Signaling in the Information Age:
How Increasingly Public Lives Affects the Cost of Signaling

B.D. Bernheim and A.L. Bodoh-Creed

January 15, 2019

Abstract

The advent of the internet, and in particular online social networks, has created an abundance of new opportunities for signaling. More abstractly, the environment we study has two features: first, receivers observe many choices by each sender, rather than a single decision; second, complete-information bliss points are heterogeneous across different types of senders. We prove that, ironically, a sufficiently large number of signaling opportunities allows senders to signal their true types at arbitrarily low overall cost. Instead of becoming ubiquitous, costly signaling becomes essentially irrelevant. We apply our results to signaling through online social networks, conspicuous consumption, and educational achievement.

1 INTRODUCTION

The internet, and in particular online social networks (OSNs) (e.g., Facebook, LinkedIn, Twitter and Weibo), allow us to lead more public, closely watched personal and professional lives.[1] In 2018, 80% of Pew survey respondents between the ages of 18 and 49 reported using Facebook in 2018, 51% of these users use the site several times a day, and an additional 23% use Facebook about once a day (Smith and Anderson [45]). This same group reported using OSNs about 6.5 hours per week (Neilsen [44]). There is a rapidly growing, interdisciplinary body of literature on how individuals use OSNs and the resulting effect on themselves and others, but this literature is largely outside of economics. Effectively using OSNs requires performing a "balancing act between self-expression and self-promotion" (van Dijck [14]),

[1] Our analysis will also apply to other online communities that allow for public announcements (e.g., Usenet newsgroups).

and online behavior can have professional consequences since 60% of employed Facebook users are “friends” with a coworker (Dourin et al [18]). It is well known that offline social networks play an important role in finding employment (e.g., Granovetter [29]), and many of the properties of offline social networks also apply to OSNs (Gee, Jones, and Berk [25], Garg and Telang [24]).[2] Moreover, 94% of recruiters used online social networks such as LinkedIn for recruiting in 2013 (Jobvite [34]). The employment ramifications are sensible given that psychologists have shown that social network profiles carry information about employment relevant personality characteristics (e.g., Buffardi and Campbell [11], Kluemper and Rosen [35]).

Our goal is to provide new insights into the welfare effects of using OSNs by treating an OSN platform as a channel for signaling. The crucial difference between our model and previous signaling theories is that we assume that OSNs allow users to send signals both more frequently and to a much broader audience than before. To be clear, we assume that OSNs convey information about potentially costly actions that are taken offline and serve as signals—in other words, it is the costliness of the offline activity that lends credibility to the information conveyed through the OSN.[3] For example, individuals can signal political preferences by posting photos of themselves attending political rallies, wealth through images from luxury vacations, or career success by promoting new research papers. The insights derived from signaling models may lead one to believe that many of the actions revealed on OSNs are distorted by the incentive to signal our types, and that the resulting pervasive waste must significantly erode the benefits of using these services. Our analysis suggests otherwise. We prove that under reasonably general conditions, a profusion of observable actions enables a sender (she) to signal her private information to a receiver (he) at negligible overall cost relative to the benefits of signaling. Ironically, the proliferation of opportunities to publicize signals via OSNs can therefore render the cost of signaling irrelevant rather than ubiquitous and lead people to behave as if private information were publicly observable. Therefore, the opportunity to lead a public life may actually reduce the cost of signaling information about ourselves credibly.

We first model a general setting in which a continuum of different types of senders have different bliss points, and deviations from bliss points entail costs.[4] We allow the

[2]There is also a literature in anthropology and sociology that uses impression management theory (Goffman [27]) to study how people interact through online social networks (e.g., boyd [10]).

[3]In analogy to models of signaling via the costly acquisition of an education, the OSNs play a role similar to a diploma in that they convey that the underlying costly activity, education, has been completed.

[4]Signaling models with heterogeneous bliss points are widely employed in the literature. Examples

sender to take a single action and include an abstract parameter λ that scales the total cost of deviating from her bliss point. The use of a weighting parameter allows us to easily illustrate the economic forces underlying our results in traditional single-action models before proceeding with our analysis of multi-action models. We prove that the ratio of the total costs of signaling to the total benefits declines to zero in the limit as λ diverges to infinity. Thus, in the limit, senders reveal their information costlessly.

In a model with a *finite* number of types and distinct bliss points, our result is easy to prove. For λ sufficiently large, no type would be willing to choose the bliss point of another type because the costs of doing so outweigh any potential benefit regardless of the receiver's inferences. This observation implies that for λ sufficiently large there exists an equilibrium wherein each sender-type chooses her bliss point, which means the sender can reveal her information without distorting her choices.

With a continuum of types, this argument no longer applies. It is straightforward to show, through similar reasoning, that each type's chosen action must converge to its bliss point as λ grows without bound. However, based on that observation alone, the total cost of signaling could in principle decline, remain constant, or grow depending on the speed of convergence. This raises the question of whether costless signaling in the finite-type case is an artifact of the model's assumption that types are discrete.

Our main result proves that costless signaling can occur in models with a continuous type space. Proving this requires new economic insights, however, and is not a result of technical assumptions. Our argument focuses on equilibria that satisfy the *dominance refinement*. First we define the *plausible set* for each action, which consists of the sender types for which the action can be rationalized by some receiver inference. The dominance refinement requires that the receiver's posterior beliefs place 0 probability on the sender having a type outside the plausible set for the observed action.[5] In particular, if a sender deviates to her bliss point and λ is large, the set of types in the bliss point's plausible set is small and contains types with nearby bliss points. Because the receiver's beliefs are restricted by the dominance refinement, the potential benefit for the sender of choosing

include Spence [49] on signaling with productive educational investments; Mailath [38] on price signaling; Banks [3] on political competition; Miller and Rock [42] on dividend signaling; Bernheim [6] on conformity; Bagwell and Bernheim [4], Ireland [33], and Corneo and Jeanne [13] on conspicuous consumption; Bernheim and Severinov [8] on bequests; Bernheim and Andreoni [2] on fairness; or Bernheim and Bodoh-Creed [7] on decisive leadership.

[5]This is a weak refinement in that it is the economic primitives of the model, and not endogenous features of the equilibrium, that limit the receiver's beliefs.

her equilibrium signal relative to her bliss point is small. But, if the benefits of signaling are small, then the cost must be small as well. We prove that the dominance refinement causes the signaling action to converge to the sender's bliss point quickly enough to cause the total cost of signaling to vanish at a rate of $O\left(\frac{1}{\sqrt{\lambda}}\right)$. If the separating equilibrium is continuous, we can strengthen this result to prove the total costs vanish at a rate of $O\left(\frac{1}{\lambda^{1-\alpha}}\right)$ for any $\alpha > 0$.

One interpretation of the abstract cost parameter is that it represents the number of actions the sender can use to signal her private information, which hints at our OSN application. First we study a model where the costs of signaling are additively separable and symmetric across N actions. We provide weak conditions under which the welfare-optimal separating equilibrium is one in which the sender chooses the same action repeatedly. Under those conditions, the optimal equilibrium is isomorphic to that of a model with a single signal where the weight attached to signaling costs is proportional to the number of actions. Applying our primary result to this setting, we conclude that the ratio of the costs of signaling to the benefits vanishes at a rate of $O\left(\frac{1}{N^{1-\alpha}}\right)$ for any $\alpha > 0$ if the equilibrium strategy is continuous.

We then study a general model that allows for nonseparable utility functions over the vector of N actions.[6] The challenge in this setting is that separating equilibria can take many forms, and it is difficult to provide a reason to select a particular equilibrium. In this general setting, we prove that the total costs of signaling vanish at a rate equal to $O\left(\frac{1}{\sqrt{N}}\right)$ if the receiver's beliefs satisfy the dominance refinement.

We provide three applications of our results. We first model the welfare effects of the increasingly public lives led by users of OSNs. Our model captures two forces that regulate signaling through OSNs. First, there is an increasing benefit to having a larger social network, particularly for tasks like job search where weak-ties matter. Second, we assume that the users have more opportunities to signal their type online. Our model implies that the total cost of signaling is the result of a race between the benefit of a larger social network, which raises the total cost of signaling, and the increased number of opportunities to signal, which lowers the total cost of signaling. Whether social networks exacerbate or eliminate the total cost of signaling will depend on which of these two forces is dominant.

Second, we consider a model of conspicuous consumption inspired by Ireland [33]. The avenues for signaling affluence have grown immensely over time. In the past wealth could

[6] Such a generalization is necessary to study applications like conspicuous consumption, where the sender's utility is defined over bundles of goods that may be substitutes for one another.

be signaled with fine clothes and jewelry or an expensive car, whereas now wealth can also be signaled through OSN posts describing a wide variety of expensive experiences (e.g., international vacations, expensive dinners). Although there are only a few studies on the relationship between OSN usage and conspicuous consumption, Thoumrungroje [52] provides survey evidence that increased social media use is correlated with an intensification of conspicuousness as a driver of consumption. Nonacademic survey data suggests that an increasing fraction of consumer spending is used for live events and that younger consumers more frequently post about live events to OSNs (Eventbrite and Harris [20]). This accords with evidence from the Bureau of Economic Analysis that spending on recreational services, food services, and accommodations doubled between 2002 and 2017, while spending on clothing increased by only 36%.[7] Since OSNs allow previously invisible consumption goods such as luxurious experiences to now be consumed conspicuously, one might have conjectured that welfare losses due to wasteful consumption are increasing. In contrast, our framework suggests that the total waste from conspicuous consumption is decreasing because people can conspicuously consume a bundle of goods closer to their bliss-point by deviating modestly across many goods instead of binging on a few.

Our final application is to the use of educational attainment to signal ability. The set of outlets for signaling ability to U.S. colleges includes an increasing number of standardized exams, extracurricular activities, and volunteer work. The "Turning the Tide" [30] report produced by the Harvard Graduate School of Education documents that college applicants exert large amounts of effort engaging in a variety of tasks with (what appears to be) the sole aim of improving their chances of admission to desirable schools. Education researchers have also shown that students in high performing schools suffer psychological costs during middle and high school (e.g., Luthar and Becker [37], Galloway et al. [23]), which suggests that the competitive pressures cause significant welfare losses. One might conclude from this that students would be better off in a system that based college admissions on a smaller number of objective measures (e.g., a universal standardized exam such as the Gaokao exam in China). However, our results imply the opposite—students are better off with a larger number of outlets for signaling and would suffer if these were restricted.

We are unaware of any other paper that has pointed out the relationship between heterogeneous bliss points, the number of observed signals, and the total cost of signaling. The economics literature on multidimensional signal focuses on providing conditions under

[7] We do not claim that this difference is entirely driven by conspicuous consumption.

which a separating equilibrium exists (e.g., Engers [19], Qunizii and Rochet [46], and Wilson [54]). The closest analogs to this paper are in the the theoretical biology literature on the "Handicap Principle," which implies that a costly signal is necessary to credibly convey information (Zahavi [55], Grafen [28]). Follow-up papers focused attention on the fact that the out-of-equilibrium cost of mimicking a signal is what makes the signal credible, which in turn led to the realization that signaling games with discrete types need not require costly equilibrium signals (Hurd [32], Számado [50] and [51]). This literature has suggested that signaling games with a continuum of types require costly signals (except for knife-edge examples) and that one should be pessimistic regarding the possibility of costless signals in generic settings with a continuum of types (Lachmann, Számado, and Bergstrom [5] and Bergstrom, Számado, and Lachmann [36]).[8]

There is also a small literature inspired by the economics and biology literature that studies signaling on OSN via the network of connections one has. Donath [15] and Donath and boyd [17] study social connections as a credible signal of identity in an otherwise anonymous virtual community, and Donath [16] draws out the implications of these ideas for OSN design. In contrast, our model of OSNs focuses on how these platforms allow a sender to reveal more offline actions to a larger set of receivers.

Section 2 introduces our benchmark model in which the sender's marginal cost of deviating from her bliss point is parameterized abstractly. Section 3 provides several examples of the benchmark model in order to build intuition for the first result, which Section 4 proves. Section 4 then provides results supporting our interpretation of the marginal cost parameter λ as the number of signals observed by the receiver , and Section 5 provides applications of our results . We close in Section 6. All proofs appear in Appendix A along with extensions of our results on multi-action signaling. Appendix B provides extra applications to job market signaling and politicians signaling decisiveness.

2 MODEL

The sender's private information is represented by her type $t \in [\underline{t}, \bar{t}] = T \subset \mathbb{R}$. For our benchmark model we assume the sender chooses an action $a \in \mathbb{R}_+$ that is observed by the receiver and serves as a signal used by the receiver to make inferences about the sender's

[8] In an evolutionary context, the preferences of the agents are endogenous. This means that the existence of a knife-edge example suggests selection pressures may push preferences to what would otherwise be a nongeneric case.

type. The sender's direct utility from action a in the absence of any signaling incentive (i.e., with complete information) is $\lambda\pi(a,t)$ where $\lambda > 0$ denotes a weight we use to scale the relative magnitude of the signaling incentives and the direct utility . The *bliss point* for an agent of type t is

$$a_{BP}(t) = \underset{a}{arg\,max}\ \lambda\pi(a,t). \tag{1}$$

$a_{BP}(t)$ is the action the sender would choose if her type were publicly observed . If the agents takes an action $a \neq a_{BP}(t)$ in equilibrium, then $\lambda\left[\pi(a,t) - \pi(a_{BP}(t),t)\right] < 0$ is the total cost of signaling.

Having observed a, the receiver uses Bayes's rule to form a posterior belief about the sender's type. The belief of the receiver following an observation of action a is denoted $\delta(a) \in \Delta(T)$ where $\Delta(T)$ is the set of Borel measures over T, and we refer to $\delta(a)$ as the receiver's *perception* of the sender. When we focus our analysis on fully separating equilibria, the receiver's equilibrium beliefs place probability 1 on the sender having the type $\widehat{t}(a)$, which is derived from the sender's strategy. When convenient we suppress the argument of δ and $\widehat{t}$ and refer to the sender as "choosing" the receiver's perception.

