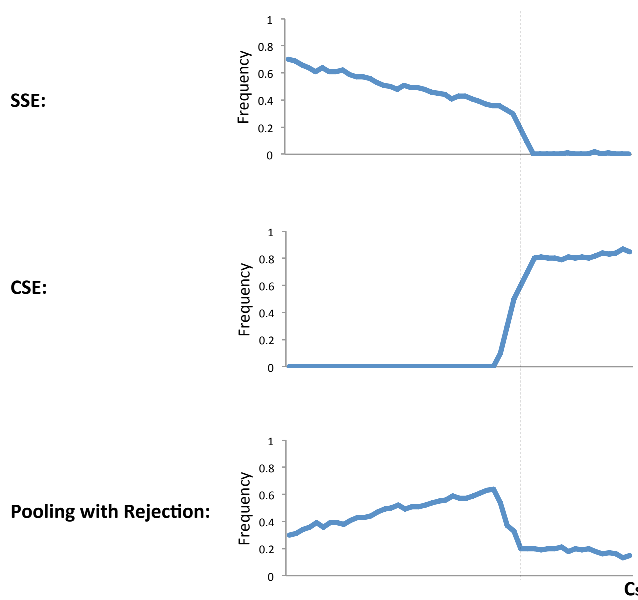


Figure 6: Frequency of equilibria as the cost of sending the signal, c_s, is varied

We randomly seed the strategy frequencies 100 times for 20 values of the parameter c_s, and allow the population to adjust according to a Wright-Fisher process for 5000 generations. After a 'burn-in' period of 500 generations, in each generation, we record whether the population is in one of the three equilibria: SSE, CSE, or pooling. For each value of c_s, we present the average frequency of each equilibrium, averaged over all 100 simulations. The dotted line corresponds to the value of c_s at which CSE ceases to be an equilibrium. CSE is frequent when it is an equilibrium, and non-existent when it is not. SSE occurs very infrequently when CSE is an equilibrium, but frequently in the region where CSE is not an equilibrium but SSE is an equilibrium. Pooling equilibria are present throughout.

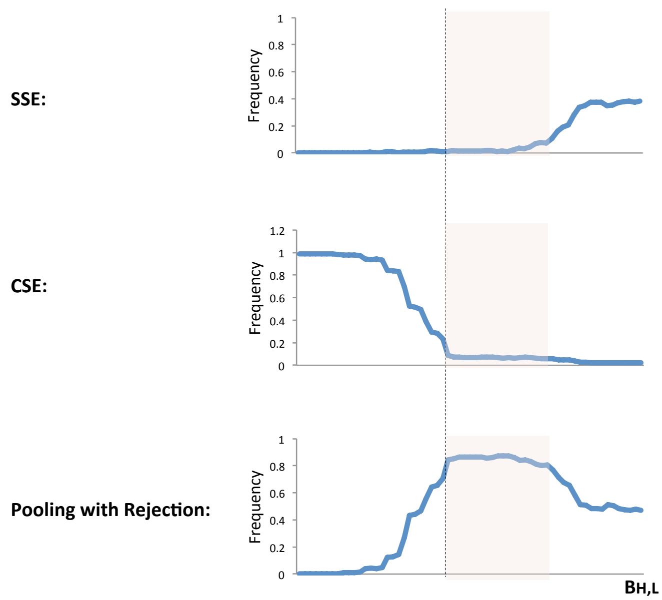


Figure 7: Frequency of equilibria as high-type senders' preference for low-type receivers is varied

We randomly seed the strategy frequencies 100 times for 20 values of the parameter $b^s_{H,L}$, and allow the population to adjust according to a Wright-Fisher process for 3500 generations. After a 'burn-in' period of 500 generations, in each generation, we record whether the population is in one of the three equilibria: SSE, CSE, or pooling. For each value of $b^s_{H,L}$, we present the average frequency of each equilibrium, averaged over all 100 simulations. The dotted line corresponds to the value of $b^s_{H,L}$ at which CSE ceases to be an equilibrium. CSE is frequent when it is an equilibrium, and non-existent when it is not. SSE occurs very infrequently when CSE is an equilibrium, but frequently in the region where CSE is not an equilibrium but SSE is an equilibrium. Pooling equilibria are present throughout. The shaded region corresponds to XXX.

(a) Early 2000s (b) 2013

Figure 8: Louis Vuitton Ads in China

relative to income–of purchasing luxury goods, which corresponds to c^s in the model, was relatively high. As we saw in Fig. 3 and Fig. 6, when c_s is high, CSE is an equilibrium and occurs frequently. But, once c_s falls sufficiently, CSE is not an equilibrium, SSE still is, and arises frequently.

Of course, there are other possibilities. A second possibility is that the proportion of high types, q, is growing, as China's middle class grows (see Fig. 4). And, a third possibility is that the relative benefits for a high type to matching with another high type (i.e. the benefits to homophily) are growing. In the model, we'd represent this by an increase in $b^s_{H,H}$ and/or a decrease in $b^s_{H,L}$ (see Fig. 5).

In addition to helping us better understand why and when people employ sophisticated signals, this paper makes a number of contributions. First, our paper expands on efforts (Veblen 1899, Charles et al. 2007) to understand fashion through the lens of incentives. Our paper further establishes that fashion is not only driven by design, aesthetics, or genius–incentives, too, determine which design elements succeed and persist. The benefits of this approach are clear: formal modeling of incentives yields a deeper understanding of fashion, and precise predictions about patterns and trends.

Second, our paper expands the existing signaling literature, further cementing the importance of signaling for understanding communication. Signaling has been used to explain elements of not only fashion, but also language (Gambetta 2009), religion (Veblen 1899, Sosis and Bressler 2003), *etc.* However, these applications have usually employed variants of Spence (1973)'s original costly signaling game (a notable exception is counter-signaling (Feltovich et al. 2001)), even though signaling is certainly a much richer phenomenon. Models such as ours will help establish it as such.

Finally, our paper contributes a growing literature that uses game theory to explore the origins of preferences and ideologies (Trivers 1971, Charles et al. 2007, Pinker, Nowak and Lee 2008, DeScioli and Wilson 2011, Hoffman, Hilbe and Nowak 2016b, Hoffman, Yoeli and Nowak 2015, Hoffman, Dalkiran and Yoeli 2016a). Like others in this literature, we eschew the strong assumption that agents are rational and choose their strategies deliberately, relying instead on the much weaker assumption that preferences and ideologies respond to incentives–those that are relatively successful become more common. Since, when this is true, a population typically stabilizes at a Nash equilibrium, this weaker assumption opens exciting possibilities for applying game theory to novel domains.

