Online Appendix to Mechanism Design with News Utility

Jetlir Duraj

This version: September 14, 2019

The online appendix is organized as follows. In section 1 we show how one can use Euler-Lagrange equations to find a solution of the optimal screening mechanism in timeline C whenever a weak condition which implies no ironing is satisfied.

In section 2 we show that the classical revelation principle holds true in our multi-agent model.

In section 3 we prove that for timeline B of the multi-agent model of the main paper optimal mechanisms are all-pay and also show that this result as well as virtually all our results from the main paper remain intact when the outside option of the agents is non-degenerate in the money dimension.

Finally, section 4 registers several additional applications of the multi-agent model of the main paper: we prove that the classical impossibility results for ex-post efficiency in bilateral trade and public good provision remain intact in our setting for timeline A; we also characterize optimal auctions for asymmetric bidders in the case of timeline A as well as for symmetric bidders in timeline B.

1 Characterization of the optimal screening mechanism for timeline C under a no ironing condition.

In this section we require some additional regularity assumptions on $v$ as well as an additional joint regularity requirement on the problem which avoids ironing considerations.

AA1: Additional Regularity of $v$. Besides the classical conditions from the Set Up subsection of the screening model (Section 2) in the main body of the paper assume that $v$ satisfies the following.

- $v'(x) \geq \frac{K}{v^{p-1}(x)}$ for some $p \geq 2$
- $\frac{(v-1)'}{(v-1)}$ and $\frac{1}{(v-1)}$ are uniformly Lipschitz continuous in every interval $I$ of the type $I = (a, +\infty)$, $a > 0$.

The conditions are weak: the classical conditions from the main body of the paper as well as AA1 are all satisfied for e.g. $v(x) = x^l$ as long as $l \in [\frac{1}{2}, 1)$.

*duraj@g.harvard.edu*
Define for use in the following the function \( f : [1, \hat{\lambda}] \times \mathbb{R}_+ \to \mathbb{R} \) given by
\[
f(\lambda, y) = \frac{g'(\lambda)}{g(\lambda)} \left[ \frac{m}{M} + 1 - \lambda - c(v^{-1})'(y) \right] - 2
\]

Similar to the proof of Proposition 4 in the appendix of the main paper we assume here first that the threshold \( \hat{\lambda} \) is given.

We are looking for a solution of the following Calculus of Variations Problem with both qualitative and also inequality constraints.

\[
\begin{align*}
\max_{s \to u(s)} & \quad \int_1^{\lambda} \left\{ \Gamma^C(\lambda) \frac{u'(\lambda)}{M} + u(\lambda) - cv^{-1}(u'(\lambda)) \right\} dG(\lambda) \\
\text{s.t.} & \quad (1) \ u(s) \text{ is continuous, non-decreasing, concave} \\
& \quad (2) \ [\Gamma^B(\lambda) - \Gamma^C(\lambda)] \frac{u'(\lambda)}{M} - u(\lambda) \geq 0, \ \text{a.e.} \ \lambda \in [1, \hat{\lambda}].
\end{align*}
\]

If we have a condition AA2 which gives that any solution of (1) without the constraint (1) also satisfies it automatically, we have a solution for (1). Note that the existence of a solution for the maximization of \( J \) over \( C(\hat{\lambda}) \) \(^1\) implies that we can use Euler-Lagrange methods (necessary conditions for an optimum) in the relaxed problem where \((u(1), u'(1))\) are fixed. The solutions of these Euler-Lagrange equations are then parametrized by \((u(1), u'(1))\) and we can then solve for the optimal \((u(1), u'(1))\) which still satisfies constraint (2). If then (1) is automatically satisfied for every \( u \) optimizing \( J \) with fixed \((u(1), u'(1))\) we have thus characterized a solution of the original problem (1) which we can then translate to an optimal mechanism for the timeline \( C \) in the same way as in the proof of Proposition 4, part 2.

We now calculate the Euler-Lagrange equation for the maximization in (1) ignoring the constraints, except for fixing the pair \((u(1), u'(1))\) \(\in \mathbb{R} \times \mathbb{R}_+\). For this, we first extend \( v^{-1} \) to a strictly convex, twice continuously differentiable function \( w : \mathbb{R} \to \mathbb{R}_+ \), with the property \( w(x) = v^{-1}(-x) \) for negative \( x \). It is easy to see that our assumptions on \( v \) from the main paper imply that \( w \) is differentiable with \( w'(0) = 0 \). For ease of notation we don’t distinguish between \( v^{-1} \) and \( w \) in the following. The integrand has the form
\[
L(\lambda, y, y') = g(\lambda) \left( \Gamma^C(\lambda) \frac{y'}{M} + y - c(v^{-1})(y') \right).
\]
Note that it is concave in \( y' \) and that it holds for the partial derivatives
\[
L_y = g(\lambda), \quad L_{y'} = g(\lambda) \left( \frac{\Gamma^C(\lambda)}{M} - c(v^{-1})'(y') \right).
\]

The second variation at a stationary point \( y(\lambda) \) is given by the integral of
\[
S(\lambda, y(\lambda), y'(\lambda)) := (y')^2(\lambda) L_{y'y'}(\lambda, y(\lambda), y'(\lambda)) + y^2(\lambda) L_{yy}(\lambda, y(\lambda), y'(\lambda)) \\
+ 2y(\lambda)y'(\lambda) L_{y'}(\lambda, y(\lambda), y'(\lambda)) = -c \cdot g(\lambda) \cdot (v^{-1})''(y'(\lambda)) \cdot (y')^2(\lambda), \quad \lambda \in [1, \hat{\lambda}].
\]
\(^1\)Recall from the main paper that this is the set of appropriately bounded Sobolev-p functions which satisfies (1)-(2) a.e. in (1).
Here $y$ is a Sobolev ‘test’ function, i.e. it satisfies $y(1) = y(\hat{\lambda}) = 0$. The boundary conditions on the test function are needed because in this step we are considering fixed $u(1), u'(1)$. The Poincaré inequanty holds for such test functions (see section 4.4.1 in [Dacorogna '08]). Upon integrating $S$ w.r.t. $\lambda$ the fact that the second variation is negative and definition of the Sobolev-p-norm implies immediately that every stationary point is a local maximum. This is because the second variation of the problem satisfies the ellipticity condition usually used as a sufficient second order condition in the calculus of variations (see pg. 100 of [Gelfand, Fomin '00]).