Given the receiver's perception, the sender receives benefits $B\left(t,\delta(a)\right)$.[9,10] Many signaling models include an endogenous response to the signal by the receiver, and $B\left(t,\delta(a)\right)$ is a reduced form representation of the sender's utility from the receiver's response to the signal. The sender's total utility is

$$U(a,t;\lambda) = B\left(t,\delta\left(a\right)\right) + \lambda\pi(a,t). \tag{2}$$

We make several assumptions that are easily verified in applications. We assume the existence of all referenced derivatives, and we use subscripts to denote partial derivatives with respect to the subscripted variable. For ease of notation, we first assume that each agent has a unique optimal choice under complete information.

Assumption 1. *$a_{BP}(t)$ is unique.*

The crux of our argument is an analysis of the costs of signaling when the equilibrium strategy approaches $a_{BP}(t)$. To provide a general analysis, we use polynomial approxima-

[9] It is straightforward to allow the action a to influence B. If we were to include a in B, then we would require that the partial derivatives $B_a(a,t,\delta)$, $B_{aa}(a,t,\delta)$, and $B_{aaa}(a,t,\delta)$ be defined for all (a,t,δ) and that $B_{aaa}(a,t,\delta)$ be uniformly bounded from above.

[10] In a complete information model, it is without loss of generality to normalize $B(t,t) = 0$ for all t. Since we are concerned with the signaling incentives, such a normalization is not innocuous .

tions of $\pi(a,t)$ around $(a_{BP}(t),t)$. Assumption 2 part (1) allows us to focus on second order effects and ignore 3^{rd} and higher order effects.[11] Part (2) of Assumption 2 extends the assumption of part (1) to cases where (for example) $\pi_{aa}=0$ and $\pi_{aaaa}<0$ — in other words, we can focus on the lowest relevant non-zero derivative and ignore higher order effects.[12]

Assumption 2. *There exists $C<\infty$ such that one of the following holds for all (a,t) in an open neighborhood of $\{(a,t): a=a_{BP}(t)\}$*

1. *For $a=a_{BP}(t)$ we have $\pi_{aa}(a,t)<0$ and $\pi_{aaa}(a,t)\le C$.*
2. *For all pairs $(a_{BP}(t),t)$ we have $\frac{\partial^i \pi(a,t)}{\partial a^i}=0$ for $i\in\{2,..,k<\infty\}$, $\frac{\partial^{k+1}\pi(a,t)}{\partial a^{k+1}}\neq 0$, and $\frac{\partial^{k+2}\pi(a,t)}{\partial a^{k+2}}\le C$.*

Our next assumption requires that first order changes in type result in first order changes in bliss points.

Assumption 3. *There exists $\beta>0$ such that for any $t>t'$, we have $a_{BP}(t)-a_{BP}(t')\ge \beta(t-t')$.*

Assumption 4 bounds the rate at which the sender's benefit can change with her perceived type in a fully separating equilibrium. Assumption 5 bounds the sender's benefit when the receiver is not confident about the sender's type, which may occur following a deviation from the equilibrium action by the sender. Together they imply that there exists $\underline{B}\le\overline{B}$ such that $B(t,\delta(a))\in[\underline{B},\overline{B}]$.

Assumption 4. *There is $\gamma>0$ such that $0\le B_{\hat{t}}(t,\hat{t})\le\gamma$.*

Assumption 5. *If the support of $\delta(a)$ is $\mathcal{S}$, then $\max_{\widehat{t}\in S} B(t,\widehat{t})\ge B(t,\delta(a))\ge \min_{\widehat{t}\in S} B(t,\widehat{t})$.*

Denote the strategy used in the separating equilibrium as $a_{SEP}(t;\lambda)$. We can write the equilibrium utility for an agent of type t who mimics the equilibrium action of an agent of type $\widehat{t}$ as

$$V(\widehat{t},t;\lambda)=B\left(t,\widehat{t}\right)+\lambda\pi\left(a_{SEP}\left(\widehat{t};\lambda\right),t\right).$$

[11] Since we are analyzing actions in the neighborhood of the bliss point, first order effects are absent as $\pi_a(a_{BP}(t),t)=0$.

[12] $\pi_{aa}(a_{BP}(t),t)=0$ implies $\pi_{aaa}(a_{BP}(t),t)=0$, since otherwise $a_{BP}(t)$ would not be a local maximum of $\pi(a,t)$. This point is elaborated in Footnote 24.

By standard arguments $a_{SEP}(t;\lambda)$ is the solution to the differential equation

$$\left.\frac{\partial a_{SEP}(\circ;\lambda)}{\partial \widehat{t}}\right|_{\widehat{t}=t} = \frac{-1}{\lambda}\left.\frac{\partial B}{\partial \widehat{t}}\right|_{(t,\widehat{t})=(t,t)} \left.\frac{1}{\pi_a(a,t)}\right|_{(a,t)=(a_{SEP}(t),t)}. \tag{3}$$

Since the receiver makes the worst possible inference for type $\underline{t}$ senders, senders of this type can do no worse than choose their bliss-point action. This implies that Equation 3 has the initial condition $a_{SEP}(\underline{t};\lambda) = a_{BP}(\underline{t})$. One can prove that this solution is, in fact, a separating equilibrium under relatively mild conditions (see Mailath and von Thadden [39]). To allow for greater generality, we directly assume that a separating equilibrium exists.

Assumption 6. *A fully separating Bayes-Nash equilibrium exists.*

3 EXAMPLES

We illustrate our framework through two examples. The first has heterogeneous bliss points, and we show that the difference between the bliss-point and the equilibrium action vanishes at the rate $O\left(\lambda^{-1}\right)$. The high speed of convergence of the equilibrium actions to the bliss points is crucial for the total cost of signaling to vanish.

Example 1. *Suppose $B(t,\widehat{t}) = \widehat{t}$, $\pi(a,t) = -(a-t)^2$, and $T = [0,1]$. The bliss points are (obviously) $a_{BP}(t) = t$. The ODE defining the fully separating equilibrium is*

$$\left.\frac{\partial a_{SEP}}{\partial \widehat{t}}\right|_{\widehat{t}=t} = \frac{1}{2\lambda(a_{SEP}(t)-t)}.$$

We use the change of variables $z(t) = a_{SEP}(t) - t$. Solving the inverse ODE yields

$$t = -\left[z + \frac{1}{2\lambda}\ln\left(\frac{1}{2\lambda} - z\right)\right] + C.$$

The initial condition $z(0) = 0$ implies

$$t = -\left[z + \frac{1}{2\lambda}\ln\left(1 - 2\lambda z\right)\right].$$

Reversing our change of variables and rearranging, we find

$$z(t) = \frac{1 - e^{2\lambda a_{SEP}(t)}}{2\lambda}.$$

The total cost of signaling is then

$$\lambda z(t)^2 \leq \lambda \left(\frac{1}{2\lambda}\right)^2 = \frac{1}{4\lambda},$$

which is $O\left(\lambda^{-1}\right)$ as claimed.

In the second example, the agents share the bliss point of $a_{BP} = 0$, $a_{SEP}(t)$ converges to $a_{BP}(t)$ at the rate $O\left(\lambda^{-0.5}\right)$, and the slow convergence causes the total cost of signaling to be bounded away from 0. Intuitively, the progressively greater bunching of low t types around 0 puts pressure on higher t types to choose higher actions. This pressure is absent in Example 1 since the different types are attracted to different bliss points. In fact, we show that while $a_{SEP}(t;\lambda) \to t$ as $\lambda \to \infty$, the total cost of signaling is invariant to λ.

Example 2. *Suppose $B(t, \widehat{t}) = \widehat{t}$, $\pi(a,t) = \frac{-a^2}{t+\gamma}, \lambda > 0$, and $T = [0,1]$. The bliss points are all $a_{BP}(t) = 0$, which means they are homogenous. The ODE defining the fully separating equilibrium is*

$$\left.\frac{\partial a_{SEP}}{\partial \widehat{t}}\right|_{\widehat{t}=t} = \frac{t+\gamma}{2\lambda a_{SEP}(t)}.$$

We can write this in a more convenient form as

$$2\lambda a_{SEP}(t) \left.\frac{\partial a_{SEP}}{\partial \widehat{t}}\right|_{\widehat{t}=t} = t + \gamma.$$

Integrating both sides and using our initial condition yields

$$\lambda a_{SEP}(t)^2 = \frac{1}{2}(t+\gamma)^2 - \frac{\gamma^2}{2}.$$

The total cost of signaling, $\lambda a_{SEP}(t)^2$, is invariant with respect to λ.

4 THEORETICAL RESULTS

4.1 Abstract Weight λ

Our first result proves that as $\lambda \to \infty$ the utility obtained by the sender in a separating equilibrium approaches the utility she receives with complete information when she chooses her bliss point. Recall that $\lambda[\pi(a_{SEP}(t;\lambda), t) - \pi(a_{BP}(t), t)]$ is the total cost of signaling, so our main result implies that the total cost vanishes in the limit as $\lambda \to \infty$. Because we hold the benefit of signaling fixed as λ grows, one can interpret this result as indicating that the total cost to benefit ratio declines to 0 as λ increases.

Some of our results refer to equilibria that satisfy a dominance refinement, which we now provide intuition for. For each sender type t and action a, the receiver can identify whether there is some posterior he could hold that would rationalize a sender of type t taking action a (relative to $a_{BP}(t)$). The set of such types is the *plausible set* for that action, and types outside of this set find taking a too costly regardless of the receiver's response. The refinement requires that the receiver places probability 1 on the sender having a type t inside the plausible set.[13]

Definition 1. *The **plausible set** for action a is $\mathcal{P}(a) = \{t : \underline{B} + \lambda\pi(a_{BP}(t), t) \geq \overline{B} + \lambda\pi(a, t)\}$. The receiver's beliefs satisfy the **dominance refinement** if the receiver's posterior places probability 1 on the sender having a type in $\mathcal{P}(a)$ after observing a.*

Before proceeding to our main result, we prove the following lemma. In addition to serving as the first step of the proof of our main result, the lemma is of additional interest since it characterizes *all* of the equilibria of our signaling model, not just the fully separating equilibrium referenced in our main result. Let $a(t;\lambda)$ denote an equilibrium of the signaling game, and let $\{a(t;\lambda_i)\}_{i=1}^{\infty}$, $\lambda_i \to \infty$, be a convergent sequence of equilibrium strategies. Lemma 1 implies that the limit must be $a_{BP}(t)$.

Lemma 1. *Let Assumptions 1, 2, and 5 hold, and assume $a_{BP}(t) > 0$ for $t > \underline{t}$. Then $\|a(t;\lambda_i) - a_{BP}(t)\| = O\left(\frac{1}{\sqrt{\lambda_i}}\right)$ as $\lambda_i \to \infty$.*

Lemma 1 distills a somewhat obvious truth: if the cost of deviating from the bliss-point increases while the benefit of signaling is bounded, then the deviations must shrink as the costs grow.[14] Lemma 1 also sheds light on the speed of convergence, which is less

[13]Our dominance refinement is similar to the one studied in Cho and Kreps [12].

[14]This result also implies that pools must vanish as λ_i increases.

obvious. Under mild technical conditions, $\lambda\left[\pi(a_{BP}(t),t)-\pi(a,t)\right]$ can be approximated by $\frac{-1}{2}\pi_{aa}\lambda(a-a_{BP}(t))^2$ for a sufficiently close to $a_{BP}(t)$, which must occur as $\lambda\to\infty$. Since the potential benefits of signaling are bounded from above by $\overline{B}-\underline{B}$ regardless of the inferences made by the receiver, we conclude that $\|a(t;\lambda_i)-a_{BP}(t)\|=O\left(\frac{1}{\sqrt{\lambda_i}}\right)$.

In a setting with a finite number of types with distinct bliss points, Lemma 1 implies that the total cost of signaling vanishes (i.e., $\lambda\left[\pi(a_{BP}(t),t)-\pi(a_{SEP}(t;\lambda),t)\right]=0$) for λ sufficiently large in *any* equilibrium that satisfies the dominance refinement. Example 3 describes the structure of such an equilibrium. For completeness, we also include an example that violates the dominance refinement and, as a result, has a nonvanishing total cost of signaling.

Example 3. *Suppose $B(t,\widehat{t})=\widehat{t}$, $\pi(a,t)=-(a-t)^2$, and $T=\{0,1\}$. The only separating equilibrium that satisfies the dominance refinement is $a_{SEP}(0;\lambda)=0$, $a_{SEP}(1;\lambda)=max\{\lambda^{-0.5},1\}$, and for the receiver to infer that the sender is of type $t=0$ if and only if $a<a_{SEP}(1;\lambda)$. For $\lambda\geq 1$ we have $a_{SEP}(1;\lambda)=1=a_{BP}(1)$ since the $t=0$ type is unwilling to choose $a_{BP}(1)$ even if it shifts the receiver's inference from placing probability 0 on the sender having type $t=1$ to placing probability 1 on the sender's type being $t=1$. Therefore the total cost of signaling vanishes for $\lambda\geq 1$.*

There also exists an equilibrium in which the total cost of signaling does not vanish. Suppose that $a_{SEP}(0;\lambda)=0$, $a_{SEP}(1;\lambda)=1+\lambda^{-0.5}$, and the receiver infers that the sender's type is $t=0$ following any deviation by the sender to $a<a_{SEP}(1;\lambda)$. In this case, the $t=1$ sender faces a cost of $-\lambda(a-1)^2=1$ for all values of λ. This equilibrium does not survive many refinements including our dominance refinement.

In contrast, for the typical model with a continuum of types, there exist no truth-telling equilibrium even for large λ. As a result, the existence of an equilibrium for which total signaling costs vanish as λ grows is far from obvious. Consider Example 1, the continuum analog of Example 3. Since the cost function is quadratic, if we had $a_{SEP}(1;\lambda_i)-a_{BP}(t)=\Theta\left(\frac{1}{\sqrt{\lambda_i}}\right)$, we would then have $\lambda\left[\pi(a_{BP}(t),t)-\pi(a_{SEP}(t;\lambda_i),t)\right]=\Theta(1)$—in other words, the costs would not vanish.[15] Thus, to prove our main result, we need to show that $\|a_{SEP}(t;\lambda_i)-a_{BP}(t)\|$ converges at a faster rate.