References

Anaya, Suleman, "Has Logo Fatigue Reached a Tipping Point?," http://www.businessoffashion.com/articles/intelligence/has-logo-fatigue-reached-a-tipping-point 2013. Accessed: January 25, 2016.

Atsmon, Yuval, Vinay Dixit, and Cathy Wu, "Tapping China's luxury-goods market," *McKinsey Quarterly*, 2011, pp. 1–5.

Bagwell, Laurie Simon and B Douglas Bernheim, "Veblen effects in a theory of conspicuous consumption," *The American Economic Review*, 1996, pp. 349–373.

Bearden, William O and Michael J Etzel, "Reference group influence on product and brand purchase decisions," *Journal of consumer research*, 1982, *9* (2), 183–194.

Berger, Jonah and Morgan Ward, "Subtle signals of inconspicuous consumption," *Journal of Consumer Research*, 2010, *37* (4), 555–569.

Charles, Kerwin Kofi, Erik Hurst, and Nikolai Roussanov, "Conspicuous consumption and race," Technical Report, National Bureau of Economic Research 2007.

DeScioli, Peter and Bart J Wilson, "The territorial foundations of human property," *Evolution and Human Behavior*, 2011, *32* (5), 297–304.

Feltovich, Nick, Rick Harbaugh, and Ted To, "Too cool for school? Signaling and countersignaling," *Signaling and Countersignaling (March 2001)*, 2001.

Ferdowsian, Andrew, "Using DyPy, A Powerful Python Library for Simulating Evolutionary Dynamics of Matrix-Form Games," https://dl.dropboxusercontent.com/u/22769456/DyPy_WhitePaper.pdf 2016. Accessed: February 21, 2016.

Fisher, Ronald Aylmer, *The genetical theory of natural selection: a complete variorum edition*, Oxford University Press, 1930.

Fudenberg, Drew and Jean Tirole, "Perfect Bayesian equilibrium and sequential equilibrium," *Journal of Economic Theory*, 1991, *53* (2), 236–260.

—— **and Lorens A Imhof**, "Imitation processes with small mutations," *Journal of Economic Theory*, 2006, *131* (1), 251–262.

Gambetta, Diego, *Codes of the underworld: How criminals communicate*, Princeton University Press, 2009.

Han, Young Jee, Joseph C Nunes, and Xavier Drèze, "Signaling status with luxury goods: The role of brand prominence," *Journal of Marketing*, 2010, *74* (4), 15–30.

Hoffman, Moshe, Aygun Dalkiran, and Erez Yoeli, "Why our norms are categorical," 2016. Harvard University Working Paper.

—— **, Christian Hilbe, and Martin Nowak**, "Why we obscure our positive traits and good deeds," 2016. Harvard University Working Paper.

—— **, Erez Yoeli, and Martin A Nowak**, "Cooperate without looking: Why we care what people think and not just what they do," *Proceedings of the National Academy of Sciences*, 2015, *112* (6), 1727–1732.

lun Tse, Ho et al., "This is not an LV bag: the simulacra of fashion in and beyond the media business in Hong Kong and mainland China," *HKU Theses Online (HKUTO)*, 2013.

Nowak, Martin A, *Evolutionary dynamics*, Harvard University Press, 2006.

Pinker, Steven, Martin A Nowak, and James J Lee, "The logic of indirect speech," *Proceedings of the National Academy of Sciences*, 2008, *105* (3), 833–838.

Sosis, Richard and Eric R Bressler, "Cooperation and commune longevity: A test of the costly signaling theory of religion," *Cross-Cultural Research*, 2003, *37* (2), 211–239.

Spence, Michael, "Job market signaling," *The Quarterly Journal of Economics*, 1973, pp. 355–374.

Trivers, Robert L, "The evolution of reciprocal altruism," *Quarterly review of biology*, 1971, pp. 35–57.

Veblen, Thorstein, *The theory of the leisure class*, Oxford University Press, 1899.

Wright, Sewall, "Evolution in Mendelian populations," *Genetics*, 1931, *16* (2), 97–159.

A Formal Definition of the Sophisticated Signaling Game

We model the sophisticated signaling game presented in section 1 as an extensive form game with incomplete information. The game is defined as follows:

- The set of players is $N = \{S, R\}$.
- The set of histories is:

$$\begin{aligned} H \quad = \quad & \{(L,L),(L,H),(H,L),(H,H), \\ & ((L,L),N),((L,L),I),((L,H),N),\ldots \\ & ((L,L),N,NS),((L,L),N,CS),\ldots \\ & ((L,L),N,NS,R),((L,L),N,NS,A),\ldots\} \end{aligned}$$

  . For example, (L, H) represents the history in which only chance has moved, and has chosen a low-type sender and a high-type receiver; $((L, H), N)$ represents the history in which both chance and the receiver have moved, chance has chosen (L, H) and the receiver has then chosen not to invest; $((L, H), N, CS)$ represents the history in which chance chooses (L, H), the receiver does not invest, and the sender chooses the common signal; and $((L, H), N, CS, A)$ is the history in which chance chooses (L, H), the receiver does not invest, the sender chooses the common signal, the receiver accepts.

- The set of terminal histories is $Z = \{((L,L),N,NS,R),((L,L),N,NS,A),\ldots\}$.
- The player function $P : h \in H \;\; Z \to N \cup \{c\}$ identifies the player who chooses an action after history h, where c represents chance. In this game, chance moves first, only the receiver moves after histories of length 1, only the sender moves after histories of length 2, and only the receiver moves after histories of length 3. The player function is thus:

$$P \quad = \quad \begin{cases} c & \text{if } h = \emptyset \\ R & \text{if } h \in \{(L,L),(L,H),(H,L),(H,H)\} \\ S & \text{if } h \in \{((L,L),N),((L,L),I),((L,H),N),\ldots\} \\ R & \text{if } h \in \{((L,L),N,NS,R),((L,L),N,NS,A),\ldots\} \end{cases}$$