Recall that the functional being maximized has the form

$$ J(u) = \int_{1}^{\hat{\lambda}} L(\lambda, u(\lambda), u'(\lambda))d\lambda. $$

Every stationary point $u$ for fixed $(u(1), u'(1))$ will thus satisfy the Euler-Lagrange equation and be automatically a maximum for the sub-problem with fixed pair $(u(1), u'(1))$. The Euler-Lagrange equation is

$$ L_y(\lambda, u(\lambda), u'(\lambda)) - \frac{d}{d\lambda} L_{y'}(\lambda, u(\lambda), u'(\lambda)) = 0, \quad \lambda \in [1, \hat{\lambda}]. $$

Plugging in and calculating we get the following second-order ODE

$$ u''(\lambda) = f(\lambda, u'(\lambda)) \frac{1}{c(v^{-1})''(u'(\lambda))}, \quad \lambda \in [1, \hat{\lambda}]. $$

This ordinary differential equation (ODE) has always a locally unique solution, as it is easily checked using the classical existence and uniqueness result for ODEs.²

From the Proposition 4 in the main body of the paper it follows that if the solution $u$ of the Euler-Lagrange equations satisfies all of the constraint (1) automatically for every pair $(u(1), u'(1)) \in \mathbb{R} \times \mathbb{R}_+$ the optimal solution characterized in the proof of Part 1 of Proposition 4 in the main paper is a solution of the Euler-Lagrange equations with appropriate pair $(u(1), u'(1))$.

We require the following condition which implies that no ironing of $u$ is needed.

**AA2 - Joint Regularity.** It holds $\frac{g'(\lambda)}{g(\lambda)} \left[ \frac{m}{M} + 1 - \lambda - c(v^{-1})'(y) \right] \leq 2$ for all $\lambda, y \in [1, \hat{\lambda}] \times \mathbb{R}_+$.

This is a relatively weak condition: it is always satisfied when $G$ is a uniform distribution or whenever $\frac{m}{M}$ is low and $g'(\lambda) > 0$ everywhere or whenever $\frac{m}{M}$ is high and $g'(\lambda) < 0$ everywhere.

AA2 and the previous assumptions imply automatically that every stationary point is concave. Thus any solution of the Euler-Lagrange equations satisfies (1). This implies that our procedure delivers a full solution to (1) by following the simple recipe whenever AA1-AA2 are satisfied:

The Euler Lagrange equation delivers $u$ up to boundary conditions. The set of feasible boundary conditions $(u(1), u'(1))$ is then given by plugging in requirement (2) of (1). Given optimal $u$, one can always back out $f$ and $q$ as in the proof of Proposition 4, part 1 in the appendix of the main paper and thus get a solution of the original problem for

---

²This is an implication of the second regularity requirement in AA1.
the fixed \( \hat{\lambda} \). We know from discussion in the proof of the above mentioned Proposition from the main paper that w.l.o.g. the set of feasible \((u(1), u'(1))\) can be taken to be compact. Once we plug the candidate \(u\) together with any feasible \((u(1), u'(1))\) into the objective function of (1), it turns into a continuous function in the pair \((u(1), u'(1))\) whose maximum existence is ensured by compactness. Finding the optimal threshold \( \hat{\lambda} \) proceeds then as in the proof of Proposition 4, part 1 in the paper.

**Example 1.** Just as in Example 2 we take \( F = \mathcal{U}[0, 1] \), resulting in \( m = \frac{1}{2} \) and \( M = \frac{1}{6} \). Moreover, we take again \( G = \mathcal{U}[1, 2] \) as well as \( v(q) = \sqrt{q} \). We use the recipe above. Note that AA1 and AA2 are trivially satisfied here.

*Step 1.*

Using the assumption \( v(q) = \sqrt{q} \) yields for every stationary point of the Lagrange equation

\[
u''(\lambda) = -\frac{1}{c}, \quad \lambda \in [1, 2].\]

This corresponds to a solution

\[
u(\lambda) = u(1) + u'(1)(\lambda - 1) - \frac{1}{c} \left( \frac{(\lambda - 1)^2}{2} \right), \quad \lambda \in [1, 2]. \tag{2}\]

*Step 2.* We find all functions of the form (2) which lie in the set of feasible \( u \)-s defined in (1).

Plugging (2) into (1) and doing routine calculations we arrive at the following expression for the revenue

\[
u'(1)(\hat{\lambda}^2 + 2\hat{\lambda} - 1) - cu'(1)(\hat{\lambda} - 1) + u(1)(\hat{\lambda} - 1) - \frac{1}{c} \left( \frac{1}{6} \hat{\lambda}^3 + \hat{\lambda}^2 - \frac{5}{2} \hat{\lambda} + \frac{4}{3} \right) \tag{3}\]

We note here for future use that \( \text{rest}(\hat{\lambda}) = \frac{1}{6} \hat{\lambda}^3 + \hat{\lambda}^2 - \frac{5}{2} \hat{\lambda} + \frac{4}{3} \) is nonnegative on \([1, 2] \) and has a unique minimum of zero at \( \hat{\lambda} = 1 \). It follows in particular that optimal profit is always non-negative.