Before discussing our main theorem, we present a lemma showing that each agent takes an action above her bliss point. The logic is that as λ grows, agents must signal using

[15]The notation $f(x)=\Theta(g(x))$ means that there exists $k_1,k_2>0$ such that $k_1g(x)\leq f(x)\leq k_2g(x)$ as $x\to\infty$.

actions close to their bliss points. If $a_{SEP}(t;\lambda) < a_{BP}(t)$, then for large enough λ the separating action will equal the bliss point of some type $t' < t$, which gives the agent of type t' incentive to deviate to $a_{BP}(t') = a_{SEP}(t;\lambda)$.[16]

Lemma 2. *Let Assumptions 1, 3, and 4 hold. Then in any separating equilibrium we have for almost all t that $a_{SEP}(t;\lambda) \geq a_{BP}(t)$ for λ sufficiently large.*

We provide two versions of our first result. We state both in terms of the rate of convergence of $\lambda\pi(a_{SEP}(t;\lambda), t)$ to $\lambda\pi(a_{BP}(t), t)$, but one can also interpret them as implying that the ratio of the total costs to the total benefits of signaling vanishes. In the proof of each theorem we derive limits on the beliefs of the receiver following any sender action. In Theorem 1, the lower bound on the rate of change of $a_{BP}(t)$ (Assumption 3) allows us to use the dominance refinement to bound the inferences the receiver can make following a sender's choice of $a = a_{BP}(t)$. The bounds on the receiver's inferences combined with our bound on the rate of change of $B(t, \circ)$ (Assumption 4) yields an upper bound on the benefit of choosing $a_{SEP}(t;\lambda)$ relative to $a_{BP}(t)$ for senders with type t. The bound on the benefits of signaling one's type relative to taking the bliss point action implictly bounds the cost of signaling, and we show this bound implies that the total cost of signaling converges to 0 as $\lambda \to \infty$.

Theorem 1. *Let Assumptions 1 - 5 hold and assume the receiver's beliefs satisfy the dominance refinement. Then $\lambda\left[\pi(a_{BP}(t), t) - \pi(a_{SEP}(t;\lambda), t)\right] = O\left(\frac{1}{\sqrt{\lambda}}\right)$.*

The second result assumes that both the bliss points and the fully separating equilibrium are continuous, which allows us to ignore the issue of off-path beliefs when making arguments similar to those used in the proof of Theorem 1. The rate of convergence shown in Theorem 2 is faster than that of Theorem 1 because we can use the continuity of $a_{SEP}(t;\lambda)$ to iterate our bounding argument to obtain tighter bounds on $\lambda\left[\pi(a_{BP}(t), t) - \pi(a_{SEP}(t;\lambda), t)\right]$.

Theorem 2. *Let Assumptions 1 - 5 hold and suppose that $a_{BP}(t)$ and $a_{SEP}(t)$ are continuous for all λ. Then $\lambda\left[\pi(a_{BP}(t), t) - \pi(a_{SEP}(t;\lambda), t)\right] = O\left(\frac{1}{\lambda^{1-\alpha}}\right)$ for any $\alpha > 0$.*

4.2 Multiple Signals with Additive Aggregators

We now consider settings where the receiver observes multiple signals from the sender simultaneously. We consider a sequence of models indexed by $N \in \mathbb{N}$, where the N^{th}

[16] Recall that Assumption 4 requires $B(t, \hat{t})$ to be increasing in $\hat{t}$.

model features a sender who takes N simultaneous actions $a \in \mathbb{R}$ yielding an action vector $\boldsymbol{a}^N \in \mathbb{R}^N$ with the i^{th} component denoted a_i^N. The utility for the sender from these N actions is $\boldsymbol{\pi}^N(\boldsymbol{a}^N, t)$. Since we focus on separating equilibria, in equilibrium the receiver has a degenerate belief placing probability 1 on the sender's type being $\widehat{t}\left(\boldsymbol{a}^N\right)$, and the sender receives utility of $B\left(t, \widehat{t}\left(\boldsymbol{a}^N\right)\right) + \boldsymbol{\pi}^N(\boldsymbol{a}^N, t)$.

We start with the simplest case wherein $\boldsymbol{\pi}^N(\boldsymbol{a}^N, t)$ is additively separable and symmetric across actions, which allows us to characterize the sender-optimal equilibrium. In this model, the utility of the sender in the N^{th} model is

$$U_N(\boldsymbol{a}^N, t) = B\left(t, \widehat{t}\left(\boldsymbol{a}^N\right)\right) + \sum_{i=1}^{N} \pi\left(a_i^N, t\right).$$

Let $a_{BP}(t) = \underset{a}{arg\,max}\ \pi(a,t)$, so $\boldsymbol{a}_{BP}(t) = (a_{BP}(t), a_{BP}(t), ..., a_{BP}(t)) \in \mathbb{R}_+^N$ denotes the bliss point of the type t agent in the N-action model.

Consider the symmetric signaling equilibrium where $a_1(t) = ... = a_N(t) = a_{SEP}(t; N)$. We can write the equilibrium utility of an agent of type t that mimics the action of an agent of type $\widehat{t}$ as

$$V_N(\widehat{t}, t) = B\left(t, \widehat{t}\right) + N\pi\left(a_{SEP}\left(\widehat{t}; N\right), t\right).$$

The function $a_{SEP}(t; N)$ is defined by the following differential equation

$$\left.\frac{\partial a_{SEP}}{\partial \widehat{t}}\right|_{\widehat{t}=t} = \frac{-1}{N} \left.\frac{\partial B}{\partial \widehat{t}}\right|_{(t,\widehat{t})=(t,t)} \left.\frac{1}{\pi_a(a,t)}\right|_{(a,t)=(a_{SEP}(t;N),t)}$$

with the initial condition $a_{SEP}(\underline{t}; N) = a_{BP}(\underline{t})$.

We now provide conditions that imply the symmetric separating equilibrium is the welfare optimal separating equilibrium for all types, which motivates our focus on it.

Theorem 3. *Assume that $\pi(a,t)$ is supermodular in (a,t) and for all t and $a > a_{BP}(t)$ we have $\pi_a(a,t) \leq 0$ and $\frac{\pi_{at}(a,t)}{\pi_a(a,t)}$ is weakly increasing in a. For any fixed N, $\boldsymbol{a}_{SEP}^N$ maximizes the payoff of each type of sender relative to any other separating equilibrium.*

The supermodularity requirement is standard. The final two requirements of Theorem 3 are satisfied for common specifications such as $\pi(a,t) = -(a-t)^2$.

The N-signal setting with symmetric equilibrium $\boldsymbol{a}_{SEP}^N(t)$ is almost identical to the model of Section 4 where N plays the role of λ. The following corollary to Theorem 2 implies that the ratio of the costs of signaling to the benefits converges to 0 as $N \to \infty$ in the symmetric separating equilibrium.

Corollary 1. *Let Assumptions 1 - 5 hold and suppose that $a_{BP}(t)$ and $a_{SEP}(t;N)$ are continuous for all N. Then $N\left[\pi(a_{BP}(t),t) - \pi(a_{SEP}(t;N),t)\right] = O\left(\frac{1}{N^{1-\alpha}}\right)$ for any $\alpha > 0$*

Corollary 5 from Section 4.3 provides a more general result based on the dominance refinement that holds for any separating equilibrium in the additive model. However, the rate of convergence is only $O\left(\frac{1}{\sqrt{N}}\right)$.

4.3 Multiple Signals with General Aggregator Functions

It is natural to wonder whether our conclusions hinge on the additivity restriction or instead hold within a broader class of environments. In this section, we argue that versions of our results hold much more generally. As before, we consider a sequence of models, the N^{th} of which allows the sender to choose N actions. The utility of the sender is

$$U_N(\boldsymbol{a}^N, t) = B\left(t, \widehat{t}\left(\boldsymbol{a}^N\right)\right) + \boldsymbol{\pi}^N\left(\boldsymbol{a}^N, t\right).$$

We define $\boldsymbol{a}_{BP}(t)$ as

$$\boldsymbol{a}^N_{BP}(t) = \underset{\boldsymbol{a}^N \in R^N}{arg\,max}\ \boldsymbol{\pi}^N(\boldsymbol{a}^N, t). \tag{4}$$

Assumption 7 requires that an increase in t produces a non-trivial increase in all dimensions of the bliss-point actions, and it resembles Assumption 3 from the one-dimensional case.

Assumption 7. *$\boldsymbol{a}^N_{BP}(t)$ is the unique solution to (4). There exists a scalar $\beta > 0$ such that for all N and any $t > t'$, we have*

$$\boldsymbol{a}^N_{BP}(t) - \boldsymbol{a}^N_{BP}(t') \geq \beta(t - t')\mathbf{1}^N,$$

where $\mathbf{1}^N = (1, 1, ..., 1) \in \mathbb{R}^N$.

Assumption 8 is the multidimensional analog of Assumption 2. Part (1) insures that the Taylor series is well defined. Part (2) ensures that a small departure in the direction of $\mathbf{1}^N$ becomes increasingly costly, without bound, as N grows.[17] In particular, when combined

[17] Alternatively, we could provide an assumption along the lines of Assumption 2, Part (2), but it would be notationally intensive.

with Assumption 7, it implies that the costs of selecting the bliss point of an agent of type $t' > t$ for an agent with type t grows without bound as $N \to \infty$. Part (3) implies that we can neglect the third and higher order effects in the Taylor expansion for $\boldsymbol{a}^N > \boldsymbol{a}_{BP}(t)$ as N grows.

Assumption 8. *The following statements hold in some open neighborhood of* $\{(\boldsymbol{a}^N, t) : \boldsymbol{a}^N = \boldsymbol{a}^N_{BP}(t)\}$

1. *The second and third order partial derivatives of* $\boldsymbol{\pi}^N$ *exist and are bounded.*
2. *There exists a sequence* $\{\phi_N\}_{N=1}^{\infty}$ *such that* $\phi_N < 0$*,* $\phi_N \to -\infty$ *as* $N \to \infty$*, and*

$$\sum_{i,j=1}^{N} \frac{\partial^2 \boldsymbol{\pi}^N(\boldsymbol{a}^N, t)}{\partial \boldsymbol{a}_i \partial \boldsymbol{a}_j} < \phi_N.$$

3. *The third order effects are negligible relative to the second order effects in the sense that:*

$$\underset{N\to\infty}{lim\,sup} \frac{\sum_{i,j,k=1}^{N} \frac{\partial^3 \boldsymbol{\pi}^N(\boldsymbol{a}^N,t)}{\partial \boldsymbol{a}_i \partial \boldsymbol{a}_j \partial \boldsymbol{a}_k}}{\sum_{i,j=1}^{N} \frac{\partial^2 \boldsymbol{\pi}^N(\boldsymbol{a}^N,t)}{\partial \boldsymbol{a}_i \partial \boldsymbol{a}_j}} \leq 0.$$

The analysis of Section 4 leveraged the fact that the action space was one dimensional, which limits the kinds of equilibria we consider, and the focus on symmetric equilibria in Section 4.2 effectively reduced the problem to the one-dimensional model. The challenge here is that we need to prove our result when allowing for equilibria that entail signaling on all, some, or perhaps just one of the dimensions of the action space. Since signaling on one dimension of the action space might seem to be equivalent to a (traditional) single-action signaling model, it is not obvious that signaling costs vanish as the number of observed actions increases.

We show that costs must vanish because the sender always has the option of deviating to her bliss-point. Since $\boldsymbol{a}^N_{BP}(t)$ is strictly increasing (by Assumption 7), the dominance refinement pins down the receiver's beliefs after observing $\boldsymbol{a}^N_{BP}(t)$.[18] The restricted beliefs of the receiver bound the benefit of signaling one's type in equilibrium, which can be converted into a bound on the costs of signaling. Assumption 8, part 2, insures that the bound on the cost of signaling tightens as $N \to \infty$.

[18] The formal statement of the dominance refinement for the multidimensional case is the same as for the single dimensional case, except that one interprets the action a as a vector.

Theorem 4. *Let Assumptions 4, 5, 7, and 8 hold. In addition, assume* $\boldsymbol{a}^N_{SEP}(t) \geq \boldsymbol{a}^N_{BP}(t)$ *and that the receiver's beliefs satisfy the dominance refinement.*[19] *Then* $\boldsymbol{\pi}^N\left(\boldsymbol{a}^N_{BP}(t), t\right) - \boldsymbol{\pi}^N\left(a^N_{SEP,i}(t), t\right) = O\left(\frac{1}{\sqrt{-\phi_N}}\right)$.

It is straightforward to show that the additive model studied in Section 4.2 satisfies the assumptions of Theorem 4. Therefore, we have the following corollary that holds for any fully separating equilibrium of the additive model, symmetric or otherwise. We let $a^N_{SEP,i}(t)$ refer to the i^{th} component of an arbitrary fully separating equilibrium of the additive model.

Corollary 2. *Consider the additive model, and assume* $\pi(a,t)$ *satisfies assumptions 1 - 5 hold,* $\boldsymbol{a}^N_{SEP}(t) \geq \boldsymbol{a}^N_{BP}(t)$*, and the receiver's beliefs satisfy the dominance refinement. Then*

$$\sum_{i=1}^{N}\left[\boldsymbol{\pi}\left(a_{BP}(t), t\right) - \boldsymbol{\pi}\left(a^N_{SEP,i}(t), t\right)\right] = O\left(\frac{1}{\sqrt{N}}\right).$$

Theorem 4 does not hold universally in settings with multidimensional signals and heterogeneous bliss points. We demonstrate this point through a counterexample that violates Assumption 8, part 2, since the second order terms vanish as the sender's action approaches her bliss point.