- We denote by $H_i = \{h | P(h) = i\}$ the set of histories in which $i \in N \cup \{c\}$ moves.
- The action function $A(h)$ identifies the actions available to player $P(h)$ after history h. The action

function is:

$$A(h) = \begin{cases} \{(L,L),(L,H),(H,L),(H,H)\} & \text{if } h = \emptyset \\ \{N,I\} & \text{if } h \in \{(L,L),(L,H),(H,L),(H,H)\} \\ \{NS,CS,SS\} & \text{if } h \in \{((L,L),N),((L,L),I),((L,H),N),\ldots\} \\ \{R,A\} & \text{if } h \in \{((L,L),N,NS,R),((L,L),N,NS,A),\ldots\} \\ \{\ \} & \text{otherwise} \end{cases}$$

- The probability distribution $\mu : H_c \to \Re$ determines the likelihood with which chance chooses each action in H_c.

$$\mu = \begin{cases} (1-p)(1-q) & \text{if } h = (L,L) \\ (1-p)q & \text{if } h = (L,H) \\ p(1-q) & \text{if } h = (H,L) \\ pq & \text{if } h = (H,H) \end{cases}$$

- For each player $i \in N$, we define the partition Π_i of H_i that identifies what information i has when it is her turn to move. Specifically, Π_i partitions H_i into sets of states which i cannot tell apart:

$$\Pi_S = \left\{ \begin{array}{l} \{((L,L),N),((L,L),I),((L,H),N),((L,H),I)\}, \\ \{((H,L),N),((H,L),I),((H,H),N),((H,H),I)\} \end{array} \right\}$$

$$\Pi_R = \left\{ \begin{array}{l} \{(L,L),(H,L)\},\{(L,H),(H,H)\}, \\ \\ \{((L,L),N,NS),((H,L),N,NS),((L,L),N,SS),((H,L),N,SS)\}, \\ \{((L,L),N,CS),((H,L),N,CS)\}, \\ \\ \{((L,L),I,NS),((H,L),I,NS)\}, \\ \{((L,L),I,SS),((H,L),I,SS)\}, \\ \{((L,L),I,CS),((H,L),I,CS)\}, \\ \\ \{((L,H),N,NS),((H,H),N,NS),((L,H),N,SS),((H,H),N,SS)\}, \\ \{((L,H),N,CS),((H,H),N,CS)\}, \\ \\ \{((L,H),I,NS),((H,H),I,NS)\}, \\ \{((L,H),I,SS),((H,H),I,SS)\}, \\ \{((L,H),I,CS),((H,H),I,CS)\} \end{array} \right\}$$

- Players strategies are functions $s_i : H_i \to A(h)$, such that if $\Pi(h) = \Pi(h')$ then $s_i(h) = s_i(h')$.
- For each player $i \in N$, payoffs, $U_i : Z \to \Re$, are determined by their own type, whether they

matched, and, if so, the quality-type of their match. For senders, payoffs are also determined by whether they sent a signal, and for receivers, by whether they invested. Payoffs are:

$$u_S(h) = \begin{cases} 0 & \text{if } h \in \left\{ \begin{array}{l} ((L,L),N,NS,R),((L,H),N,NS,R),((L,L),I,NS,R),((L,H),I,NS,R), \\ ((H,L),N,NS,R),((H,H),N,NS,R),((H,L),I,NS,R),((H,H),I,NS,R) \end{array} \right\} \\ -c^s & \text{if } h \in \left\{ \begin{array}{l} ((L,L),N,CS,R),((L,H),N,CS,R),((L,L),I,CS,R),((L,H),I,CS,R), \\ ((L,L),N,SS,R),((L,H),N,SS,R),((L,L),I,SS,R),((L,H),I,SS,R), \\ ((H,L),N,CS,R),((H,H),N,CS,R),((H,L),I,CS,R),((H,H),I,CS,R), \\ ((H,L),N,SS,R),((H,H),N,SS,R),((H,L),I,SS,R),((H,H),I,SS,R) \end{array} \right\} \\ b^s_{L,L} - c^s & \text{if } h \in \{((L,L),N,CS,A),((L,L),I,CS,A),((L,L),N,SS,A),((L,L),I,SS,A)\} \\ b^s_{L,H} - c^s & \text{if } h \in \{((L,H),N,CS,A),((L,H),I,CS,A),((L,H),N,SS,A),((L,H),I,SS,A)\} \\ b^s_{H,L} - c^s & \text{if } h \in \{((H,L),N,CS,A),((H,L),I,CS,A),((H,L),N,SS,A),((H,L),I,SS,A)\} \\ b^s_{H,H} - c^s & \text{if } h \in \{((H,H),N,CS,A),((H,H),I,CS,A),((H,H),N,SS,A),((H,H),I,SS,A)\} \\ b^s_{L,L} & \text{if } h \in \{((L,L),N,NS,A),((L,L),I,NS,A)\} \\ b^s_{L,H} & \text{if } h \in \{((L,H),N,NS,A),((L,H),I,NS,A)\} \\ b^s_{H,L} & \text{if } h \in \{((H,L),N,NS,A),((H,L),I,NS,A)\} \\ b^s_{H,H} & \text{if } h \in \{((H,H),N,NS,A),((H,H),I,NS,A)\} \end{cases}$$

$$u_R(h) = \begin{cases} 0 & \text{if } h \in \{((L,L),N,NS,R),((L,H),N,NS,R),((H,L),N,NS,R),((H,H),N,NS,R)\} \\ -c^r & \text{if } h \in \{((L,L),I,NS,R),((L,H),I,NS,R),((H,L),I,NS,R),((H,H),I,NS,R)\} \\ b^r_{L,L} - c^r & \text{if } h \in \{((L,L),I,NS,A),((L,L),I,CS,A),((L,L),I,SS,A)\} \\ b^r_{L,H} - c^r & \text{if } h \in \{((L,H),I,NS,A),((L,H),I,CS,A),((L,H),I,SS,A)\} \\ b^r_{H,L} - c^r & \text{if } h \in \{((H,L),I,NS,A),((H,L),I,CS,A),((H,L),I,SS,A)\} \\ b^r_{H,H} - c^r & \text{if } h \in \{((H,H),I,NS,A),((H,H),I,CS,A),((H,H),I,SS,A)\} \\ b^r_{L,L} & \text{if } h \in \{((L,L),N,NS,A),((L,L),N,CS,A),((L,L),N,SS,A)\} \\ b^r_{L,H} & \text{if } h \in \{((L,H),N,NS,A),((L,H),N,CS,A),((L,H),N,SS,A)\} \\ b^r_{H,L} & \text{if } h \in \{((H,L),N,NS,A),((H,L),N,CS,A),((H,L),N,SS,A)\} \\ b^r_{H,H} & \text{if } h \in \{((H,H),N,NS,A),((H,H),N,CS,A),((H,H),N,SS,A)\} \end{cases}$$

- For each history $h' \in \Pi_i(h)$, $\beta_i(h'|\Pi_i(h))$ is player i's assessment that $h = h'$. We restrict

$$\sum_{h' \in \Pi_i(h)} \beta_i\left(h'|\Pi_i(h)\right) = 1$$

  and $\beta_i\left(h'|\Pi_i(h)\right) \geq 0$.