Plugging (2) in the constraint (2) of the program (1) results in the restriction

\[
\frac{(\lambda - 1)(\lambda - 3)}{\lambda + 1} \left( u'(1) - \frac{1}{c} (\lambda - 1) \right) \geq u(1) + u'(1)(\lambda - 1) - \frac{1}{c} \left( \frac{(\lambda - 1)^2}{2} \right), \quad \lambda \in [1, \hat{\lambda}]. \tag{4}\]

Thus, the problem is reduced to maximizing (3) w.r.t. (4) and the constraint \( u'(1) \geq 0 \). The constraint (4) can be written as

\[
u(1) \leq u'(1) \frac{4(1 - \lambda)}{\lambda + 1} - \frac{(\lambda - 1)^2(\lambda - 7)}{2c(\lambda + 1)} = \left( \frac{(7 - \lambda)(\lambda - 1)}{2c} - 8u'(1) \right) \frac{\lambda - 1}{\lambda + 1}, \quad \lambda \in [1, \hat{\lambda}]. \tag{5}\]

Define the continuous function \( h(c, u'(1), \hat{\lambda}) = \min_{\lambda \in [1, \hat{\lambda}]} \left( \frac{(7 - \lambda)(\lambda - 1)}{2c} - 8u'(1) \right) \frac{\lambda - 1}{\lambda + 1} \).

The overall maximization problem for fixed threshold \( \hat{\lambda} \) is then

4
\[ R^C(\hat{\lambda}) = \max_{u(1), u'(1)} \] 
\[
\begin{align*}
&u'(1)(\hat{\lambda}^2 + 2\hat{\lambda} - 1) - cu'(1)(\hat{\lambda} - 1) + u(1)(\hat{\lambda} - 1) - \frac{1}{c} \left( \frac{1}{6} \hat{\lambda}^3 + \hat{\lambda}^2 - \frac{5}{2} \hat{\lambda} + \frac{4}{3} \right) \\
\text{s.t.} & (1) \quad u'(1) \geq 0 \\
& (2) \quad u(1) \leq h(c, u'(1), \hat{\lambda}).
\end{align*}
\]

The classical maximum theorem gives that \( R^C \) is continuous in \( \hat{\lambda} \) and one can solve for the optimal threshold to get the complete characterization of the optimal mechanism.

\section{On the revelation principle for the case of uncertainty about intrinsic types}

The classical revelation principle states that in certain Bayesian mechanism design situations it is sufficient for the analysis of implementability of a social choice function to focus on a particular class of mechanisms called direct mechanisms.\(^3\) In a direct mechanism the agents declare their types (private information) and the designer uses the reported type profile to choose a consumption allocation and transfers.

The following formal definitions recall the set up usually used to prove the revelation principle, with the (innocuous) twist of adding the timeline \( T \) as a component of the mechanism.

\textbf{Definition 1.} 1) A mapping \( F : \Theta \rightarrow A \times \mathbb{R}^N \) is called a social choice function (SCF).

2) Given strategy spaces \( S_i \) for each agent \( i \), a game form \( g \) is a mapping

\[
g : S_1 \times \ldots \times S_N \rightarrow A \times \mathbb{R}^N, \quad s \mapsto (q^g(s), t^0_g(s), \ldots, t^N_g(s)).
\]

3) A mechanism is a triple \( M = (S_1 \times \ldots \times S_N, g, T) \) consisting of a product of strategy spaces \( S_1 \times \ldots \times S_N \), a game form \( g \) and a timeline \( T \in \{A, B, C, D\} \).

In a game form \( g \) an interim strategy of agent \( i \) is a mapping \( \sigma_i : \Theta_i \rightarrow S_i \). We denote the set of such strategies by \( \Sigma_i \).

Recall the definitions from the main text of expected valuation \( V_i \), expected transfer \( T_i \) and the related \( T_i^+ \) as well as the gain loss utility terms in the good and money dimensions, \( \Gamma_i \) and \( \omega_i \). Those were mappings defined on the type spaces \( \Theta = \Theta_1 \times \ldots \Theta_N \). For a given game form \( g \) and parallel to all these auxiliary terms from the main body of the paper define the terms \( V_i^g, T_i^g, T_i^{g+} \) as well as \( \Gamma_i^g \) and \( \omega_i^g \) as mappings from the set of all strategies \( \Sigma = \Sigma_1 \times \ldots \Sigma_N \) to the space of functions \( \mathbb{R}^{\Theta_i} \). The value of these mappings for a profile of strategies \( \sigma = (\sigma_1, \ldots, \sigma_N) \in \Sigma \) are given by replacing one-for-one in the formulas of the main text \( q(\theta) \) and \( t_i(\theta) \) by respectively \( q^g(\sigma(\theta)) \) and \( t^g_i(\sigma(\theta)) \).

Given strategies \( \sigma_{-i} \) of the other agents, the prior on the type space \( \Theta \) and using the functions \( q^g(\sigma(\theta)) \) and \( t^g_i(\sigma(\theta)) \), \( j = 1, \ldots, n \) each choice of strategy \( \sigma_i \) for agent \( i \) induces for each type \( \theta_i \) a distribution over consumption and transfers. This is given by taking the image of \( (q^g(\sigma(\theta)), t^g_i(\sigma(\theta))) \) under the posterior over \( \Theta_{-i} \) given her type \( \theta_i \).

\(^3\)See for example Chapter 6 of [Myerson '97] or Chapter 7 in [Fudenberg, Tirole '91].
Given a timeline $T$, a decision utility $\sigma_i \mapsto W^T_{t_i}(\sigma_i, \sigma_{-i})$ is defined analogously to the play decision utilities $W_i$ in the main text by making the following replacements pointwise for each fixed type profile $\theta \in \Theta$: $q(\theta)$ by $q^g(\sigma(\theta))$ and $t_i(\theta)$ by $t^g_i(\sigma(\theta))$.

It follows that a mechanism induces a Bayesian game $((1,\ldots,N), \Theta, \times_{i=1}^N, (\Sigma = \Sigma_i)_{i=1}^N)$ among the agents.

**Definition 2.** We say the mechanism $\mathcal{M}$ implements the SCF $\mathcal{F}$ in Bayesian Nash equilibrium, if the induced Bayesian game has a Bayesian Nash equilibrium $\sigma^* \in \Sigma$ such that for all $\theta \in \Theta$ it holds $g(\sigma^*(\theta)) = \mathcal{F}(\theta)$.

We finally introduce direct mechanisms and the related incentive compatibility definition. As now timelines are important these are ‘indexed’ by the variable $T$.