Example 4. *Consider a utility function of the form*

$$U_N(\boldsymbol{a}^N, t) = B\left(t, \widehat{t}\left(\boldsymbol{a}^N\right)\right) - \prod_{m=1}^{N}\left(a_m - t\right),$$

where $t \in [0,1]$ *and* $a_i \geq 0$*. If we let* $B\left(t, \widehat{t}\right) = \widehat{t}$*, then the exact solution to the ODE defining the fully separating equilibrium is*

$$a_{SEP}(t) = t^{1/N} + t.$$

This implies that the cost of signaling is

$$\left(a_{SEP}(t) - t\right)^N = t.$$

[19] At the cost of much additional notation, we could alter Assumption 8, parts (2) and (3), to account for cases where $\boldsymbol{a}^N_{SEP}(t) \ngeq \boldsymbol{a}^N_{BP}(t)$. We would have to assume that the sum of the second and third order partial derivatives multiplied by the associated double and triple products of $a_{SEP,i}(t) - a_{BP,i}(t)$ have the appropriate sign and limit properties.

which is invariant with respect to N (and hence nonvanishing).

5 Applications

5.1 Social Networks

Online social networks (OSNs) provide opportunities to publicize professional accomplishments to signal one's ability, obtain information about job openings, or foster stronger-ties with colleagues. We focus on the first activity and how this signaling interacts with the scale of the social network.[20] Our goal is to study how the scale of network (i.e., the number of receivers that can see each action) interacts with the number of actions that can be observed by each receiver to determine the total cost of signaling.

We make use of the additive model of Section 4.2. The sender has N opportunities to signal, and we refer to each opportunity as a *project*. The set of projects an agent undertakes are the sort of tasks she would perform even in the absence of a social networking site. However, the networking site allows a sender to communicate the outcome or quality of more of these projects to receivers than she would be able to if the OSN were absent. The action a associated with a particular project represents the quality of the project, and the sender's type t represents the productivity of her effort. The direct utility of action a is $\pi(a, t)$. These direct payoffs could be from an intrinsic preference for high quality projects or from career benefits that derive from the quality of the project. However, these benefits must be independent of the inferences drawn about the receiver's type from the quality of the projects. We assume the user's bliss point $a_{BP}(t)$ is increasing in t. The increasing bliss-point requires that higher ability agents have higher marginal benefits from and/or lower marginal costs of producing project quality.

In a world without OSNs, some of these projects would gain wide publicity. For example, in an academic context, it is likely that most of one's academic peers would become aware of publications in major general interest journals. However, other accomplishments such as publications in field journals, grants, citations in the press, or teaching awards are less widely known (if at all) without an OSN like Twitter to publicize them. In a business context, it is likely that winning a major contract or a significant promotion would

[20] Social networks are potentially important conduits for information flow (i.e., the second activity), but no credibility is at stake in obtaining/giving information unless an incentive conflict is at work. The third goal, strengthening ties, could involve signaling commitment to the relationship, but the scale of the social network is not obviously relevant to fostering dyadic ties.

become widely known amongst one's professional peers, but completing a product update or internal promotions remain unknown outside of one's own firm.

The sender chooses to exert effort across R projects. We consider counterfactuals where the number of projects that receivers can observe, $N \leq R$, is increasing with the wide-spread adoption of OSNs. The major addition to the additive model studied above is that we now allow the benefit of signaling to depend on both the receiver's inferences about the sender and the number of social connections the sender has, which we denote F. We have in mind applications where both the perceived productivity and number of connections increase the benefit from signaling, $B(\widehat{t}, F)$. Assumption 9 implies that the marginal benefit of signaling is increasing in F.

Assumption 9. *There is a function $\gamma(F) > 0$ such that $0 \leq B_{\widehat{t}}(\widehat{t}, F) \leq \gamma(F)$.*

We now provide two results. The first applies to any fully separating equilibrium. In this case, a straightforward modification of the proofs of Theorem 1 and Corollary 5 yields the following result.

Theorem 5. *Consider the additive model, and assume $\pi(a, t)$ satisfies assumptions Assumptions 1 - 3 and 9 hold and the receiver's beliefs satisfy the dominance refinement. Then*

$$\sum_{i=1}^{N} \left[\boldsymbol{\pi} \left(a_{BP}(t), t \right) - \boldsymbol{\pi} \left(a^{N}_{SEP,i}(t), t \right) \right] = O\left(\sqrt{\frac{\gamma(F)^3}{N}} \right).$$

For equilibria that are symmetric across actions and continuous, a straightforward modification of the proofs of Theorem 2 and Corollary 1 yields the following result.

Theorem 6. *Let Assumptions 1 - 3 and 9 hold and suppose that $a_{BP}(t)$ and $a_{SEP}(t)$ are continuous for all N. Then*

$$N\left[\pi(a_{BP}(t), t) - \pi(a_{SEP}(t), t)\right] = O\left(\left[\frac{\gamma(F)^2}{N} \right]^{1-\alpha} \right)$$

for any $\alpha > 0$.

Despite the different rates of convergence, the message from these results is that the effect of living a more public life on the welfare loss from signaling is ambiguous since it depends on how F scales with N. For example, suppose that the sender is an academic whose primary professional connections are a small group of researchers working on tightly related topics. If the introduction of the social network causes the sender's existing social network to move onto the platform without adding additional connections, then the network will reduce the total cost of signaling as long as the network provides more signaling opportunities (e.g., opportunities to publicize the sender's nonacademic publications or discuss the sender's research in relation to current news). On the other hand, if the sender is able to widely expand her social network when she migrates to the platform (i.e., F increases), but the number of chances to send signals does not increase, then the social network increases the total cost of signaling. However, as long as the sender can choose F, then she can do no worse from moving online as she could always choose to not add connections and take advantage of the increased number of opportunities to signal.

5.2 Conspicuous Consumption

Online social networks have made it possible for senders to make visible previously invisible consumption. For example, it is easy to make the expense of a dinner or a vacation widely known through a Facebook or Twitter post, whereas credible signals of these expenditures are difficult to circulate widely without using OSNs. Inspired by Ireland [33], we now study a model wherein a sender's type t represents her budget, and the sender signals her type through conspicuous consumption of goods that are visible to external observers.

We consider a setting where the agent can purchase bundles of R different goods sold at prices $p_1, ..., p_R$. We denote the particular bundle purchased using a quantity vector $\boldsymbol{a}^R = (a_1, ..., a_R)$, and the first N of the goods are visible to receivers and the remaining $R - N$ are not visible. We assume the agent has a continuous, weakly monotone utility function $f^R(a_1, ..., a_R)$ over the bundles of goods. We let $\pi^R(a_1, ..., a_R, t) = f^R(a_1, ..., a_R) + g(t - \sum_{i=1}^R p_i a_i)$, which we interpret as the direct utility from consuming the potentially visible goods and a quasilinear term reflecting utility from other uses of the remaining budget. The bliss point for the sender solves the following optimization problem

$$\boldsymbol{a}^N_{BP}(t) = \underset{\boldsymbol{a}^R}{arg\,max}\ f^R(a_1, ..., a_R) + g\left(t - \sum_{i=1}^{R} p_i a_i\right).$$

We assume $\boldsymbol{a}_{BP}(t)$ is unique and that there exists β such that for $t > t'$ we have $\boldsymbol{a}_{BP}(t) - \boldsymbol{a}_{BP}(t') \geq \beta(t - t'))\mathbf{1}^R$, which requires that all of the potentially visible goods are normal goods. Finally, and unlike Ireland [33], we assume that the benefit of a receiver inferring the sender has type t is $B(\widehat{t})$.[21]

We have to use the results on general aggregator functions in Section 4.3 to analyze this model since consumption utility is not separable across goods. In this context, Assumption 8, part 1, is a technical assumption requiring that the utility function be well behaved. Assumption 8, part 2, is consistent with f and g being concave functions, as economic intuition would seem to require, but we must assume that the sum of the second-order terms diverges to $-\infty$. Assumption 8, part 3, is technical and requires that the third order terms of a Taylor expansion of π^R around $(\boldsymbol{a}_{BP}N(t), t)$ are insignificant relative to the second-order terms.

Our model is best interpreted as one in which the number of potentially visible goods, R, is large, and our results predict changes in the total cost of conspicuous consumption as a greater number of potentially visible goods become actually visible. Theorem 4 yields the following corollary, which implies that the welfare loss from conspicuous consumption in any fully separating equilibrium must vanish as the number of visible goods increases. In other words, as N increases and more of a person's consumption is observed, one can consume a bundle closer to one's ideal, suffer lower welfare losses, and still signaling affluence.

Corollary 3. *Let Assumptions 4, 5, 7, and 8 hold. In addition, assume $\boldsymbol{a}^N_{SEP}(t)$ is a separating equilibrium, $\boldsymbol{a}^N_{SEP}(t) \geq \boldsymbol{a}^N_{BP}(t)$, and that the receiver's beliefs satisfy the dominance refinement. Then*

$$\boldsymbol{\pi}^N\left(\boldsymbol{a}^N_{BP}(t), t\right) - \boldsymbol{\pi}^N\left(\boldsymbol{a}^N_{SEP}(t), t\right) = O\left(\frac{1}{\sqrt{-\phi_N}}\right).$$

.

5.3 University Applications

We now consider how our results play out in a model of signaling through educational attainment. A student's type t represents her underlying ability, and the vector of actions $\boldsymbol{a}$ are signals that the universities use to infer her ability. Examples include standardized exam scores, course grades, extracurricular activities, and pre-college employment. We assume that the direct payoffs from the activities are symmetric and additively separable,

[21] In the original model of Ireland [33], the signaling benefit is the inferred utility of an agent with type $\widehat{t}$. Therefore, the benefit is a function of the distorted consumption of visible goods. Footnote 9 describes the assumptions needed for our results to hold when the benefit function directly depends on the agent's action (i.e., $B(\widehat{t}, \boldsymbol{a}^R)$). We focus on this simpler model for expositional ease.

so $\pi(\boldsymbol{a}^N, t) = \sum_{i=1}^N \left[P(a_i^N - C(a_i^N, t)\right]$ where $P(a)$ is the direct benefit of the action and $C(a, t)$ is the direct cost. The bliss point for each action is

$$a_{BP}(t) = \underset{a}{arg\,max}\ P(a) - C(a, t).$$

We assume that $a_{BP}(t)$ is unique, $a_{BP}(\underline{t}) > 0$, and $C(a, t)$ is strictly supermodular in $(-a, t)$ so that $a_{BP}(t)$ is strictly increasing. To insure that the productivity effect of effort is nontrivial, we also assume there exists $\beta > 0$ such that for $t > t'$ we have $a_{BP}(t) - a_{BP}(t') \geq \beta(t - t')$. The sender's utility is

$$B(\widehat{t}(\boldsymbol{a}^N)) + \sum_{i=1}^{N} \left[P(a_i^N) - C(a_i^N, t)\right],$$

where for simplicity we have assumed the receiver's beliefs, $\widehat{t}(\boldsymbol{a}^N)$, are degenerate. We assume that

$$\frac{\pi_{at}(a, t)}{\pi_a(a, t)} = \frac{-C_{at}(a, t)}{P_a(a, t) - C_a(a, t)}$$

is weakly increasing in a, and from Theorem 3 we conclude that an equilibrium where the same action is taken across all N metrics is welfare optimal for the sender.

Letting $a_{SEP}(t)$ be the symmetric action taken across the N metrics, the following result follows from Corollary 1, which implies that the total cost of signaling vanishes as the number of human capital metrics increases.

Corollary 4. *Let Assumptions 1 - 5 hold and suppose that $a_{BP}(t)$ and $a_{SEP}(t; N)$ are continuous for all N. Then for any $\alpha > 0$ we have*

$$N\left[\pi(a_{BP}(t), t) - \pi(a_{SEP}(t), t)\right] = O\left(\frac{1}{N^{1-\alpha}}\right)$$

College admissions systems based solely or primarily on a single centralized exam are used in several countries including China (Gaokao), Turkey (YGS-LYS), and Brazil (ENEM).[22] We could model such an admissions policy in our setting by letting the students take N actions, but universities only observe a_1^N, which represents the outcome of the standardized test. Whatever effort students put into other activities is purely for their

[22]Centralized exams are usually implemented as a contest, so test-takers with good scores mechanically are admitted to better universities without the schools needing to make inferences. The issue of wasteful effort is well known in contest settings (Hoppe, Moldovanu, and Sela [31]).

direct benefits. In the equilibrium of this model, the students choose their bliss point action for the unobserved actions $2, ..., N$ (i.e., $a_i^N(t) = a_{BP}(t)$ for $i > 1$). The choice of the signal, a_1^N, is determined by the usual single-action separating equilibrium. In this case, the total cost of signaling does not vanish since there will alway be nontrivial distortions to the first action.

In contrast, college admissions in the United States are usually based on a holistic assessment of each applicant by each school. These holistic assessments integrate multiple aspects of achievement including exam scores, class grades, extracurricular activities, and personality. Corollary 6 implies that a strictly exam-based admissions system such as the Chinese Gaokao will have high welfare costs relative to a more holistic scheme. This does not necessarily mean the conclusion from studies like the "Turning the Tide" report ([30]) are wrong or that that students in the U.S. system are not suffering welfare losses. However, our results imply that encouraging colleges to look at fewer kinds of student accomplishments during the admissions process is counterproductive.

6 CONCLUSION

One of the many effects of living a public life through the internet and OSNs is that the number of opportunities to signal and the set of receivers that observe the signal have both grown. Given that signaling requires costly effort to credibly convey information, the prospect of sending many more signals through OSNs might suggest that welfare losses from wasteful signaling are growing. On the contrary, our results show that the abiltiy to signal through many channels simultaneously can actually mitigate welfare losses. Instead of increasing the cost of signaling, a public life may allow one to credibly reveal private information essentially costlessly.

Appendix B provides two additional applications. The first application studies the welfare cost to politicians from signaling decisiveness, which we study in the context of the model of Bernheim and Bodoh-Creed [7]. The past 20 years have brought major changes to the media landscape including the rise of OSNs such as Twitter, which let politicians directly communicate with voters, and 24 hour networks that focus on politics. We prove that as the number of channels for publicizing decisive actions increases, the hastiness of the politicians' actions decreases and the welfare of the politicians rises. The second application is to job market signaling, which reframes Section 5.3 to apply to signaling ability to employers through educational attainment. The conclusion is that employers can reduce the cost of signaling by holistically evaluating potential employees as opposed to (for example) focusing exclusively on course grades.