- For each history h, the function $Z(s|h) : s \to Z$ identifies the terminal history reached when players play s from h.

- Finally, we define players' expected utility at history h as:

$$U_i(s_s, s_r|\Pi_i(h)) = \sum_{h' \in \Pi_i(h)} \beta_i(h'|\Pi_i(h)) \cdot u_i(Z(s|h')).$$

B Categorizing Equilibria of the Sophisticated Signaling Game

Throughout, we will use the following equilibrium solution concepts.

Definition 1. *A* ***Perfect Bayesian Equilibrium*** *(PBE) is a pair of strategies and beliefs (σ, μ) such that:*

(PBE1) σ_i is optimal with respect to the beliefs and opponents' strategies at every information set

(PBE2) At information sets that are reached with positive probability, the beliefs agree with Bayes' rule

Definition 2. *A* ***Sequential Equilibrium*** *(SE) is a PBE such that (σ, μ) is consistent. That is, there exists a sequence σ^n of totally mixed strategy profiles such that $\sigma^n \to \sigma$ and the induced beliefs by Bayes' rule μ^n obey $\mu^n \to \mu$*

Next, we characterize when there are SE in which senders send common signals, sophisticated signals, or no signals, respectively.

Definition 3. *A Sequential Equilibrium (σ, μ) is a* ***Common Separating Equilibrium*** *(CSE) if the high type senders play CS and low type senders play NS.*

Strategy Profile 1. *Actions and beliefs in the CSE equilibria:*

1. *In each information set containing histories of length 1, receivers believe that senders are of high type with probability p, and do not invest.*

2. *In each information set containing histories of length 2, senders believe $((*, H), NI)$[4] with probability q and $((*, L), NI)$ with probability $1 - q$. High type senders send CS and low type senders send NS.*

3. *In each information set containing histories of length 3, receivers beliefs and actions are as follows:*

   (a) *For information sets containing histories of the form $((*, *), *, CS)$, place probability 1 that sender is high type and accept.*

   (b) *For information sets containing histories of the form $((*, *), I, NS)$, place probability 0 that sender is of high type. For the action, consider the payoff from being paired with a low-type sender (for example, the high-type receiver will consider b^r_{LH}). If it is positive, accept. If it is negative, reject. If it is zero, any mixed strategy is possible.*

[4]Throughout, we use the symbol $*$ as a placeholder for any action.

(c) *For the 4-element information sets containing histories of the form* $((*, *), NI, NS$ or $SS)$, *place probability 1 that sender is low type and NS is chosen. For the action, look at the respective utility of working with a low-type sender. If it is positive, accept. If it is negative, reject. If it is zero, any mixed strategy is possible.*

(d) *For information sets containing histories of the form* $((*, *), I, SS)$, *the belief can be anything. Let* $\alpha \in [0, 1]$ *be the belief that the sender is high-type (the same belief is used for both information sets). For the action, compute the respective* α*-weighted expected utility (for example, the high-type receiver will compute* $\alpha b^r_{HH} + (1 - \alpha) b^r_{LH}$*). If it is positive, accept. If it is negative, reject. If it is zero, any mixed strategy is possible.*

Definition 4. *A Sequential Equilibrium* (σ, μ) *is a* ***Sophisticated signaling Equilibrium*** *(SSE) if high type senders play SS and low type senders play NS.*

Note that the conditions on sender benefits are exactly the same as those for Common Separating Equilibria. Next, we present the strategies-beliefs profiles that yield these Sophisticated signaling Equilibria.

Strategy Profile 2. *Description of profile:*

1. *In each information set containing histories of length 1, receivers believe that senders are of high type with probability p, and invest.*

2. *In each information set containing histories of length 2, senders believe* $((*, H), I)$ *with probability* q *and* $((*, L), I)$ *with probability* $1 - q$*. High type senders perform SS and low type senders perform NS.*

3. *In each information set containing histories of length 3, receivers beliefs and actions are as follows:*

   (a) *For information sets containing histories of the form* $((*, *), I, SS)$, *place probability 1 that sender is high type, and perform accept.*

   (b) *For information sets containing histories of the form* $((*, *), I, NS)$, *place probability 0 that sender is high type, and perform reject.*

   (c) *For the 4-element information sets containing histories of the form* $((*, *), NI, NS$ or $SS)$, *place probability p that sender is high type sending SS and probability* $1 - p$ *that sender is low type sending NS. For the action, look at the respective p-weighted expected utility (for example, the high-type receiver will compute* $pb^r_{HH} + pb^r_{LH}$*). If it is positive, accept. If it is negative, reject. If it is zero, any mixed strategy is possible.*

(d) *For information sets containing histories of the form* $((*,*),*,CS)$*, the belief can be anything. Let* $\alpha \in [0,1]$ *be the belief that the sender is high-type (the same belief is used for all four information sets). For the action, compute the respective* α*-weighted expected utility. If it is positive, accept. If it is negative, reject. If it is zero, any mixed strategy is possible. To guarantee that this profile is a Sequential Equilibrium, we will set* α *sufficiently high (e.g. 1) so that the action is accept.*

Strategy Profile 3. *Description of profile:*

1. *In each information set containing histories of length 1, receivers believe that senders are of high type with probability* p*. High type receiver invests and low type does not.*

2. *In each information set containing histories of length 2, senders believe* $((*,H),I)$ *with probability* q *and* $((*,L),NI)$ *with probability* $1-q$*. High type senders perform SS and low type senders perform NS.*

3. *In each information set containing histories of length 3, receivers beliefs and actions are as follows:*

   (a) *For information sets containing histories of the form* $((*,*),I,SS)$*, place probability 1 that sender is high type, and perform accept.*