**Definition 3.** 1) A mechanism for timeline $T$ is called a direct mechanism if it is of the form

$$\mathcal{M} = (\Theta_1 \times \ldots \Theta_N, g, T).$$

2) A direct mechanism $\mathcal{M} = (\Theta_1 \times \ldots \Theta_N, g, T)$ is called truthful if the profile of identity functions $id_i : \Theta_i \to \Theta_i$ given by $id_i(\theta_i) = \theta_i, i = 1,\ldots,N$ is a Bayesian Nash equilibrium of the game induced by the mechanism.

Given a general mechanism $\mathcal{M} = (\Theta_1 \times \ldots \Theta_N, g, T)$ implementing a SCF $\mathcal{F}$ in timeline $T$, there exists a BNE $\sigma^*$ of the induced game such that $g(\sigma^*(\theta)) = \mathcal{F}(\theta)$. Define the corresponding direct mechanism $\mathcal{M}^d$ by

$$\mathcal{M}^d = (\Theta_1 \times \ldots \Theta_N, g^d, T)$$

with

$$g^d(\theta) = g(\sigma^*(\theta)).$$

It is easy to see, with the same argument as in the original classical proof, that $\mathcal{M}^d$ is truthful. With this the Revelation Principle follows.

**Proposition 1** (Revelation Principle for mechanisms with news utility). A mechanism $\mathcal{M}$ implements a SCF $\mathcal{F}$ if and only if there exists a truthful direct mechanism $\mathcal{M}^d$ implementing the same SCF.

We note that for a fixed SCF, implementability (through direct or indirect mechanisms) may be possible in different timelines. It may happen that a SCF is implementable in all timelines as well as not in all of them. Proposition 8 in the main paper is an example of this.

This version of the Revelation Principle may not give strict implementation in the sense that there may be other equilibria of the associated direct mechanism $\mathcal{M}^d$ which don’t correspond to the Bayesian Nash equilibria of a mechanism $\mathcal{M}$. We avoid the question of strict implementation in this setting and tacitly make the assumption that whenever a direct mechanism $\mathcal{M}^d$ has multiple equilibria besides the truthful one, the agents coordinate on the truthful one.

Just as in the classical setting, the revelation principle allows us to focus the analysis on direct mechanisms only. We use this extensively in the main body of the paper as well as here in the online appendix.
3 All-pay optimality in timeline B

Theorem 1 (All-Pay Structure for B). Assume that the agents are loss averse in the money dimension. For a fixed incentive compatible and individually rational mechanism $M = (q, t_1, \ldots, t_N)$ in timeline B there is another incentive compatible and individually rational mechanism with the same allocation rule and degenerate transfers $M' = (q, t'_1, \ldots, t'_N)$ which yields weakly higher expected revenue than $M$. The revenue under $M'$ is strictly higher whenever the transfers in $M$ are non-degenerate.

In particular, revenue maximizing mechanisms in any of the timelines B, C and D feature degenerate transfers.

Proof of Theorem 1. Timeline B: We write $O_i(\theta_i)$ for an agent’s utility from an outside option to avoid distinguishing in the following between the cases $v_i(\emptyset) \in \{\inf v_i, \sup v_i\}$. Obviously $O_i$ is non-negative and non-decreasing.

First, one sees from the Mirrlees representation, that in an incentive compatible mechanism the individual rationality requirement is equivalent to

$$V_i(\theta_i) \geq \max_{\Theta_i}\{O_i(\theta_i) - \int_{\theta_i}^{\Theta_i} W_i(s) ds\}$$

for all $\theta_i$. (6) will bind in any revenue optimal mechanism. This implies that individual rationality only determines $V_i(\theta_i)$ which in turn determines only $\Upsilon_i(\theta_i)$.

Recall the defining relation of the perceived payments

$$\Upsilon^B_i(\theta_i) = (1 + \eta^m_i) T_i(\theta_i) + \Lambda_i^m(T_i^+(\theta_i) + \omega_i(\theta_i))$$

This relation shows that, for fixed $V_i$ and $\Upsilon_i$, the latter determined up to constant by the optimal incentive compatible allocation rule, setting $\omega_i \equiv 0$ by choosing degenerate payments always leaves room to increase $T_i(\theta_i)$ and $T_i^+(\theta_i)$ further by adding to them a non-negative constant. In other words, for fixed $\Upsilon_i$ the $T_i$-s found in the proof of Proposition 6 of the main body of the paper are the largest possible.

While in classical optimal mechanism design models with quasilinear utilities, transfer schedules can always be rearranged to be degenerate, Theorem 1 says, that this is strictly optimal from a revenue perspective in the presence of news utility and loss aversion for timelines B. A similar optimality result for all-pay mechanisms has been noted in the classical literature for classical agents whose utility is separable in the consumption and money dimension and who are risk averse in the money dimension (see for example [Maskin, Riley ’84]). Theorem 1 shows that there could be other reasons for the optimality of all-pay mechanisms: the agents may experience news utility and the implementation of the mechanism may feature delays between the participation decision and the realization of the outcome of the mechanism.

4 Additional Results for multiple agents and uncertainty about intrinsic types

4.1 Public Goods

There is a classical impossibility result for the classical model of public good provision (see [Güth, Hellwig ’86] or [Börgers ’15]). It shows that ex-post efficiency can never be
implemented in a model of public good provision with quasilinear utilities and independent types. It is natural to expect the same result to hold with timeline A due to the isomorphism between this timeline and the classical quasilinear model. The following Proposition confirms this intuition. We focus for simplicity on symmetric agents.

**Proposition 2.** Assume that $(1 + \eta^0)\bar{\theta} < c(n) < (1 + \eta^0)\bar{\theta}$ and timeline A. Then every incentive compatible and IR mechanism which implements ex-post efficiency exhibits negative budget balance.