Finally, our results also have implications for models that do not result in fully separating equilibrium. For example, Bernheim [6] models social conformity as a partially pooling equilibrium wherein agents that sufficiently value social esteem pool on the bliss-point action of the most esteemed type. Banks [3] studies political conformity where politicians' political platforms pool on the median voter's preferred policy to signal that the politicians are moderates that have views similar to those of the median voter. The natural extension of these models to a multi-action environment would define a conformity equilibrium as one in which agents signal by conforming across most or all of the actions. Under this interpretation, Lemma 1 implies that the set of agents willing to conform shrinks as the number of actions increases.[23] This means that as the number of actions increases, conformity switches from being a tool that enforces social/political norms to one that identifies "true believers" in the ideal that is being conformed to.

References

[1] "2016 Nielsen Social Media Report," retrieved from *https://www.nielsen.com/content/dam/corporate/us/en/reports-downloads/2017-reports/2016-nielsen-social-media-report.pdf* on 5 July 2018.

[2] Andreoni, J. and B.D. Bernheim (2009) "Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects," *Econometrica*, 77 (5), pp. 1607 - 1636.

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

[4] Bagwell, L. and B.D. Bernheim (1986) "Veblen Effects in a Theory of Conspicuous Consumption," *The American Economic Review*, 86 (3), pp. 349 - 373.

[5] Bergstom, C.; S. Számado; and M. Lachmann (2002) "Separating Equilibria in Continuous Signalling Games," *Philosophical Transactions: Biological Sciences*, 357, pp. 1595-1606.

[23] Interpreting Lemma 1 in this way requires identifying λ with the N as we do in Section 4.2.

[6] B.D. Bernheim (1994) "A Theory of Conformity," *The Journal of Political Economy*, 102 (4), pp. 841-877.

[7] Bernheim, B.D. and A. Bodoh-Creed (2017) "A Theory of Decisive Leadership," *mimeo.*

[8] Bernheim, B.D. and S. Severinov (2016) "Bequests as Signals: An Explanation for the Equal Division Puzzle," *Journal of Political Economy*, 111 (4), pp. 733-764.

[9] Besley, T. (1995) "Nonmarket Institutions for Credit and Risk Sharing in Low-Income Countries," *Journal of Economic Perspectives*, 9 (3), pp. 115-127.

[10] boyd, d. (2004) "Friendster and Publicly Articulated Social Networks," *Conference on Human Factors and Computing Systems*, Vienna: ACM.

[11] Buffardi, L. and W.K. Campbell (2008) "Narcissism and Social Networking Web Sites," *Personality and Social Psychology Bulletin*, 34 (10), pp. 1303-1314.

[12] Chi, I-K. and D. Kreps (1987) "Signaling Games and Stable Equilibria," *The Quarterly Journal of Economics*, 102 (2, pp. 179-222.

[13] Corneo, G. and O. Jeanne (1997) "Conspicuous Consumption, Snobbism, and Conformism," *Journal of Public Economics*, 66 (1), pp. 55 - 71.

[14] van Dijck, Jose (2013) "You have one identity: performing the self on Facebook and LinkedIn," *Media, Culture & Society*, 35 (2), pp. 199-215.

[15] Donath, J. (2002) "Identity and deception in the virtual community." *Communities in Cyberspace*. Eds. P. Kollock and M. Smith. London: Taylor & Francis Group, pp. 27-58.

[16] Donath, J. (2008) "Signals in Social Supernets," *Journal of Computer-Mediated Communication*, 13, pp. 231-351.

[17] Donath, J. and d boyd (2004) "Public displays of connection," *BT Technology Journal*, 22 (4), pp. 71-82.

[18] Drouin, M.; K. O'Connor; G. Schmidt; and D. Miller (2015) "Facebook fired: Legal perspectives and young adults opinions on the use of social media in hiring and firing decisions," *Computers in Human Behavior*, 46, pp. 123-128.

[19] Engers, M. (1987) "Signalling with Many Signals," *Econometrica*, 55 (3), pp. 663 - 674.

[20] Eventbrite (2017) "Millenials: Fueling the Experience Economy," downloaded on 16 August 2018 from *https://www.eventbrite.com/blog/academy/millennials-fueling-experience-economy/*.

[21] Fafchamps, M. and S. Lund (2003) "Risk-sharing networks in rural Philippines," *Journal of Development Economics*, 71, pp. 261-287.

[22] Frankel, A. and N. Kartik (2017) "Muddled Information," *mimeo*.

[23] Galloway, M.; J. Conner; and D. Pope (2013) "Nonacademic Effects of Homework in Privileged, High-Performing High Schools," *The Journal of Experimental Education*, 81 (4), pp. 490 - 510.

[24] Garg, R. and R. Telang (2017) "To Be or Not to Be Linked: Online Social Networks and Job Search by Unemployed Workforce," *Management Science*, forthcoming.

[25] Gee, L.; J. Jones; and M. Burke (2017) "Social Networks and Labor Markets: How Strong Ties Relate to Job Finding on Facebook's Social Network," *Journal of Labor Economics*, 35 (2), pp. 485-518

[26] Getty, T. (1998) "Handicap signalling: when fecundity and viability do not add up," *Animal Behavior*, 56, pp. 127-130.

[27] Goffman, E. (1959) *The presentation of self in everyday life*. Garden City, NY: Doubleday Anchor.

[28] Grafen, A. (1990) "Biological Signals as Handicaps," *Journal of Theoretical Biology*, 144, pp. 517-546.

[29] Granovetter, M. (1973) "The Strength of Weak Ties," *American Journal of Sociology*, 78 (6) , pp. 1360-1380.

[30] Harvard Graduate School of Education, "Turning the Tide: Inspiring Concern for Others and the Common Good through College Admissions." Retrieved from *http://mcc.gse.harvard.edu/collegeadmissions* on 27 February 2016.

[31] Hoppe, H.; B. Moldovanu; and A. Sela (2009) "The Theory of Assortative Matching Based on Costly Signals," *The Review of Economic Studies*, 76 (1), pp. 253-281.

[32] Hurd, P. (1995) "Communication in Discrete Action-Response Games," *Journal of Theoretical Biology*, 174, pp. 217-222.

[33] Ireland, N. (1994) "On limiting the market for status signals," *Journal of Public Economics*, 53, pp. 91-110.

[34] Jobvite (2013), *2013 Social Recruiting Survey*, downloaded from https://www.jobvite.com/jobvite-news-and-reports/2013jobvitesocialrecruitingsurvey/.

[35] Kluemper, D. and P. Rosen (2009) "Future employment selection methods: evaluating social networking web sites," *Journal of Managerial Psychology*, 24 (6), pp. 567-580.

[36] Lachmann, M.; S. Számado; and C. Bergstrom (2001) "Cost and conflict in animal signals and human language," *Proceedings of the National Academy of Sciences of the United States of America*, 98, pp. 13189-13194.

[37] Luthar, S. and B. Becker (2002) "Privileged but Pressured? A Study of Affluent Youth," *Child Development*, 73 (5), pp. 1593 - 1610.

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

[39] Mailath, G. and E. von Thadden (2013) "Incentive compatibility and differentiability: New results and classic applications," *Journal of Economic Theory*, 148, pp. 1841-1861.

[40] Maynard, S. (1982) *Evolution and the Theory of Games*. Cambridge: Cambridge University Press.

[41] Milgrom, P. and J. Roberts (1986) "Price and Advertising Signals of Product Quality," *Journal of Political Economy,* 94 (4), pp. 796 - 821.

[42] Miller, M. and K. Rock (1985) "Dividend Policy Under Asymmetric Information," *Journal of Finance*, 40 (4), pp. 1031- 1051.

[43] Montgomery, J., (1991) "Social Networks and Labor-Market Outcomes: Toward an Economic Analysis" *The American Economic Review*, 81 (5), pp. 1408-1418.

[44] Neilsen (2016), *Neilsen 2016 Social Media Report*, downloaded from http://www.nielsen.com/us/en/insights/reports/2017/2016-nielsen-social-media-report.html.

[45] Smith, A. and M. Anderson (2018) "Social Media Use in 2018," (Pew Research Center, Washington, D.C.).

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

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

[48] Spence, M. (1973) "Time and Communication in Economic and Social Interaction," *The Quarterly Journal of Economics*, 87 (4), pp. 651-660.

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

[50] Számado, S. (1999) "The Validity of the Handicap Principle in Discrete Action-Response Games," *Journal of Theoretical Biology*, 198, pp. 593-602.

[51] Számado, S. (2011) "The cost of honesty and the fallacy of the handicap principle," *Animal Behavior*, 81, pp. 3-10.

[52] Thoumrungroje, A. (2014) "The Influence of Social Media Intensity and EWOM on Conspicuous Conumption," *Procedia - Social and Behavioral Sciences*, 148, pp. 7-15.

[53] Wernerfelt, B. (1988) "Umbrella Branding as a Signal of New Product Quality: An Example of Signaling by Posting a Bond," *The RAND Journal of Economics*, 19 (3), pp. 458 - 466.

[54] Wilson, R. (1985) "Multi-dimensional Signaling," *Economics Letters*, 19, pp. 17 - 21.

[55] Zahavi, A. (1979) "Mate Selection—A Selection for a Handicap," *Journal of Theoretical Biology*, 53, pp. 205-214.

A PROOF APPENDIX

Lemma 1. *Let Assumptions 1, 2 and 5 hold. Then* $\|a(t;\lambda_i) - a_{BP}(t)\| = O\left(\frac{1}{\sqrt{\lambda_i}}\right)$ *as* $\lambda_i \to \infty$.

Proof. The bound on $a(t;\lambda_i) - a_{BP}(t)$ will be derived from the following inequality, which must hold in equilibrium:

$$\lambda_i \left[\pi(a_{BP}(t), t) - \pi(a(t;\lambda_i), t)\right] \leq \overline{B} - \underline{B} \tag{5}$$

We first prove our result for the case where $\pi_{aa}(a_{BP}(t), t) < 0$, and then consider what occurs when $\pi_{aa}(a_{BP}(t), t) = 0$.

The uniqueness of the bliss point for type t (Assumption 1) implies that $\pi(a_{BP}(t), t) - \pi(a(t;\lambda_i), t) > 0$, and the continuity of $\pi(\circ, t)$ implies $a(t;\lambda_i) - a_{BP}(t) \to 0$ as $\lambda \to \infty$. The Taylor expansion of $\pi(a, t)$ around $(a_{BP}(t), t)$ is:

$$\pi(a_{BP}(t), t) - \pi(a(t;\lambda_i), t) = \frac{-1}{2}\pi_{aa}(a_{BP}(t), t)\left(a(t;\lambda_i) - a_{BP}(t)\right)^2 - \frac{\pi_{aaa}(\xi)}{6}\left(a(t;\lambda_i) - a_{BP}(t)\right)^3 \tag{6}$$

where $\xi \in [a_{BP}(t), a(t;\lambda_i)]$. Suppose $\pi_{aaa}(\xi)\left(a(t;\lambda_i) - a_{BP}(t)\right) < 0$. Then

$$\pi(a_{BP}(t), t) - \pi(a(t;\lambda_i), t) \geq \frac{-1}{2}\pi_{aa}(a_{BP}(t), t)\left(a(t;\lambda_i) - a_{BP}(t)\right)^2$$

Suppose $\pi_{aaa}(\xi)\left(a(t;\lambda_i) - a_{BP}(t)\right) > 0$. Then since $\|\pi_{aaa}(a, t)\| \leq C$ (Assumption 2) and $a(t;\lambda_i) - a_{BP}(t) \to 0$ as $\lambda_i \to \infty$, we can choose λ_i^* such that for all $\lambda_i > \lambda_i^*$

$$\left\|\frac{\pi_{aaa}(\xi)}{6}\left(a(t;\lambda_i) - a_{BP}(t)\right)\right\| \leq \frac{-1}{4}\pi_{aa}(a_{BP}(t), t)$$

which means we have

$$\pi(a_{BP}(t), t) - \pi(a(t;\lambda_i), t) \geq \frac{-1}{4}\pi_{aa}(a_{BP}(t), t)\left(a(t;\lambda_i) - a_{BP}(t)\right)^2 \tag{7}$$

In either case, using equation 5 we can write

$$\left(a(t;\lambda_i) - a_{BP}(t)\right)^2\left(\frac{-1}{4}\pi_{aa}(a_{BP}(t), t)\right) \leq \frac{\overline{B} - \underline{B}}{\lambda}$$

which in turn yields

$$\|a(t;\lambda_i) - a_{BP}(t)\| \leq \frac{1}{\sqrt{\lambda}}\sqrt{\frac{-4(\overline{B}-\underline{B})}{\pi_{aa}(a_{BP}(t),t)}} \tag{8}$$

In the case where $\pi_{aa}(a_{BP}(t),t) = 0$, the first higher-order partial derivative that is nonzero must be of an even-numbered order.[24] If the order of the first nonzero derivative with respect to a is k, we can use Assumption 2, Part (2) to make an argument analogous to that provided above to show convergence at the rate

$$\|a(t;\lambda_i) - a_{BP}(t)\| = O\left(\frac{1}{\lambda_i^{k/2}}\right)$$

□

Lemma 2. *Let Assumptions 1, 3, and 4 hold. Then in any fully separating equilibrium we have for almost all t that $a_{SEP}(t;\lambda) \geq a_{BP}(t)$ for λ sufficiently large.*

Proof. First note that the claim holds for $\underline{t}$ since $a_{SEP}(\underline{t}) = a_{BP}(t)$ in any fully separating equilibrium. Now consider an arbitrary $t > \underline{t}$ and suppose our claim fails to hold. Then for any $\lambda^* > 0$ there exists $\lambda > \lambda^*$ such that $a_{SEP}(t) < a_{BP}(t)$. Equation 8 implies that for such a choice of λ sufficiently large that $a_{SEP}(\underline{t}) \leq a_{SEP}(t) < a_{BP}(t)$. But from the continuity of $a_{BP}(t)$ (Assumption 1), there exists $t' \in [\underline{t}, t)$ such that $a_{BP}(t') = a_{SEP}(t)$. Since the equilibrium is fully separating, it must be that $a_{SEP}(t') \neq a_{BP}(t')$ or types t and t' would pool. Since $B(t,\widehat{t})$ is increasing in $\widehat{t}$ (Assumption 4), t' can profitably deviate to $a_{BP}(t')$ and have the receiver infer her type to be $\widehat{t} = t > t'$. This contradiction proves our claim. □

Theorem 1. *Let Assumptions 1 - 5 hold and assume the receiver's beliefs satisfy the dominance refinement. Then $\lambda\left[\pi(a_{BP}(t),t) - \pi(a_{SEP}(t),t)\right] = O\left(\frac{1}{\sqrt{\lambda}}\right)$ for any $\alpha > 0$.*

[24] To see this more formally, if $\pi_{aa}(a_{BP}(t),t) = 0$, then Equation 6 has the form

$$\pi(a_{BP}(t),t) - \pi(a(t;\lambda_i),t) = \frac{-1}{6}\pi_{aaa}(a_{BP}(t),t)\,(a_{BP}(t) - a(t;\lambda_i))^3 - \frac{\pi_{aaaa}(\xi)}{24}\,(a_{BP}(t) - a(t;\lambda_i))^4$$

The fourth order term is negligible relative to the third order term for λ sufficiently large (i.e., $a_{BP}(t)$ and $a(t;\lambda_i)$ sufficiently close), so if $\pi_{aaa}(a_{BP}(t),t) \neq 0$ utility would be increased by either a slight increase or decrease in a from $a_{BP}(t)$. But then $a_{BP}(t)$ cannot be optimal, and from this contradiction we conclude that $\pi_{aa}(a_{BP}(t),t) = 0$ entails $\pi_{aaa}(a_{BP}(t),t) = 0$. We conclude that if $\pi_{aa}(a_{BP}(t),t) = 0$, then the first higher-order partial derivative that is nonzero must be of an even-numbered order.