   (b) *For information sets containing histories of the form* $((*,*),I,NS)$*, place probability 0 that sender is high type. High-type receivers will perform reject, while low-type receivers will look at* b^r_{LL}*. If it is positive, accept. If it is negative, reject. If it is zero, any mixed strategy is possible.*

   (c) *For the 4-element information sets containing histories of the form* $((*,*),NI,NS$ or $SS)$*, place probability* p *that sender is high type sending SS and probability* $1-p$ *that sender is low type sending NS. For the action, look at the respective* p*-weighted expected utility (for example, the high-type receiver will compute* $pb^r_{HH} + pb^r_{LH}$*). If it is positive, accept. If it is negative, reject. If it is zero, any mixed strategy is possible.*

   (d) *For information sets containing histories of the form* $((*,*),*,CS)$*, the belief can be anything. Let* $\alpha \in [0,1]$ *be the belief that the sender is high-type (the same belief is used for all four information sets). For the action, compute the respective* α*-weighted expected utility. If it is positive, accept. If it is negative, reject. If it is zero, any mixed strategy is possible. To guarantee that this profile is a Sequential Equilibrium, we will choose* $\alpha = p$*, make receivers play the same strategy in* $((*,*),*,CS)$ *as in their 4-element information sets.*

Definition 5. *A Sequential Equilibrium* (σ,μ) *is a* ***Pooling Equilibrium*** *PE) if, under* σ*, the high type senders and low type senders choose the same action.*

Strategy Profile 4. *Description of profile:*

1. *In each information set containing histories of length 1, receivers believe that senders are of high type with probability* p*, and do not invest.*

2. *In each information set containing histories of length 2, senders believe* $((*, H), NI)$ *with probability* q *and* $((*, L), NI)$ *with probability* $1-q$*. Senders perform NS.*

3. *In each information set containing histories of length 3, receivers beliefs and actions are as follows:*

   (a) *For each two-element information sets, place probability* p *that sender is high type. For the action, look at the respective* p*-weighted expected utility (for example, the high-type receiver will compute* $pb^r_{HH} + pb^r_{LH}$*). If it is positive, accept. If it is negative, reject. If it is zero, any mixed strategy is possible, but to guarantee that this profile is a Sequential Equilibrium we will demand accept here.*

   (b) *For the 4-element information sets, place probability* p *that sender is high type sending NS and probability* $1-p$ *that sender is low type sending NS. For the action, look at the respective* p*-weighted expected utility. If it is positive, accept. If it is negative, reject. If it is zero, any mixed strategy is possible, but to guarantee that this profile is a Sequential Equilibrium we will demand accept here.*

C Proofs

C.1 Proof of Proposition 2 (CSE)

Proof. We proceed to solve for all possible Common Separating Equilibria under assumptions (A1)-(A4). We also assume $b^r_{LH} \neq 0$ and $b^r_{LL} \neq 0$.

C.1.1 Second period actions

By definition 2.1.1, we have to set high type senders to play CS and low type senders to play NS.

C.1.2 First period beliefs

For each of the first period information sets, we are forced by Bayes' rule to implement belief with probability p that sender is high type and $1-p$ that sender is low type.

C.1.3 Third period beliefs

The beliefs here are pinned down by (SE) to be those described in profile 2.1.3. For example, consider the information set $\{(HH, I, CS)\,,\ (LH, I, CS)\}$. Any in sequence σ^n of totally mixed strategy profiles where the sender strategies converge to those in subsubsection 2.2.1, the belief that sender is high type will approach 1. The same reasoning applies to all other third period information sets, to yield the beliefs described in profile 2.1.3. The only instance where (SE) does not pin down the beliefs at information sets containing histories of the form $((*, *), I, SS)$, and we have yet to shown that the beliefs in both information sets must be the same.

Suppose the high-type receiver has belief α that sender is high-type and the low-type receiver has belief β that sender is high-type. Consider any totally mixed strategy. In this strategy, let y_H and y_L be the respective probability that the high and low type senders send SS. When a high-type receiver sees $((*, H), I, CS)$, the induced belief that sender is high type is $py_H/(py_H + (1-p)y_L)$. When a low-type receiver sees $((*, L), I, CS)$, the induced belief that sender is high type is also $py_H/(py_H + (1-p)y_L)$. Thus, if (SE) were to hold and we exhibit σ as a limit of totally mixed σ^n, then it must be that the beliefs in both information sets are the same.

C.1.4 Third period actions

The third period beliefs pin down third period actions to be those described in profile 2.1.3 - simply compare the utility of accept or reject given the beliefs.

C.1.5 First period actions

It is optimal for receiver to NI because both I and NI lead to same period 2 and 3 paths, but investing incurs a cost.

C.1.6 Second period beliefs

This is pinned by Bayes' rule. For each of the second period information sets, belief with probability 0 that receiver is high/low and invested, probability q that receiver is high type and not invested, probability $1-q$ that receiver is low type and not invested.

C.1.7 Consistency of beliefs

At this point, we have shown that any candidate profile for Common Separating Equilibrium must be that described in profile 2.1.3. It remains to check that the profile is a Sequential Equilibrium. (PBE2) is true by construction, so next we check that (SE) holds. We define a sequence of totally mixed strategies σ^n as follows:

- Receivers invest with probability $\frac{1}{n}$,

- High-type senders do SS with probability $\frac{\alpha}{np}$, NS with probability $\frac{1}{n}$ and CS with probability $1-\frac{\alpha}{np}-\frac{1}{n}$. Low-type senders do SS with probability $\frac{1-\alpha}{n(1-p)}$, NS with probability $1-\frac{1-\alpha}{n(1-p)}-\frac{1}{n}$ and CS with probability $\frac{1}{n}$,
- Receivers accept with probability $\frac{n-1}{n}$ at nodes where σ says accept, and accept with probability $\frac{1}{n}$ at nodes where σ says reject (if σ says mixed, then use the same probability as in σ).

It is a direct check that $\sigma^n \to \sigma$ and that the induced beliefs $\mu^n \to \mu$.

C.1.8 Final step

It remains to check that (PBE1) holds. Since the actions of first and third periods were already chosen to be optimal, the only possible violation of (PBE1) is non-optimality of second period actions.