**Proof.** From the Mirrlees representation in Proposition 6 of the paper it is easy to see, that the socially optimal mechanism has degenerate payments and sets $V_i(\bar{\theta}) = 0$. Moreover, due to the calculation

$$\Upsilon_i(\theta) = W_i(\theta)\theta - \int_{\theta}^{\theta} W_i(s)ds = \int_{\theta}^{\theta} (W_i(\theta) - W_i(s)) ds + W_i(\theta)\theta \geq 0,$$

where

$$W_i(\theta) = (1 + \eta^0)Q_i(\theta) - \Lambda^0Q_i(\theta)(1 - Q_i(\theta)),$$

it follows

$$T_i(\theta) = \frac{\Upsilon_i(\theta)}{1 + \lambda^m\eta^m}.$$

We get for the ex-ante Budget surplus of the mechanism BS

$$BS = \sum_{i=1}^{N} \int_{0}^{\theta} T_i(s)dF(s) - Nc(N)(1 - F^*N(nc'(N)))$$

$$= \frac{N}{1 + \lambda^m\eta^m} \int_{0}^{\theta} W(s)\gamma(s)dF(s) - Nc(N)(1 - F^*N(nc'(N))).$$

Here $\gamma(s) = s - \frac{F(s)}{f(s)}$ is the virtual valuation. Using the convolution property $\int_{0}^{\theta} F^{*(N-1)}(Nc'(N) - s)dF(s) = F^*N(Nc'(N))$, which trivially implies $\int_{0}^{\theta} 1 - F^{*(N-1)}(Nc'(N) - s)dF(s) = 1 - F^*N(Nc'(N))$, one gets

$$\frac{BS(N)}{N} = \frac{1 + \eta^0}{1 + \lambda^m\eta^m} \int_{0}^{\theta} Q(s)(\gamma(s) - c'(N)(1 + \lambda^m\eta^m)) dF(s).$$

We have that

$$\int_{0}^{\theta} Q(s)(\gamma(s) - c'(N)(1 + \lambda^m\eta^m)) dF(s) < \int_{0}^{\theta} Q(s)(\gamma(s) - c'(N)) dF(s)$$

Now recall that

$$Q(s) = 1 - F^{*(N-1)}(Nc'(N) - s).$$

In particular $N\int_{0}^{\theta} Q(s)(\gamma(s) - c'(N)) dF(s)$ is the revenue from the classical mechanism with quasilinear utilities without news utility effects and involving average costs equal to $c'(N)$. But recall that we are considering the case $c'(N) > \bar{\theta}$. From the classical result (see [Güth, Hellwig '86] or section 3.3 in Chapter 3 of [Börgers '15]) it follows that this revenue is negative. Overall it follows that the result is not reversed with symmetric agents and timeline A. 

\[\square\]
In the following simple Proposition we illustrate the effect of news utility for revenue maximizing mechanisms in the public goods setting.

**Proposition 3.** Assume the designer wants to maximize expected revenue and the timeline is $A$ and that $\theta \geq 0$. The optimal provision rule is

$$q(\theta_1, \ldots, \theta_n) = \begin{cases} 1 & \text{if } (1 + \eta^g) \sum_{i=1}^n \gamma(\theta_i) \geq Nc(N) \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

The proof is just a small modification of the classical proof for revenue maximization. See for example [Börgers ’15].

The optimal rule reflects the surprise effect of news utility: given that the public good is always of value to the agent, the additional positive utility effect due to news utility has the same effect as a down-scaling of the costs of provision by a factor of $1 + \eta^g$.

### 4.2 Ex-post efficiency in bilateral trade and timeline $A$

The primitives for this setting are the ones from [Myerson, Satterthwaite ’83]: there is a buyer and a seller whose valuations are independently drawn from respective distributions $F_B, F_S$ over $[0, 1]$.

In our setting with news utility suppose as a first step there is complete and symmetric information about the intrinsic valuation $\theta_B$ of the buyer and $\theta_S$ of the seller. We define efficiency as maximizing the sum of agents’ utilities under complete information. We focus on timeline $A$. We have in mind a marketplace where transactions are done without delay, e.g. for already produced consumption goods or financial securities.

Denoting $q$ the probability of trade in a direct mechanism we assume for the intrinsic utilities of the seller and buyer: $v_S(q) = 1 - q$, $v_S(\emptyset) = 1$, $v_B(q) = q$ and $v_B(\emptyset) = 0$. In the presence of news utility and when agents are asymmetric, they may have differing marginal utilities of money even under complete information. Therefore, one needs to take into account the differing marginal utilities of money when defining ex-post efficiency in our setting. Namely, if it happens that $\lambda^m_S \eta^m_S < \eta^m_B$ then any positive transfer from the seller to the buyer is welfare-enhancing. A similar situation, but with transfers from the buyer to the seller holds if $\lambda^m_B \eta^m_B < \eta^m_S$. To avoid these purely monetary effects due to quasi-linearity under complete information, we constrain in the following the behavioral parameters of the agents in the money dimension.

**Assumption (NoM): No money effects under complete information.** We assume

$$\lambda^m_S \eta^m_S \geq \eta^m_B \geq \lambda^m_B.$$

---

4Recall, that $\theta \geq 0$ always.

5[Myerson, Satterthwaite ’83] consider distributions $F_B, F_S$ defined over not necessarily equal but always overlapping supports. We avoid these technicalities here and focus on the case of equal supports.

6This is in contrast to the assumption in [Myerson, Satterthwaite ’83] who assume equal marginal utility of money.

7Here we don’t consider agents with limited liability w.r.t money. Similar results to the ones presented here hold in a more general model with limited liability but that setting is more cumbersome and doesn’t yield additional interesting insights.
A special case of the assumption NoM is the case when $\eta_m^S = \lambda^m_B \eta_B^m$. We call this condition *equal perceived marginal utility of money*, whenever it is satisfied.

Under *equal perceived marginal utility of money* and timeline A in a trading mechanism it holds that the perceived negative disutility of the buyer for one additional expected money unit paid to the seller is equal in absolute value to the perceived utility of the seller for one additional expected money unit received.