Proof. For the duration of the proof we assume that λ is sufficiently large that $a_{SEP}(t) \geq a_{BP}(t)$ as per Lemma 2. The goal of this proof is to tighten the bound provided by Lemma 1. To that end, suppose agent t deviates from $a_{SEP}(t)$ to $a_{BP}(t)$. The proof of Lemma 1 showed that even if a deviation from her bliss point could shift $B(t, \delta(a))$ from $\underline{B}$ to $\overline{B}$, the largest she would be willing to deviate is of $O(\lambda^{-0.5})$. If there exists t' such that $a_{SEP}(t') = a_{BP}(t)$, then the receiver infers (incorrectly) that the sender is of type t'. Using Equation 8 we can write:

$$a_{BP}(t) = a_{SEP}(t') \leq a_{BP}(t') + \frac{1}{\sqrt{\lambda}}\sqrt{\frac{-4\left(\overline{B} - \underline{B}\right)}{\pi_{aa}(a_{BP}(t'), t')}}$$

From Assumption 3 we have

$$\begin{aligned} \beta(t - t') &\leq a_{BP}(t) - a_{BP}(t') \leq \frac{1}{\sqrt{\lambda}}\sqrt{\frac{-4\left(\overline{B} - \underline{B}\right)}{\pi_{aa}(a_{BP}(t'), t')}} \\ t' &\geq t - \frac{1}{\beta\sqrt{\lambda}}\sqrt{\frac{-4\left(\overline{B} - \underline{B}\right)}{\pi_{aa}(a_{BP}(t'), t')}} \end{aligned} \tag{9}$$

If there is no t' such that $a_{SEP}(t') = a_{BP}(t)$, then Lemma 1 combined with the dominance refinement implies that the receiver must believe that the sender has some type t' that satisfies $\|a_{BP}(t') - a_{BP}(t)\| = O(\lambda^{-0.5})$. Using Equation 8 we can write this formally as:

$$a_{BP}(t) \leq a_{BP}(t') + \frac{1}{\sqrt{\lambda}}\sqrt{\frac{-4\left(\overline{B} - \underline{B}\right)}{\pi_{aa}(a_{BP}(t'), t')}}$$

As before, Assumption 3 implies

$$t' \geq t - \frac{1}{\beta\sqrt{\lambda}}\sqrt{\frac{-4\left(\overline{B} - \underline{B}\right)}{\pi_{aa}(a_{BP}(t'), t')}} \tag{10}$$

We now use Equation 9 or 10 as appropriate to bound the effect on the signaling incentive $B(t, \widehat{t})$ more tightly. The core idea is that the the cost of signaling can be no larger than the benefit received by having the receiver infer that the sender has type t

instead of type t'. Using Assumption 4 we have:

$$\begin{aligned}\pi(a_{BP}(t),t) - \pi(a_{SEP}(t),t) &\leq \frac{1}{\lambda}\left[B(t,t) - B\left(t, t - \frac{1}{\beta\sqrt{\lambda}}\sqrt{\frac{-4\left(\overline{B} - \underline{B}\right)}{\pi_{aa}(a_{BP}(t'),t')}}\right)\right] \\ &\leq \frac{\gamma}{\lambda^{1.5}}\left[\frac{1}{\beta}\sqrt{\frac{-4\left(\overline{B} - \underline{B}\right)}{\pi_{aa}(a_{BP}(t'),t')}}\right]\end{aligned}$$

This then yields:

$$\begin{aligned}\lambda\left[\pi(a_{BP}(t),t) - \pi(a_{SEP}(t),t)\right] &\leq \frac{\gamma}{\sqrt{\lambda}}\left[\frac{1}{\beta}\sqrt{\frac{-4\left(\overline{B} - \underline{B}\right)}{\pi_{aa}(a_{BP}(t'),t')}}\right] \qquad (11) \\ &= O\left(\frac{1}{\sqrt{\lambda}}\right) \qquad (12)\end{aligned}$$

□

Theorem 2. *Let Assumptions 1 - 5 hold and suppose that $a_{BP}(t)$ and $a_{SEP}(t)$ are continuous for all λ. Then $\lambda\left[\pi(a_{BP}(t),t) - \pi(a_{SEP}(t),t)\right] = O\left(\frac{1}{\lambda^{1-\alpha}}\right)$ for any $\alpha > 0$.*

Proof. The goal of this proof is also to tighten the bound provided by Lemma 1. The key difference with the proof of Theorem 1 is that due to the continuity of $a_{BP}(t)$ and $a_{SEP}(t)$, if an agent deviates from $a_{SEP}(t)$ to $a_{BP}(t)$, then there exists a t' such that the receiver believes that sender has type $t' < t$ — namely t' such that $a_{SEP}(t') = a_{BP}(t)$. This allows us to avoid the issue of off-path beliefs.

To that end, suppose agent t deviates from $a_{SEP}(t)$ to $a_{BP}(t)$. The type t' such that $a_{SEP}(t') = a_{BP}(t)$ defines the inference made by the receiver following the deviation by type t. Equation 8 yields:

$$a_{SEP}(t') = a_{BP}(t) \leq a_{BP}(t') + \frac{1}{\sqrt{\lambda}}\sqrt{\frac{-4\left(\overline{B} - \underline{B}\right)}{\pi_{aa}(a_{BP}(t'),t')}}$$

From Assumption 3 we have

$$\begin{aligned} \beta(t-t') &\leq a_{BP}(t) - a_{BP}(t') \leq \frac{1}{\sqrt{\lambda}}\sqrt{\frac{-4\left(\overline{B}-\underline{B}\right)}{\pi_{aa}(a_{BP}(t),t)}} \\ t' &\geq t - \frac{1}{\beta\sqrt{\lambda}}\sqrt{\frac{-4\left(\overline{B}-\underline{B}\right)}{\pi_{aa}(a_{BP}(t'),t')}} \end{aligned}$$

This allows us to bound the effect on the signaling incentive $B(t,\widehat{t})$ more tightly. Using Assumption 4 we have:

$$\begin{aligned} \pi(a_{BP}(t),t) - \pi(a_{SEP}(t),t) &\leq \frac{1}{\lambda}\left[B(t,t) - B\left(t, t - \frac{1}{\beta\sqrt{\lambda}}\sqrt{\frac{-4\left(\overline{B}-\underline{B}\right)}{\pi_{aa}(a_{BP}(t'),t')}}\right)\right] \\ &\leq \frac{\gamma}{\lambda^{1.5}}\left[\frac{1}{\beta}\sqrt{\frac{-4\left(\overline{B}-\underline{B}\right)}{\pi_{aa}(a_{BP}(t'),t')}}\right] \end{aligned}$$

If Assumption 2, Part (1) holds, we can then write

$$\frac{-1}{4}\pi_{aa}(a_{BP}(t),t)\left[a_{SEP}(t) - a_{BP}(t)\right]^2 \leq \pi(a_{BP}(t),t) - \pi(a_{SEP}(t),t) \tag{13}$$

$$\leq \frac{1}{\lambda^{1.5}}\frac{\gamma}{\beta}\sqrt{\frac{-4\left(\overline{B}-\underline{B}\right)}{\pi_{aa}(a_{BP}(t'),t')}} \tag{14}$$

where the first inequality can be derived using an argument essentially identical to that used to derive Equation 7 from the proof of Lemma 1. Simplifying we have:

$$a_{SEP}(t) - a_{BP}(t) \leq \frac{1}{\lambda^{3/4}}\sqrt{\frac{\gamma}{\beta}}\left(\frac{-4}{\pi_{aa}(a_{BP}(t'),t')}\right)^{3/4}\left(\overline{B}-\underline{B}\right)^{1/4} \tag{15}$$

If Assumption 2, Part (2) applies, we can make an analogous argument that yields an even tighter bound on $a_{SEP}(t) - a_{BP}(t)$ as in the proof for Lemma 1.

Iterating this process K times yields

$$a_{SEP}(t) - a_{BP}(t) \leq \frac{C_K}{\lambda^{1-0.5^K}}$$

When we use this in our Taylor expansion we get

$$\begin{aligned}\pi(a_{BP}(t),t)-\pi(a_{SEP}(t),t) &= (a_{SEP}(t)-a_{BP}(t))^2\left(\frac{-1}{2}\pi_{aa}(a_{BP}(t),t)-\right.\\ &\qquad\left.\frac{\pi_{aaa}(\xi)}{6}\left(a_{SEP}(t)-a_{BP}(t)\right)\right)\\ &= \frac{C_K^2}{\lambda^{2-0.5^{K-1}}}\left(\frac{-1}{2}\pi_{aa}(a_{BP}(t),t)-\frac{\pi_{aaa}(\xi)}{6}\left(\frac{C_K}{\lambda^{1-0.5^K}}\right)\right)\end{aligned}$$

Using the negligibility of the third order terms, we find

$$\begin{aligned}\lambda\left[\pi(a_{BP}(t),t)-\pi(a_{SEP}(t),t)\right] &\le \frac{C_K^2}{\lambda^{1-0.5^{K-1}}}\left(\frac{-1}{2}\pi_{aa}(a_{BP}(t),t)-\frac{\pi_{aaa}(\xi)}{6}\left(\frac{C_K}{\lambda^{1-0.5^K}}\right)\right)\\ &= O\left(\frac{1}{\lambda^{1-0.5^{K-1}}}\right)\end{aligned}$$

as desired. □

Theorem 3. *Assume that $\pi(a,t)$ is supermodular in (a,t) and for all t and $a > a_{BP}(t)$ we have $\pi_a(a,t) \le 0$ and $\frac{\pi_{at}(a,t)}{\pi_a(a,t)}$ is weakly increasing in a. For any fixed N, $\mathbf{a}_{SEP}^N$ maximizes the payoff of each type of sender relative to any other separating equilibrium.*

Proof. Suppose we have a separating equilibrium with action functions $\mathbf{a}(t) = (a_1(t), ..., a_N(t))$. Defining

$$\Gamma(\mathbf{a},\mathbf{t}) \equiv \sum_{i=1}^{N} \pi\left(a_i, t\right) \tag{16}$$

we can write the first-order condition for type t's optimal choice as:

$$\left.\frac{\partial B(t,\widehat{t})}{\partial \widehat{t}}\right|_{\widehat{t}=t} + \sum_{i=1}^{N} \frac{\partial \pi\left(a_i,t\right)}{\partial a_i}\frac{da_i(t)}{dt} = 0 \tag{17}$$

We are interested in determining type t's total payoff in equilibrium. If we let $V\left(t,\widehat{t}\right)$ denote the payoff of a type t sender having chosen the action of type $\widehat{t}$, we have by definition:

$$V\left(t,\widehat{t}\right) = B\left(t,\widehat{t}\right) + \Gamma(\mathbf{a}(\widehat{t}),t) \tag{18}$$

and the Envelope Theorem yields:

$$\frac{dV\left(t,t\right)}{dt}=\frac{dB\left(t,t\right)}{dt}+\frac{\partial\Gamma(\mathbf{a}(t),t)}{\partial t}$$

Notice that only the final term depends on the particular separating equilibrium. Let $\mathbf{a}^0$ denote the symmetric separating equilibrium with payoffs V^0, and $\mathbf{a}^A$ denote an asymmetric separating equilibrium with payoffs V^A. To demonstrate that payoffs in the symmetric separating equilibrium are strictly higher than in the asymmetric separating equilibrium, we will establish the following Property (capitalized for clarity of subsequent references): if it were the case for some t that either (i) $V^0(t,t) = V^A(t,t)$ and $\mathbf{a}^0(t) \neq \mathbf{a}^A(t)$, or (ii) $V^0(t,t) < V^A(t,t)$, then we would have $\frac{dV^0(t,t)}{dt} > \frac{dV^A(t,t)}{dt}$.[25]

To understand why this Property delivers the desired conclusion, note that $V^A(t',t') - V^0(t',t')$ would shrink as t' rises over $[\underline{t},t]$ if the property holds. But then we would have a violation of the boundary condition $V^0(\underline{t},\underline{t}) = V^A(\underline{t},\underline{t}) = B(\underline{t},\underline{t}) - \Gamma(\boldsymbol{a}_{BP}(\underline{t}),\underline{t})$ where $\boldsymbol{a}_{BP}\left(t\right) = (a_{BP}(t),...,a_{BP}(t)) \in \mathbb{R}^N$. In light of Equation 18, we can rewrite the Property as follows: if it were the case for some t that either (i)′ $\Gamma(\mathbf{a}^0(t),t) = \Gamma(\mathbf{a}^A(t),t)$ and $\mathbf{a}^0(t) \neq \mathbf{a}^A(t)$, or (ii)′ $\Gamma(\mathbf{a}^0(t),t) > \Gamma(\mathbf{a}^A(t),t)$, then we would have $\frac{\partial\Gamma(\mathbf{a}^0(t),t)}{\partial t} > \frac{\partial\Gamma(\mathbf{a}^A(t),t)}{\partial t}$.