First check the optimality of high-type sender. $CS \succeq NS, SS$. The utilities from playing the respective actions are:

$$\begin{aligned}
CS &: qb^s_{HH} + (1-q)\, b^s_{HL} - c^s, \\
NS &: \mathfrak{a}_H q b^s_{HH} + \mathfrak{a}_L (1-q)\, b^s_{HL}, \\
SS &: \mathfrak{a}_H q b^s_{HH} + \mathfrak{a}_L (1-q)\, b^s_{HL} - c^s
\end{aligned}$$

where $\mathfrak{a}_i = 1$ if $b^r_{Li} > 0$ and $\mathfrak{a}_i = 0$ if $b^r_{Li} < 0$ for $i \in \{H, L\}$ (think of $\mathfrak{a}_i$ as "indicator for acceptance"). We want $c^s \leq (1-\mathfrak{a}_1)\, qb^s_{HH} + (1-\mathfrak{a}_2)(1-q)\, b^s_{HL}$.

As for the low-type sender, we want $NS \succeq CS, SS$. The utilities from playing the respective actions are:

$$\begin{aligned}
CS &: qb^s_{LH} + (1-q)\, b^s_{LL} - c^s, \\
NS &: \mathfrak{a}_H q b^s_{LH} + \mathfrak{a}_L (1-q)\, b^s_{LL}, \\
SS &: \mathfrak{a}_H q b^s_{LH} + \mathfrak{a}_L (1-q)\, b^s_{LL} - c^s.
\end{aligned}$$

We want $(1-\mathfrak{a}_1)\, qb^s_{LH} + (1-\mathfrak{a}_2)(1-q)\, b^s_{LL} \leq c^s$.

Thus, the strategy profile is a PBE iff

$$(1-\mathfrak{a}_1)\, qb^s_{LH} + (1-\mathfrak{a}_2)(1-q)\, b^s_{LL} \leq c^s \leq (1-\mathfrak{a}_1)\, qb^s_{HH} + (1-\mathfrak{a}_2)(1-q)\, b^s_{HL}.$$

Substituting the various values for $\mathfrak{a}_H, \mathfrak{a}_L$ depending on what b^r_{LH}, b^r_{LL} are, we obtain the various conditions of proposition 2.1.2. This completes the proof. □

C.2 Proof of Proposition 1 (SSE)

Proof. We proceed to solve for all possible Sophisticated signaling Equilibria under assumptions (A1)-(A4). For simplicity, we also ignore the knife-edge cases where

$$pb^r_{HH} - c^r_H = \max\{pb^r_{HH} + (1-p)\, b^r_{LH}, 0\},$$
$$pb^r_{HL} - c^r_L = \max\{pb^r_{HL} + (1-p)\, b^r_{LL}, 0\}.$$

C.2.1 Case 0

Suppose

$$pb^r_{HH} - c^r_H < \max\{pb^r_{HH} + (1-p)\, b^r_{LH}, 0\},$$
$$pb^r_{HL} - c^r_L < \max\{pb^r_{HL} + (1-p)\, b^r_{LL}, 0\}.$$

We solve for all possible equilibria solutions here.

Consider the high-type receiver. If she invests, she incurs the cost of investing, gets to work with high-type senders, and perhaps with low-type senders (if the utility is non-negative). That is, she gets

$$pb^r_{HH} - c^r_H + \max\{(1-p)b^r_{LH}, 0\}.$$

If she does not invest, she decides whether to work with the average senderm and her utility is

$$\max\{pb^r_{HH} + (1-p)\, b^r_{LH}, 0\}.$$

By the conditions of case 0, she will not invest. A similar reasoning shows that the low-type receiver will not invest.

Then, the high-type sender will want to deviate from SS to NS because it will not change whether they get accepted or rejected, but they will save on the cost of sending the signal.

Thus, there is no equilibrium here.

C.2.2 Case 1

Suppose

$$pb^r_{HH} - c^r_H > \max\{pb^r_{HH} + (1-p)\, b^r_{LH}, 0\},$$
$$pb^r_{HL} - c^r_L > \max\{pb^r_{HL} + (1-p)\, b^r_{LL}, 0\}.$$

We solve for all possible equilibria solutions here. Note that these conditions imply $b^r_{LH} < 0$ and $b^r_{LL} < 0$.

To begin, the first period beliefs are pinned down by Bayes' rule to be those described in profile 3.1.4.

For (SE) to hold, almost all the third period beliefs are pinned down to be those described in profile 3.1.4. The only exception are the beliefs at the information sets containing histories of the form $((*, *), *, CS)$ which can be anything, but if (SE) were to hold, then the beliefs at all four information sets must be the same, arguing the same way as in subsubsection 2.2.3. Let the common belief be α, and keep in mind that we have yet to impose further restrictions on α.

The third period actions are thus pinned down to be those in profile 3.1.4, except at the information sets containing histories of the form $((*, *), *, CS)$ which is determined if we know α.

Let us now consider the first period actions. If she invests, she would get $pb^r_{HH} - c^r_H$. If she does not invest, she would get $\max\{pb^r_{HH} + (1-p)\, b^r_{LH}, 0\}$. By assumption on the conditions, she will invest. Similarly, the low-type receiver will invest. This is as stated in profile 3.1.4. Then, the second period beliefs are pinned down by Bayes' rule to be those in profile 3.1.4.

At this point, we have shown that any candidate profile for Sophisticated signaling Equilibrium under case 1 must be that described in profile 3.1.4, except the freedom on α. To check consistency of beliefs, use a sequence of totally mixed strategies σ^n similar to that of subsubsection 2.2.7:

- Receivers invest with probability $\frac{n-1}{n}$
- High-type senders do SS with probability $1 - \frac{\alpha}{np} - \frac{1}{n}$, NS with probability $\frac{1}{n}$ and CS with probability $\frac{\alpha}{np}$. Low-type senders do SS with probability $\frac{1}{n}$, NS with probability $1 - \frac{1-\alpha}{n(1-p)} - \frac{1}{n}$ and CS with probability $\frac{1-\alpha}{n(1-p)}$
- Receivers accept with probability $\frac{n-1}{n}$ at nodes where σ says accept, and accept with probability $\frac{1}{n}$ at nodes where σ says reject (if σ says mixed, then use the same probability as in σ)

The last thing to check is (PBE1), particularly the optimality of second period actions. For the high-type senders, the utilities are given by

$$\begin{aligned}
&SS : qb^s_{HH} + (1-q)\, b^s_{HL} - c^s, \\
&NS : 0, \\
&CS : f_1 \cdot qb^s_{HH} + f_2 \cdot (1-q)\, b^s_{HL} - c^s
\end{aligned}$$

where $f_1, f_2 \in [0, 1]$ depend on α affecting whether the receiver accepts or rejects at histories with CS.