Under assumption NoM maximizing efficiency under complete information requires implementing the allocation rule

$$ q(\theta_B, \theta_S) = \begin{cases} 1, & \text{if } (1 + \eta_B^g) \theta_B - \eta_S^g \lambda_S^g \theta_S \geq \theta_S \\ 0, & \text{otherwise.} \end{cases} \quad (11) $$

In comparison to the classical setting, news utility increases the perceived valuation of the buyer in the consumption dimension and decreases the perceived valuation of the seller in the consumption dimension. This leads c.p. to a higher willingness to pay for the buyer and c.p. a higher ‘ask price’ by the seller. Note that in the case of symmetric buyer and seller, efficiency requires less trade than in the absence of news utility.

We introduce the ‘friction’ coefficient in the good dimension

$$ c^g = \frac{1 + \eta_S^g \lambda_S^g}{1 + \eta_B^g}. $$

The ex-post efficiency rule under NoM can then be rewritten as

$$ q(\theta_B, \theta_S) = \begin{cases} 1, & \text{if } \theta_B \geq c^g \theta_S \\ 0, & \text{otherwise.} \end{cases} $$

The case $c^g = 1$ corresponds to ex-post efficiency rule in the classical case. It is easy to see that depending on news utility parameter values in the good dimension (11) can prescribe both more and less trade than in the classical case.

On an intuitive level it is unclear whether the impossibility result from [Myerson, Satterthwaite ’83] still holds under news utility effects, because with news utility there are additional, different effects at play. Intuitively, this is unclear even in timeline A, which is closest to the classical model. For example, loss aversion in the money dimension will be present for both seller and buyer whenever the mechanism to be implemented requires money transfers. If the seller expects to gain in the money dimension and the buyer to lose in the money dimension by participating in the mechanism, they will experience news utility/disutility in the money dimension in addition to the news utility in the consumption dimension due to exchange of the good. These two news utility effects are missing in the classical setting and under the NoM assumption, they don’t appear in the ex-post efficiency rule in our setting.

The following Proposition confirms that the balance of the additional effects under NoM is indeed so that the impossibility result is retained.

**Proposition 4.** For timeline A and under NoM, the ex-post efficiency rule cannot be implemented without a budget deficit.

**Proof.** We look at the ex-post efficiency rule as defined above.
Note that in our model \( c^g \) can be any positive real number. We also introduce \( c^m = \frac{1 + \lambda_B^m \eta_B^m}{1 + \eta_S^g} \), which can be interpreted as a friction coefficient of loss aversion in the money dimension.

For timeline A, where news utility in the money dimension in the first period plays a role for expected transfers, it holds with the ex-post efficiency rule in the equilibrium of the mechanism that a necessary condition for revenue to be maximized is that the seller always receives non-negative expected transfers and the buyer always makes non-negative expected transfers. Given this, and under \( \text{NoM} \), we see that the best case scenario for the highest budget surplus achievable is when \( \lambda_B^m \eta_B^m = \eta_S^m \). Thus, if we assume this equality we achieve an upper bound for the revenue of ex-post efficiency which is uniform across all parameter values which satisfy \( \text{NoM} \).

We assume \( \eta_S^m = \lambda_B^m \eta_B^m \) in the following.

We have

\[
W_B(s) = (1 + \eta_B^g) \frac{s}{c^g} F_S(s).
\]

Incentive compatibility requires \( W_B(s) \) to be non-decreasing in \( s \), which is given. Individual rationality for the buyer is standard, so let’s look at individual rationality for the seller. Note that in general we have

\[
W_S(s) = (1 + \lambda_S^g \eta_S^g)(1 - Q_S(s)) - \lambda_S^g \eta_S^g.
\]

With the ex-post efficiency rule this is

\[
W_S(s) = (1 + \lambda_S^g \eta_S^g) F_B(c^g s) - \lambda_S^g \eta_S^g.
\]

In an incentive compatible and individual rational mechanism, the latter property is equivalent to

\[
V_S(\theta) \geq \max_{\theta \in \Theta_S} \{ \theta - \int_\theta^\theta W_S(t) dt \}.
\]

Using Mirrlees representation this becomes equivalent to

\[
V_S(\theta) \geq \max_{\theta \in \Theta_S} \{ \theta + \int_\theta^\theta W_S(t) dt \},
\]

which can be rewritten into

\[
V_S(\theta) \geq -\lambda_S^g \eta_S^g \bar{\theta} + \max_{\theta \in \Theta_S} \{ (1 + \lambda_S^g \eta_S^g) \theta + \int_{\theta}^\theta (1 + \lambda_S^g \eta_S^g)(1 - Q_S(s)) ds \}
\]

(12)

One has \( (1 - Q_S(s)) - \frac{\lambda_S^g}{1 + \eta_S^g \lambda_S} Q_S(s) (1 - Q_S(s)) = (1 - Q_S(s)) (1 - \frac{\Lambda_S^g}{1 + \eta_S^g \lambda_S} Q_S(s)) \in [0, 1] \), so that the maximum is reached at the highest value \( \bar{\theta} \). In all, it follows individual rationality is equivalent to

\[
V_S(\bar{\theta}) \geq \bar{\theta}.
\]

Just as in the classical setting, individual rationality for all types is determined by the highest type of the seller.\(^8\)

\(^8\)See chapter 3 of [Börgers ‘15].
Recalling the Mirrlees representation from Section 2 in the paper one sees by standard partial integration as in the classical Bayesian mechanism design setting \(^9\) that

\[
T_B + T_S = \frac{1}{1 + \lambda_B \eta_B^m} \int_{\Theta_B} W_B(s) \gamma_B(s) dF_B(s) + \frac{1}{1 + \eta_S^{m}} \int_{\Theta_S} W_S(t) \gamma_S(t) dF_S(t) - \frac{\nu_B(\theta_B)}{1 + \eta_S^{m}} - \frac{\nu_B(\theta_B)}{1 + \lambda_B \eta_B^m}.
\]

It is easy to see that \(\nu_B(\theta_B) = 0\) and \(\nu_S(\bar{\theta}_S) = 1\) in a revenue-optimal mechanism, so that in all, with the assumption \(\eta_S^m = \lambda_B \eta_B^m\), one gets

\[
T_B + T_S = \frac{1 + \eta_B^g}{1 + \eta_S^g} \left[ \int_0^1 F_S(s) c^g \gamma_B(s) dF_B(s) + c^g \int_0^1 (F_B(s) - 1) \gamma_S(s) dF_S(s) \right].
\]