We now establish the Property. Supposing condition (i)′ were satisfied for some $t > \underline{t}$, we would begin by defining:[26]

$$\overline{a}_m = \begin{cases} a_m^A(t) \text{ if } a_m^A(t) \geq a_{BP}(t) \\ a \geq a_{BP}(t) \text{ s.t. } \pi\left(a,t\right) = \pi(a_m^A(t),t) \text{ otherwise} \end{cases}$$

where this can be done in an abritrary order over the dimensions of $\boldsymbol{a}^A$. Let $Q \equiv \left\{m \mid a_m^A(t) < a_{BP}(t)\right\}$. Then from supermodularity we have:

$$\frac{\partial\Gamma(\overline{\boldsymbol{a}},t)}{\partial t} - \frac{\partial\Gamma(\boldsymbol{a}^A(t),t)}{\partial t} = \sum_{m\in Q}\pi_t(\overline{a}_m,t) - \pi_t(a_m^A,t) \geq 0$$

with strict inequality if Q is non-empty.

If $\boldsymbol{a}^0(t) = \overline{\boldsymbol{a}}$, we are done. If not, then since $\Gamma(\boldsymbol{a}^A(t),t) = \Gamma(\overline{\boldsymbol{a}},t)$ by construction, there must exist i and j such that $\overline{a}_i > a^0(t) > \overline{a}_j$. Define the function $\widetilde{\boldsymbol{a}}(a_i)$ as follows:

[25] Suppose our claim is true. Then if either condition (i) or (ii) holds for t, then condition (ii) must hold for all $t' \in (\underline{t},t)$.

[26] This step sets computes a cost-equivalent signal to $\boldsymbol{a}^A$ that has the intuitive property that $a_m^A \geq a_{BP}(t)$.

$\widetilde{a}_i(a_i) = a_i$, $\widetilde{a}_k(a_i) = \overline{a}_k$ for $k \neq i, j$, and $\Gamma(\widetilde{a}(a_i), t) = \Gamma(\overline{\boldsymbol{a}}, t)$. In other words, $\widetilde{a}_j(a_i)$ indicates how a_j must vary in response to changes in a_i to keep the value of Γ constant at its equilibrium value. Implicit differentiation reveals that:

$$\left.\frac{d\widetilde{a}_j}{da_i}\right|_{a_i=\overline{a}_i} = -\frac{\pi_a\left(\overline{a}_i, t\right)}{\pi_a\left(\widetilde{a}_j(\overline{a}_i), t\right)} < 0$$

Plainly, there exists a unique value $a_i^e > a_{BP}(t)$ such that $\widetilde{a}_j(a_i^e) = a_i^e$. For $a_i \in [a_i^e, \overline{a}_i(t))$ we have:

$$\begin{aligned}
\frac{d}{da_i}\left(\frac{\partial\Gamma(\widetilde{\mathbf{a}}(a_i), t)}{\partial t}\right) &= \frac{d}{da_i}\left(\sum_{i=1}^{N}\pi_t\left(\widetilde{a}_i(a_i), t\right)\right)_{a_i=\overline{a}_i} \\
&= \pi_{at}\left(a_i, t\right) + \pi_{at}\left(\widetilde{a}_j(a_i), t\right)\left.\frac{d\widetilde{a}_j}{da_i}\right|_{a_i=a_i} \\
&= \pi_{at}\left(a_i, t\right) - \pi_{at}\left(\widetilde{a}_j(a_i), t\right)\frac{\pi_a\left(a_i, t\right)}{\pi_a\left(\widetilde{a}_j(a_i), t\right)} < 0
\end{aligned}$$

where we have used the fact that since $\overline{a}_i \geq a_i \geq a_i^e \geq \widetilde{a}_j(a_i) > a_{BP}(t)$ (which implies $\pi_a\left(\overline{a}_i, t\right) < 0$) and our assumption that for $a > a_{BP}(t)$ we have $\pi_a(a, t) \leq 0$ and:

$$\frac{\pi_{at}\left(a_i, t\right)}{\pi_a\left(a_i, t\right)} > \frac{\pi_{at}\left(\widetilde{a}_j(\overline{a}_i), t\right)}{\pi_a\left(\widetilde{a}_j(\overline{a}_i), t\right)}$$

It follows that $\frac{\partial\Gamma(\widetilde{a}(a_i^e),t)}{\partial t} > \frac{\partial\Gamma(\overline{a},t)}{\partial t}$ since $\overline{a}_i > a^0(t)$ is being reduced in this equalization step. Through repeated application of this equalization argument, we conclude that $\frac{\partial\Gamma(a^0(c),t)}{\partial t} > \frac{\partial\Gamma(\overline{a},t)}{\partial t} \geq \frac{\partial\Gamma(a^A(t),t)}{\partial t}$, as desired.

Next, supposing condition (ii)$'$ were satisfied for some $t > \underline{t}$, we would begin by defining a' s.t. $a_1' = a_2' = ... = a_N' > a_{BP}(t)$ and $\Gamma(a', t) = \Gamma(a^A(t), t)$. By the same argument as for condition (i)$'$, we infer $\frac{\partial\Gamma(a',t)}{\partial t} \geq \frac{\partial\Gamma(a^A(t),t)}{\partial t}$.[27] Because $\Gamma(a^0(t), t) > \Gamma(a^A(t), t) = \Gamma(a', t)$ by assumption, we have $a_m^0(t) > a_m'$. From our assumption of supermodularity we conclude:

$$\frac{\partial\Gamma(a^0(c), t)}{\partial t} - \frac{\partial\Gamma(a', t)}{\partial t} = \sum_{m=1}^{N}\pi_t(a_m^0, t) - \pi_t(a_m', t) \geq 0$$

It follows that $\frac{\partial\Gamma(a^0(c),t)}{\partial t} > \frac{\partial\Gamma(a^A(c),t)}{\partial t}$, as desired.

[27] The inequality is weak because we include the possibility that $a' = a^A(t)$.

Having established that the Property holds, the Proposition follows for the reasons given above. □

Theorem 4. *Let Assumptions 4, 5, 7, and 8 hold. In addition, assume* $a_{SEP}(t) \geq a_{BP}(t)$ *and that the receiver's beliefs satisfy the dominance refinement. Then* $\boldsymbol{\pi}^N\left(\boldsymbol{a}_{BP}^N(t), t\right) - \boldsymbol{\pi}^N\left(\boldsymbol{a}_{SEP}^N(t), t\right) = O\left(\frac{1}{\sqrt{-\phi_N}}\right)$.

Proof. For the duration of the proof let $S_N(t) = \sum_{i,j=1}^N \frac{\partial^2 \boldsymbol{\pi}^N(\boldsymbol{a}_{BP}^N(t),t)}{\partial a_i \partial a_j} (< 0)$. We begin by imposing a bound on the size of the possible cost of signaling. Consider the Taylor expansion of $\pi^N(\boldsymbol{a}_{BP}^N(t), t) - \pi^N(a_{SEP}^N(t), t)$

$$\frac{-1}{2}\sum_{i,j=1}^{N} \frac{\partial^2 \boldsymbol{\pi}^N(\boldsymbol{a}_{BP}^N(t), t)}{\partial a_i \partial a_j}(a_{SEP,i}^N(t) - a_{BP,i}^N(t))(a_{SEP,j}^N(t) - a_{BP,j}^N(t)) + \tag{19}$$
$$\frac{-1}{6}\sum_{i,j,k=1}^{N} \frac{\partial^3 \boldsymbol{\pi}^N(\boldsymbol{\xi}, t)}{\partial \boldsymbol{a}_i \partial \boldsymbol{a}_j \partial \boldsymbol{a}_k}(a_{SEP,i}^N(t) - a_{BP,i}^N(t))((a_{SEP,j}^N(t) - a_{BP,j}^N(t))(a_{SEP,k}^N(t) - a_{BP,k}^N(t))$$

where $\boldsymbol{\xi}$ lies on the line connecting $\boldsymbol{a}_{BP}^N(t)$ and $\boldsymbol{a}_{SEP}^N(t)$.

Consider a deviation by type t from $a_{BP}(t)$ to $a_{BP}(t) + \delta \mathbf{1}^N$ where $\delta \geq 0$. Using algebra similar to that employed in the proof of Lemma 1, it is straightforward to show using Equation 19 in conjunction with Assumption 8 that

$$\boldsymbol{\pi}^N\left(\boldsymbol{a}_{BP}^N(t), t\right) - \boldsymbol{\pi}^N\left(\boldsymbol{a}_{BP}^N(t) + \delta \mathbf{1}^N, t\right) \leq \overline{B} - \underline{B} \tag{20}$$

implies

$$\delta \leq \sqrt{\frac{-4(\overline{B} - \underline{B})}{S_N(t)}} = \overline{\delta} \tag{21}$$

Suppose that an agent of type t deviates from $\boldsymbol{a}_{SEP}^N(t)$ to $\boldsymbol{a}_{BP}^N(t)$ and there exists a type t' such that $\boldsymbol{a}_{BP}^N(t) = \boldsymbol{a}_{SEP}^N(t')$. The receiver infers (incorrectly) that the sender is of type t'. Using Equation 21 we can write:

$$\boldsymbol{a}_{BP}^N(t) = \boldsymbol{a}_{SEP}^N(t') \leq \boldsymbol{a}_{BP}^N(t') + \sqrt{\frac{-4(\overline{B} - \underline{B})}{S_N(t')}}\mathbf{1}^N$$

From Assumption 7 we have

$$\begin{aligned} \beta(t-t')\mathbf{1}^N &\leq \boldsymbol{a}_{BP}^N(t)-\boldsymbol{a}_{BP}^N(t') \leq \sqrt{\frac{-4(\overline{B}-\underline{B})}{S_N(t')}}\mathbf{1}^N \\ t' &\geq t-\frac{1}{\beta}\sqrt{\frac{-4(\overline{B}-\underline{B})}{S_N(t')}} \end{aligned} \tag{22}$$

If there is no t' such that $\boldsymbol{a}_{BP}^N(t) = \boldsymbol{a}_{SEP}^N(t')$, then Equation 21 combined with the dominance refinement implies that the receiver must believe that the sender has some type t' that satisfies:

$$\boldsymbol{a}_{BP}^N(t) \leq \boldsymbol{a}_{BP}^N(t') + \sqrt{\frac{-4(\overline{B}-\underline{B})}{S_N(t')}}\mathbf{1}^N$$

As before, Assumption 7 implies

$$t' \geq t-\frac{1}{\beta}\sqrt{\frac{-4(\overline{B}-\underline{B})}{S_N(t')}} \tag{23}$$

We now use Equation 22 or 23 as appropriate to bound the effect on the signaling incentive $B(t,\widehat{t})$ more tightly. The core idea is that the the cost of signaling can be no larger than the benefit received by having the receiver infer that the sender has type t instead of type t'. Using Assumption 4 we have:

$$\begin{aligned} \boldsymbol{\pi}^N\left(\boldsymbol{a}_{BP}^N(t),t\right)-\boldsymbol{\pi}^N\left(\boldsymbol{a}_{SEP}^N(t),t\right) &\leq B(t,t)-B\left(t,t-\frac{1}{\beta}\sqrt{\frac{-4(\overline{B}-\underline{B})}{S_N(t')}}\right) \\ &\leq \frac{\gamma}{\beta}\sqrt{\frac{-4(\overline{B}-\underline{B})}{S_N(t')}} \\ &= O\left(\frac{1}{\sqrt{-\phi_N}}\right) \end{aligned}$$

□

Theorem 5. *Consider the additive model, and assume $\pi(a,t)$ satisfies assumptions Assumptions 1 - 3 and 9 hold and the receiver's beliefs satisfy the dominance refinement.*

Then

$$\sum_{i=1}^{N} \left[\boldsymbol{\pi} \left(a_{BP}(t), t \right) - \boldsymbol{\pi} \left(a_{SEP,i}^{N}(t), t \right) \right] = O\left(\sqrt{\frac{\gamma(F)^3}{N}} \right).$$

Proof. Consider equations 11 and 12 in the proof of Theorem 1. The claim follows once we replace γ with $\gamma(F)$ and note that $\overline{B} - \underline{B} \leq \gamma(F)\bar{t}$ by Assumption 9. □

Theorem 6. *Let Assumptions 1 - 3 and 9 hold and suppose that $a_{BP}(t)$ and $a_{SEP}(t)$ are continuous for all N. Then*

$$N\left[\pi(a_{BP}(t), t) - \pi(a_{SEP}(t), t)\right] = O\left(\left[\frac{\gamma(F)^2}{N} \right]^{1-\alpha} \right)$$

for any $\alpha > 0$.