For the low type senders,

$$\begin{aligned}
&SS : qb^s_{LH} + (1-q)\, b^s_{LL} - c^s, \\
&NS : 0, \\
&CS : f_1 \cdot qb^s_{LH} + f_2 \cdot (1-q)\, b^s_{LL} - c^s.
\end{aligned}$$

Comparing the requirements for SS and NS, we see that a necessary condition is

$$qb^S_{HH} + (1-q)\, b^S_{HL} \geq c^s \geq qb^s_{LH} + (1-q)\, b^s_{LL}.$$

We see that this is also sufficient for senders to not deviate to CS, by setting $\alpha = 1$ so $f_1 = f_2 = 1$.

In summary, under case 1, there is a Sophisticated signaling Equilibrium if and only if (SSE1) holds. Moreover, proposition 3.1.3 is proven.

C.2.3 Case 2

Suppose

$$pb^r_{HH} - c^r_H > \max\left\{pb^r_{HH} + (1-p)\, b^r_{LH}, 0\right\},$$
$$pb^r_{HL} - c^r_L < \max\left\{pb^r_{HL} + (1-p)\, b^r_{LL}, 0\right\}.$$

We solve for all possible equilibria solutions here. Note that these conditions imply $b^r_{LH} < 0$.

To begin, the first period beliefs are pinned down by Bayes' rule to be those described in profile 3.1.6. For (SE) to hold, the third period beliefs are pinned down to be those described in profile 3.1.6, except for the value of α. The third period actions are thus pinned down to be those in profile 3.1.6, except at the $((*,*),*,CS)$ information sets.

Let us now consider the first period actions. By the same argue as in cases 0 and 1 respectively, we see that the high type receiver invests and the low type does not. This is as stated in profile 3.1.6. Then, the second period beliefs are pinned down by Bayes' rule to be those in profile 3.1.6.

For consistency of beliefs, easily modify the sequence σ^n from case 1.

The last thing to check is (PBE1), particularly the optimality of second period actions. For high-type senders, the utilities are given by

$$SS : qb^s_{HH} + g \cdot (1-q)\, b^s_{HL} - c^s,$$
$$NS : g \cdot (1-q) b^s_{HL},$$
$$CS : f_1 \cdot qb^s_{HH} + f_2 \cdot (1-q)\, b^s_{HL} - c^s$$

where $f_1, f_2 \in [0,1]$ depend on α affecting whether the receiver accepts or rejects, and $g \in [0,1]$ depends on whether p causes the low-type receiver to accept or reject at the 4-element information set.

For the low-type senders,

$$SS : qb^s_{LH} + g \cdot (1-q)\, b^s_{LL} - c^s,$$
$$NS : g \cdot (1-q) b^s_{LL},$$
$$CS : f_1 \cdot qb^s_{LH} + f_2 \cdot (1-q)\, b^s_{LL} - c^s.$$

Comparing SS and NS, we see that a necessary condition for there to be an equilibrium under case 2 is

$$qb^s_{HH} \geq c^s \geq qb^s_{LH}.$$

We shall show that this is also sufficient. If we set $\alpha = p$ and hence we can set $f_2 = g$, then we claim that neither senders will deviate to CS. Indeed, CS is a poorer option compared to SS in each case.

In summary, under case 2, there is a Sophisticated signaling Equilibrium if and only if (SSE2) holds. Moreover, proposition 3.1.5 is proven.

C.2.4 Case 3

Suppose

$$pb^r_{HH} - c^r_H < \max\left\{pb^r_{HH} + (1-p)\, b^r_{LH}, 0\right\},$$
$$pb^r_{HL} - c^r_L > \max\left\{pb^r_{HL} + (1-p)\, b^r_{LL}, 0\right\}.$$

We solve for all possible equilibria solutions here. Note that these conditions imply $b^r_{LL} < 0$.

All the steps up to the checking of (PBE1) is a mirror of case 2 (with high type and low type receivers swapped), so we omit them. Now we check (PBE1). For high-type senders, the utilities are given by

$$SS : g \cdot qb^s_{HH} + (1-q)\, b^s_{HL} - c^s,$$
$$NS : g \cdot qb^s_{HH},$$
$$CS : f_1 \cdot qb^s_{HH} + f_2 \cdot (1-q)\, b^s_{HL} - c^s$$

where $f_1, f_2 \in [0, 1]$ depend on α affecting whether the receiver accepts or rejects, and $g \in [0, 1]$ depends on whether p causes the low-type receiver to accept or reject at the 4-element information set.

For the low-type senders,

$$SS : g \cdot qb^s_{LH} + (1-q)\, b^s_{LL} - c^s,$$
$$NS : g \cdot qb^s_{LH},$$
$$CS : f_1 \cdot qb^s_{LH} + f_2 \cdot (1-q)\, b^s_{LL} - c^s.$$

Comparing SS and NS, we see that a necessary condition for there to be an equilibrium under case 3 is

$$(1-q)\, b^s_{HL} \geq c^s \geq (1-q)\, b^s_{LL}.$$

We shall show that this is also sufficient. If we set $\alpha = p$ and hence we can set $f_1 = g$. First consider the high-type sender. If $f_2 = 1$, then SS and CS are indifferent. If $f_2 = 0$, then CS is dominated by

NS which is dominated by SS. Then any value $f_2 \in [0,1]$ leads to CS $\preceq$ SS. Next consider the low-type sender. If $f_2 = 0$, then CS is dominated by NS. If $f_2 = 1$, then CS is indifferent to SS which is dominated by NS. Then any value $f_2 \in [0,1]$ leads to CS $\preceq$ NS.

In summary, under case 3, there is a Sophisticated signaling Equilibrium if and only if (SSE3) holds. Moreover, proposition 3.1.7 is proven. □

C.3 Proof of Proposition 3 (Pooling Equilibrium)

Proof. We can systematically check that profile 4.1.3 is indeed a Sequential Equilibrium. An intuitive proof is thus: in a situation where NS is sent, there is no differentiation between between high and low type senders. This implies there is no value in investing by receivers. Receivers and Senders hence choose actions that maximize their expected payoffs based on priors placed on the distribution of high and low types in the population.