A substitution \(t = \frac{s}{c^g}\) in the first integral leads to

\[
T_B + T_S = \frac{1 + \eta_B^g}{1 + \eta_S^g} \left[ c^g \int_0^{\frac{c^g}{c^g}} F_S(s) c^g \gamma_B(s) c^g dF_B(s) c^g + c^g \int_0^1 (F_B(s) - 1) \gamma_S(s) dF_S(s) \right].
\]

But note that the expression in the brackets, up to the multiplicative constant \(c^g\) is the revenue of a classical, quasilinear model as in [Myerson, Satterthwaite ‘83], where the type of the seller comes from the distribution \(F_S\) with support on \([0, 1]\) and the type of the buyer comes from the distribution \(s \rightarrow F_B(c^g s)\) with support on \([0, \frac{1}{c^g}]\). Since the two supports overlap, the conditions for Corollary 1 in [Myerson, Satterthwaite ‘83] are satisfied and the revenue of the mechanism with \(\eta_S^m = \lambda_B \eta_B^m\) is negative. Recalling the discussion at the beginning of the proof establishes the result.

\[\square\]

### 4.3 Optimal Auctions in Timeline A with asymmetric agents

Recall that we keep the identical support assumption as well as the regularity properties of the distributions \(F_i\) of the intrinsic values.\(^{10}\) The only requirement we relax is that \(F_i = F_j\) for all \(i, j\). The perceived valuation of bidder \(i\) is

\[
\mathcal{W}_i(\theta) = (1 + \eta_i^g) Q_i(\theta),
\]

where \(Q_i(\theta)\) is the interim probability that type \(\theta\) of bidder \(i\) gets the good. Pointwise maximization give the following optimal rule for timeline A.

**Proposition 5.** The optimal auction for timeline A is an auction with agent-dependent reservation prices and has the following allocation rule.

\[
\text{Give the good to the bidder with the highest } \frac{1 + \eta_i^g}{1 + \lambda_i \eta_i^m} \gamma_i(\theta_i) \text{ as long as } \gamma_i(\theta_i) \geq 0. \quad (14)
\]

\(^9\)See for example chapter 3 in [Börgers ’15].

\(^{10}\)\(F_i\) have strictly positive, continuous density everywhere. Moreover, the virtual valuation of \(F\) given by the function \(\gamma(t) : [\bar{\theta}, \bar{\theta}] \rightarrow \mathbb{R}, \gamma(t) = t - \frac{1 - F(t)}{f(t)}\) is strictly increasing.

12
The proof is a simple adaptation of the results in the classical case. See e.g. [Börgers ’15].

The higher the weight of news utility in the good dimension and the lower loss aversion and the weight of news utility in the money dimension the higher the revenue contribution of a bidder. When there is no news utility in the good and money dimension, i.e. \( \eta_g^i = \eta_m^i = 0 \) (14) falls back to the classical formula from [Myerson ’81], where only informational rents matter. The reservation types, i.e. the types \( \theta_i^* \) where \( \frac{1+\eta_g^i}{1+\lambda_i \eta_m^i} \gamma_i(\theta_i) = 0 \) are then the same as in the classical analysis.

In the general setting the efficient allocation would require giving the good to the agent with the highest \((1 + \eta_g^i)\theta_i\). Recall that in the absence of news utility and with symmetric bidders the optimal auction is constrained efficient in the sense of either giving the good to the agent with the highest intrinsic valuation for the good or not selling at all. In our model this remains true if we assume bidders with symmetric preferences. If we assume instead different degrees of loss aversion in the money dimension, but otherwise identical distributions \( F \) of intrinsic valuations the optimal auction may require allocating the good to a bidder whose \((1 + \eta_g^i)\theta_i\) is not the highest, in case she is relatively insensitive to news utility in the money dimension, i.e. in case her \( \lambda_i \eta_m^i \) is relatively small in comparison to other agents.

4.4 Optimal Auctions in Timeline B with symmetric agents

We assume \( N \geq 2 \) symmetric bidders. For each bidder there are two news utility effects in timeline B. A positive surprise effect of participation, which stems from the positive news of potential future consumption in comparison to the situation of non-participation and the negative expected realization effect originating from the shift of expectations induced by the delay, once the agent has participated. The second effect is negative due to loss aversion, since bad news are overweighted in comparison to good news. The second news effect may dominate when the intrinsic valuation for the good is low and the agent’s preferences are sufficiently loss averse. In contrast to the classical set-up without loss aversion, a bidder’s perceived valuation in the good dimension can thus become negative, if her loss aversion in the good dimension is high enough. The auctioneer optimally excludes agents whose perceived valuation \( W_i \) is negative, since he would need to subsidize their participation otherwise. Exclusion for such low types can be done through a reservation price, just as in the classical setting. But while in the classical setting the reservation type is only set to limit the rent of types with high intrinsic valuation for the good, the optimal reservation price with news utility has to be high enough to take into account agents who have negative perceived valuation. It follows that c.p. the reservation type will be weakly higher than in the classical setting without loss aversion.

Denote by \( \theta^{(1)} \) the value where \( \gamma_i(\theta^{(1)}) = 0 \). This is the reservation type in the classical analysis (see also [Myerson ’81] or [Börgers ’15]). Define \( q \) the allocation rule which gives the good to the type with the highest reported \( \theta_i \) and denote by \( \theta^{(2)} \) the value where the perceived valuation \( W_i^B \) of \( q \) satisfies \( W_i^B(\theta^{(2)}) = 0 \). We established intuitively that it is never optimal to give the good to a bidder with \( W_i(\theta) < 0 \). Recalling Theorem 1, we prove the following result.