Proof. Consider equation 15. Replacing γ with $\gamma(F)$ and $\overline{B} - \underline{B}$ with $\gamma(F)\bar{t}$ as per Assumption 9, we find an equation proportional to $(\gamma(F)/\lambda)^{0.75}$. A third iteration of the argument yields an equation with a similar form to equation 15 that is proportional to $(\gamma(F)/\lambda)^{0.875}$, while a fourth iteration yields an equation proportional to $(\gamma(F)/\lambda)^{0.9375}$. Combining this with the Taylor approximation of $\pi(a, t)$ in equation 13 yields our result. □

B Online Appendix: Additional Applications

B.1 Political Decisiveness

Bernheim and Bodoh-Creed [7] presents a model of politician decisiveness. We provide a reduced form version of the model here for conciseness. The type of politician, denoted t, reflects the politician's innate aversion to delay. This can be interpreted as either the politician's perception of the opportunity costs of delaying a policy decision or the strength of a politician's policy preferences (i.e., weak policy preferences imply a high t). As intuition would suggest, politicians with a high aversion to delay prefer to make policy decisions more quickly. The politician's ideal policy is $p^* = \theta + x_p \in \mathbb{R}$, where θ represents a policy preference that is common across the population and x_p is a policy preference idiosyncratic to the politician. Both θ and x_p are uncertain. A politician in office gets to tune the policy based on what she learns during the period of delay about both her idiosyncratic preferences and the common component of preferences.

The agency conflict stems from two features. First, the politician delays her decision to learn about both the common and idiosyncratic components of her preferences. While voters would like the politician to learn about the common component, the voters do not gain any benefit from what the politician learns about her idiosyncratic preferences. The net effect is that the politician delays longer than voters would prefer. Second, the politician tunes her policy choice to her own idiosyncratic preferences and not those of the voters, so the politician's idiosyncratic preferences act as a source of risk for voters. Because of the agency conflict, politicians have an incentive to signal their aversion to delay.

If only one decision can be observed, the signaling incentives can cause the politician to make the decision more quickly than either the politician or the voters would prefer. However,the rise of cable news and online media has increased the volume of political coverage, which has in turn increased the opportunities for politicians to make their electorate aware of their actions. In addition, as the current U.S. president has demonstrated, OSNs such as Twitter provide ample opportunities to advertise purported accomplishments. Pew Research reports that 93% of American report getting news online, and users often finds news on social networks such as Facebook or Twitter.[28] The online webpages for the top 50 newspapers reported 11.7 million unique viewers in 2016, and natively online news orga-

[28] Downloaded on 3 August 2018 from http://assets.pewresearch.org/wp-content/uploads/sites/13/2018/07/11183646/State-of-the-News-Media˙2017-Archive.pdf

nizations were able to attract 22.8 million unique users per news organization per month.[29] And none of these statistics count individuals who are made aware of headlines through social networks. Our analysis suggests that the increasing attention paid to politicians has reduced the costs they must incur to signal their decisiveness (or many other traits they might wish to signal).[30]

In period 1, the voters observe a politician make a decision in lower office. In particular, the voters observe how quickly the politician makes the decision, and will later use this as a signal of the politician's decisiveness. In period 2, the politician competes against a randomly drawn opponent for higher office. Politicians are motivated to signal t because voters prefer more decisive politicians that are inclined to make decisions more quickly. In other words, a candidate of type t that signals her type is $\widehat{t}$ will win the election if voters believe her opponent has a lower value of t (i.e., a lower aversion to delay). In period 3, the winning politician holds office and reaps the rewards from her decisiveness. The losing politician gets the same utility as the voters.

Let a denote the amount of time taken to make a decision in period 1. Note that in this model, the politician wants to signal a high aversion to delay by choosing a *lower* (i.e., quicker) action. This is, of course, a minor reconfiguration of the model, but the definition of many of the assumptions must be slightly adjusted to account for the fact that $a_{SEP}(t) < a_{BP}(t)$. The direct benefit of action a is $\frac{-1}{a+\phi}$, which captures the fact that delaying the decision (increasing a) reduces the expected difference between the policy chosen and the ideal policy. The direct cost of delay is at. Therefore we have

$$\pi(a,t) = \frac{-1}{a+\phi} - at$$

The first order condition yields $a_{BP}(t) = \sqrt{\frac{1}{a}} - \phi$. Therefore, the model satisfies Assumptions 1-3, and $\frac{\pi_{at}(a,t)}{\pi_a(a,t)}$ is weakly decreasing in a (which is the desired condition since $a_{SEP}(t) < a_{BP}(t)$).

To simplify our exposition, we let the payoff of winning the election given a type t be denoted $B_{win}(t)$. Given that a politician wins, his payoff does not depend on who his

[29] Natively digital means the primary outlet is online. Websites that received at least 10 million unique visitors per month were analyzed.

[30] On the other hand, the opening of these alternate channels for news reporting has been at the expense of local television and local and national print journalism, however. This means that our argument applies more readily to national level politicians, and, in fact, signaling costs may get higher as local news outlets close.

opponent was in period 2, so $B_{win}(t)$ depends only on the winner's type. The payoff in expectation for a politician of type t that is believed by voters to have type $\widehat{t}$ is $B_{loss}(\widehat{t}, t)$. $B_{loss}(\widehat{t}, t)$ depends on $\widehat{t}$ to account for the expected type of the winner, and it depends on t to account for the fact that the amount of time the winner takes to make decisions has different welfare effects on different types of losing politicians. If the distribution of t amongst the politicians is $F(t)$, then we can write

$$B(t, \widehat{t}) = F(\widehat{t})B_{win}(\widehat{t}) + (1 - F(\widehat{t}))B_{loss}(t, \widehat{t})$$

where $F(\widehat{t})$ is the probability that sender wins the election in a separating equilibrium (i.e., that the opponent's type is less than $\widehat{t}$). We assume that $B(t, \widehat{t})$ satisfies Assumptions 4 and 5, but one can prove this from the microfoundations provided in Bernheim and Bodoh-Creed [7].

In reality, politicians make many decisions while in office, and we classify a political institution as more or less transparent according to whether voters can observe a greater or lesser fraction of these decisions. Greater transparency leads politicians to "spread" their signals across many decisions. Formally, suppose the voters observe N of these choices. If we assume that $\pi(a, t)$ is the same for each decision, Theorem 3 implies that the symmetric equilibrium where $a_{SEP}(t)$ is the choice for each of the N observed decisions is welfare optimal for the politicians. Bodoh-Creed and Bernheim [7] show that that $a_{BP}(t)$ and $a_{SEP}(t)$ are continuous for all N. Corollary 1 then implies

Corollary 5. $N\left[\pi(a_{BP}(t), t) - \pi(a_{SEP}(t), t)\right] = O\left(\frac{1}{N^{1-\alpha}}\right)$ *for any* $\alpha > 0$

This corollary shows that as transparency increases (N grows), the welfare losees from signaling fade. This implies that increasing transparency is good for politicians. However, the case is ambiguous for voters, as discussed at length in Bodoh-Creed and Bernheim [7]. Too little transparency causes politicians to behave hastily and make decision more quickly than voters would prefer. However, too much transparency reduces the signaling distortion, which allows politicians to signal their type while acting less decisively than voters would prefer.

B.2 Job Market Signaling

We now consider how our results play out in a model of job market signaling. The model of Spence [49] assumed that students vary in terms of an underlying attribute t that is the private information of the students. The students signal this attribute through the choice of years of eduction, a, which we interpret as a measure of human capital. The individual's productivity for an employer is $S(a, t)$, so the choice of a is productive. In a competitive separating equilibrium $a_{SEP}(t)$, the wages are $S(a_{SEP}(\widehat{t}), \widehat{t})$, and we assume that students obtain no intrinsic benefit from human capital except as it influences their wages. We also assume that $S(a, \widehat{t}) = B(\widehat{t}) + P(a)$, and accruing human capital level a has a cost equal to $C(a, t)$.[31]

Now we encode this application into our model. The benefit of signaling is captured by $B(\widehat{t})$. The direct utility from choosing a given t is $\pi(a, t) = P(a) - C(a, t)$. The bliss point is

$$a_{BP}(t) = \underset{a}{arg\,max}\ P(a) - C(a, t)$$

We assume that $a_{BP}(t)$ is unique, $a_{BP}(\underline{t}) > 0$, and $C(a, t)$ is strictly supermodular in $(-a, t)$ so that $a_{BP}(t)$ is strictly increasing. To insure that the productivity effect of effort is nontrivial, we also assume there exists $\beta > 0$ such that for $t > t'$ we have $a_{BP}(t) - a_{BP}(t') \geq \beta(t - t')$.

Letting $\widehat{t}(a)$ denote the inferences of the firms after observing a, utility maximization in a fully separating equilibrium requires the following first-order condition be satisfied

$$B_{\widehat{t}} \frac{d\widehat{t}(a)}{da} + P_a(a) - C_a(a, t) = 0$$

where the first term captures the marginal signaling incentive, the second term reflects the marginal productivity of effort, and the third time is the marginal cost of effort. In a complete information model, the first order condition that defines the agents' bliss points is

$$S_a(a_{BP}(t), t) + \pi_a(a_{BP}(t), t) = 0$$

Since the signaling incentive provides an additional incentive to accrue more capital relative to the complete-information model, the students accrue too much human capital relative to the first best, even though the human capital is productive.

The model laid out here assumes that students can only differentiate themselves to

[31] Assuming separability of $(a, \widehat{t})$ is done so that we can encode this application into our additive model. If $(a, \widehat{t})$ are not separable, then we need to let $B(\boldsymbol{a}^N, \widehat{t}) = S(\boldsymbol{a}^N, \widehat{t})$ in our multi-action setting, impose conditions along the lines of Footnote 9, and use Theorem 4 on general aggregator functions. The conspicuous consumption example in Section 5.2 has a structure of this form.

employers through a single action, which may initially appear to be a useful and innocuous simplification. However, our results reveal that this simplifying assumption actually has substantive economic implications. A more realistic model would allow the students to distinguish themselves along many different dimensions of human capital accumulation. For example, subject-specific human capital can be measured with grades. Leadership and communication skill is reflected in leadership positions in student organizations. The ability to work in teams can be assessed via internships or references from previous employers.

Assume there are N human capital metrics and, for simplicity, assume that the direct utility and cost functions are the same across all of these metrics. Leting a_i^N be the i^{th} component of $\boldsymbol{a}^N$, the sender's utility is

$$B(\widehat{t}(\boldsymbol{a}^N)) + \sum_{i=1}^{N} \left[P(a_i^N) - C(a_i^N, t)\right]$$

where for simplicity we have assumed the receiver's beliefs, $\widehat{t}(\boldsymbol{a}^N)$, are degenerate. We assume that

$$\frac{\pi_{at}(a,t)}{\pi_a(a,t)} = \frac{-C_{at}(a,t)}{P_a(a,t) - C_a(a,t)}$$

is weakly increasing in a, and from Theorem 3 we conclude that an equilibrium where the same action is taken across all N metrics is welfare optimal. Letting $a_{SEP}(t)$ be the symmetric action taken across the N metrics, t Corollary 1 implies the following corollary that the total cost of signaling vanishes as the number of human capital metrics increases.

Corollary 6. *Suppose that $a_{BP}(t)$ and $a_{SEP}(t)$ are continuous for all N. Then for any $\alpha > 0$ we have $N\left[\pi(a_{BP}(t), t) - \pi(a_{SEP}(t), t)\right] = O\left(\frac{1}{N^{1-\alpha}}\right)$*

C Online Appendix: Nonadditive Aggregators

Consider the following utility function:

$$U_N(\boldsymbol{a}^N, t) = B\left(t, \widehat{t}\left(\boldsymbol{a}^N\right)\right) + \boldsymbol{\pi}_{Agg}\left(a_1, .., a_N, t\right)$$

As an example, suppose the cost function has the form:

$$\boldsymbol{\pi}_{Agg}\left(a_1, .., a_N, t\right) = \sqrt{\sum_{m=1}^{N} \pi(a_m, t)}$$

where $\sqrt{\pi(a,t)}$ satisfies Assumptions 2 and 3.[32] Then for a separating equilibrium with symmetric actions we can write:

$$U_N(\boldsymbol{a}^N, t) = B\left(t, \widehat{t}\left(\boldsymbol{a}^N\right)\right) + \sqrt{N}\sqrt{\pi(a_{SEP}(t), t)}$$

Theorem 2 implies $U_N(a_{BP}(t), t; \lambda) - U_N(a_{SEP}(t; \lambda), t; \lambda) = O\left(\frac{1}{N^{0.5-\alpha}}\right)$ for any $\alpha > 0$.

More generally, suppose one can write the equilibrium utility for a symmetric separating equilibrium, $\boldsymbol{a}_{SEP}(t) = (a_{SEP}(t), a_{SEP}(t), ..., a_{SEP}(t))$, as:

$$U_N(\boldsymbol{a}_{SEP}(t), t) = B\left(t, t\right) + \lambda(N) g\left(a_{SEP}(t), t\right)$$

Assume the functions g and λ satisfy the following two assumptions:

Assumption 10. $g_{aa}(a,t) < 0$ *and there exists* $C < \infty$ *such that* $g_{aaa} \leq C$*.*

Assumption 11. $\lambda(N) \to \infty$ *as* $N \to \infty$*.*

Assumption 10 is analogous to Assumption 2 and allows us to focus our analysis on the second order expansion of the function $g(a,t)$. Assumption 11 requires that $\lambda(N)$, which is analogous to the decision weight of Section 4, diverges to infinity as $N \to \infty$. The additively separable model of Subsection 4.2 and the example leading this subsection satisfy these requirements.

Under these assumptions, an argument analogous to the proof of Theorem 2 shows that

[32] Theorem 3 applies in any situation where $\boldsymbol{\pi}^N$ is a function of $\sum_{m=1}^{N} \pi(a_m, t)$, which would justify our focus on symmetric separating equilibria.

aggregate signaling costs vanish as N grows.[33] For completeness, we state this result as the following corollary, where $a_{BP}(t) = \underset{a}{arg\,max}\ g\,(a, t)$.[34]

Corollary 7. *Let Assumptions 1, 4 - 3, 10, and 11 hold and assume that* $a_{SEP}(t)$ *and* $a_{BP}(t)$ *are continuous. Then*

$$\lambda(N)\,[g(a_{BP}(t), t) - g(a_{SEP}(t), t)] = O\left(\frac{1}{\lambda(N)^{1-\alpha}}\right)$$

for any $\alpha > 0$.

[33]To see this point, if one replaces λ with $\lambda(N)$ throughout the proof of Theorem 2, the same technical argument applies.

[34]Theorem 1 admits an analogous corollary, which we omit for the sake of brevity.