Specifically, first period action is optimal because I and NI lead to the same outcome but the latter saves on the cost of investment. Second period action is optimal because SS, NS, CS all lead to the same outcome but NS saves on the cost of sending a signal. Third period action is optimal because they were decided based on comparing the utilities of accept and reject. Beliefs obey Bayes' rule wherever required. Finally, to check that (SE) holds, we define the following sequence of totally mixed strategies σ^n as follows:

- Receivers invest with probability $\frac{1}{n}$,
- Senders do SS with probability $\frac{1}{n}$, NS with probability $1 - \frac{2}{n}$ and CS with probability $\frac{1}{n}$,
- Receivers accept with probability $1 - \frac{1}{n}$ at nodes where σ says accept, and accept with probability $\frac{1}{n}$ at nodes where σ says reject.

It is a direct check that $\sigma^n \to \sigma$ and that the induced beliefs $\mu^n \to \mu$. □

D Evolutionary Dynamics

D.0.1 Classification of Outcomes

We classify as an equilibrium any state that is within a small tolerance (0.001) of the following conditions.

SSE: We categorize as SSE all the states in which all high type senders are playing *SS*, low type senders are playing *NS*, high type receivers are playing *(I, (A*$|SS$*, R*$|NS$*))*, and low type receivers are playing *(N, R)*.

CSE: We categorize as CSE all the states in which all high type senders are playing *CS*, low type senders are playing *NS*, high type receivers are playing *(N, (A*$|CS$*, R*$|NS$*))*, and low type receivers are playing *(N, (A*$|CS$*, R*$|NS$*))*.

Pooling with Rejection: We categorize as Pooling with Rejection all the states in which all high type senders are playing *NS*, low type senders are playing *NS*, high type receivers are playing *(N, R)*, and low type receivers are playing *(N, R)*.

Pooling with Rejection: We categorize as Pooling with Rejection all the states in which all high type senders are playing *NS*, low type senders are playing *NS*, high type receivers are playing *(N, A)*, and low type receivers are playing *(N, A)*.

D.0.2 Simulation Details

For the simulations presented in Fig. 6 of the manuscript, we employ the following parameters: $p = 0.3$, $q = 0.2$, $b^s_{LL} = 12$, $b^s_{LH} = 40$, $b^s_{HL} = 12$, $b^s_{HH} = 150$, $b^r_{LL} =$ -100, $b^r_{HL} = 50$, $b^r_{LH} =$ -100, $b^r_{HH} = 100$, $c^r_H = 10$, $c^r_L = 30$. We let c_s vary from 10 to 20, in intervals of 0.5. For each value of c_s, we perform 100 simulations of 5000 generations. In each simulation, we randomly seed the strategy frequencies. At each generation, strategies are assigned according to the formula, $f(u) = e^{\eta u}$ where η is the selection strength and u is the expected payoff from a strategy. Each player mutates with probability XXX. After 500 generations, in each generation, we record the whether population is in any of the equilibria identified in section D.0.1. We present the average frequency of the equilibria, averaged over all 100 simulations.

Notice that at $c_s = 20$, both CSE and SSE (see propositions 2 and 1):

$$\begin{aligned} b^r_{LH} &= -100 < 0, b^r_{LL} = -100 < 0, \\ qb^s_{LH} + (1-q)b^s_{LL} &= 0.2(40) + 0.8(12) = 17.6, \\ &\leq 20 = c_s \leq 0.2(150) + 0.8(12) = qb^s_{HH} + (1-q)b^s_{HL}. \end{aligned}$$

Once c_s decreases past $c_s = 17.6$, CSE is no longer an equilibrium. However, SSE remains as an equilibrium:

$$\begin{aligned} pb^r_{HH} - c^r_H = 20 &> \max\{pb^r_{HH} + (1-p)b^r_{LH} = -40, 0\}, \\ pb^r_{HL} - c^r_L = -15 &< \max\{pb^r_{HH} + (1-p)b^r_{LH} = -55, 0\}, \\ qb^s_{HH} &= 30 \geq c_s \geq 8 = qb^s_{LH} \text{ for } c_s \in [10, 20]. \end{aligned}$$

For the simulations presented in Fig. 7 of the manuscript, we employ the following parameters: $p = 0.2$, $q = 0.2$, $c_s = 18$, $b^s_{LL} = 12$, $b^s_{LH} = 30$, $b^s_{HH} = 150$, $b^r_{LL} =$ -100, $b^r_{HL} = 50$, $b^r_{LH} =$ -100, $b^r_{HH} = 100$, $c^r_H = 10$, $c^r_L = 30$. We let b^s_{HL} vary from 10 to -50 in intervals of 1. For each value of b^s_{HL}, we perform 100 simulations of 3500 generations. In each simulation, we randomly seed the strategy frequencies. At each generation, strategies are assigned according to the formula, $f(u) = e^{\eta u}$ where η is the selection strength and u is the expected payoff from a strategy. Each player mutates with probability XXX. After

500 generations, in each generation, we record the whether population is in any of the equilibria identified in section D.0.1. We present the average frequency of the equilibria, averaged over all 100 simulations.

Notice that at $b^s_{HL} = 10$, both CSE and SSE are equilibria (see propositions 2, 1):

$$\begin{aligned} b^r_{LH} &= -100 < 0, b^r_{LL} = -100 < 0, \\ qb^s_{LH} + (1-q)b^s_{LL} &= 0.2(30) + 0.8(12) = 15.6, \\ &\leq 18 = c_s \leq 0.2(150) + 0.8(10) = qb^s_{HH} + (1-q)b^s_{HL}. \end{aligned} \tag{3}$$

As $b^s_{H,L}$ decreases, CSE ceases to be an equliibrium. This takes place at $b^s_{H,L} = -15$. At that value, notice that equation (3) no longer holds. At this point, only SSE remains the viable equilibrium:

$$\begin{aligned} pb^r_{HH} - c^r_H &= 20 > \max\{pb^r_{HH} + (1-p)b^r_{LH} = -40, 0\}, \\ pb^r_{HL} - c^r_L &= -15 < \max\{pb^r_{HH} + (1-p)b^r_{LH} = -55, 0\}, \\ qb^s_{HH} &= 30 \geq c_s \leq qb^s_{LH} \text{ for } c_s = 18. \end{aligned}$$