**Proposition 6.** The optimal auction with symmetric bidders and timeline B is an all-pay auction with the following allocation rule

\[
\text{Give the good to the bidder with the highest } \gamma_i(\theta_i) \text{ as long as } \theta \geq \theta^* = \max\{\theta^{(1)}, \theta^{(2)}\}.
\]
In the optimal auction agents with $\theta < \theta^*$ make zero payments. The reservation price is $\frac{1}{1 + \lambda^m} W_i(\theta^*) \theta^*$.

**Proof of Proposition 6.** Let $Q_i(\theta_i)$ be the interim probability that type $\theta_i$ of bidder $i$ gets the good. Note that

$$W_i(\theta) = (1 + \eta^\theta)Q_i(\theta) - \Lambda^\theta Q_i(\theta)(1 - Q_i(\theta))$$

This is convex in $Q$. Moreover, $Q_i$ are linear in the $q_i$-s. Namely, they are given by

$$Q_i(\theta_i) = \int_{\theta_i}^{\bar{\theta}} f(\int q_i(\theta_i, \cdot) dF^{N-1}(\cdot)) \gamma(\theta_i) dF(\theta_i),$$

and the max is taken over functions $q : [\theta, \bar{\theta}]^N \rightarrow \Delta$, where $\Delta = \{q : \sum_i q_i \leq 1, q_i \geq 0\}$. The latter is a convex set.

The integrand in the maximization problem is of the form

$$\sum_{i=1}^{N} f \left( \int q_i(\theta_i, \cdot) dF^{N-1}(\cdot) \right) \gamma(\theta_i)$$

with a $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ strictly convex, given by $f(x) = (1 + \eta^\theta)x - \Lambda^\theta x(1 - x)$. Note that if $\Lambda^\theta \leq (1 + \eta^\theta)$, then $f \geq 0$ and otherwise $f$ is parabolic with zeros at $x = 0$ and $x = 1 - \frac{1 + \eta^\theta}{\Lambda^\theta}$. We consider both cases in the following.

**Observation 1:** It is sufficient to find the optimal allocation rule $q$ only for $\theta$-s s.t. no two components of the vector $\theta$ are equal.

This is because $N$-dimensional Lebesgue measure puts zero measure to sets where two components of $\theta$ are equal. Once the optimal rule is specified for the case of no-ties, it can be easily extended to an optimal rule with ties and so that measurability is preserved. In the following we therefore assume we are in the case where $\theta$ is strictly ordered.

**Observation 2:** It is without loss of generality to look at symmetric allocation rules. The reason is the same as in classical setting (see footnote 11 in [Maskin, Riley '84]). This means the optimal allocation rule $q$ will be permutation invariant in the sense that for each permutation $\pi$ of $\{1, \ldots, N\}$ we have

$$q_i(\theta_1, \theta_2, \ldots, \theta_N) = q_{\pi(i)}(\theta_{\pi(1)}, \ldots, \theta_{\pi(N)}).$$

From this it follows $Q_i(\theta) = Q_{\pi(1)}(\theta)$ for all $\theta, i$. It follows that it is enough to determine the optimal allocation rule on $\{\theta_1 > \theta_2 > \cdots > \theta_n\}$. We concentrate on this subset in the following.

**Observation 3:** The optimal allocation rule will have

either $\sum_i q_i = 0$ or $\sum_i q_i = 1$.

That is, the good is either not sold or sold to one agent only.
To see this, suppose that it holds $1 > \sum_i q_i > 0$. Then one can strictly increase revenue \(^\text{11}\) by pushing all the weight to $\theta_1$.

**CASE 1:** $\Lambda^g \leq (1 + \eta^g)$

Recall from the main text the definition of the threshold type $\theta^{(1)}$ with $\gamma(\theta^{(1)})$.

**Observation 4:** It is optimal to set $q_i(\theta_i, \theta_{-i}) = 0$ for all $\theta_{-i}$, if $\theta_1 \leq \theta^{(1)}$.

This is clear, since in (16) the summands where $\gamma(\theta_i) \leq 0$ are then zero.

**Observation 5:**

Whenever $\theta_1 > \theta^{(1)}$, it is optimal to set $q_i(\theta_1, \theta_{-1}) = 1$.

This is again clear, due to (16).

By observations 1 and 2, this concludes the construction of the optimal $q$ for **CASE 1**.

**CASE 2:** $\Lambda^g > (1 + \eta^g)$

Recall that $W_i(Q)$ is a parabola with zeros at 0 and $1 - \frac{1+\eta^g}{\Lambda^g}$. Moreover, for $Q \geq \frac{1}{2} \left( 1 - \frac{1+\eta^g}{\Lambda^g} \right)$ the parabola is increasing in $Q$, so that for incentive compatibility to hold it should be the case that $Q_i(\theta)$ is nondecreasing, if $Q_i(\theta) \geq \frac{1}{2} \left( 1 - \frac{1+\eta^g}{\Lambda^g} \right)$. Note from **CASE 1**, that the resulting $Q_1$ there is given by

$$Q_i(\theta) = \begin{cases} F^{N-1}(\theta) & \text{if } \theta > \theta^{(1)} \\ 0 & \text{otherwise.} \end{cases}$$

Let $\theta^{(2)}$ be so that $F^{N-1}(\theta^{(2)}) = \left( 1 - \frac{1+\eta^g}{\Lambda^g} \right)$. It is clear that the optimal allocation rule is the same as in Case 1, whenever $\theta_1 > \max\{\theta^{(1)}, \theta^{(2)}\}$.

If it happens that $\theta_1 \in (\theta^{(1)}, \theta^{(2)})$, then having $\sum_i q_i = 1$ (which is w.l.o.g. due to Observation 3 and which is prescribed by the optimal allocation rule in **CASE 1**) results in $W_i(\theta_1) < 0$ and $\gamma_i(\theta_1) > 0$. Moreover, due to individual rationality we have

$$V_i(\theta) = -\int_\theta^{\theta^{(2)}} W_i(s) ds > 0.$$  \hspace{1cm} (17)

This term lowers expected payments for all types as can be seen from the Mirrlees representation in Proposition 6 of the main body of the paper.

Consider instead setting $\sum_i q_i = 0$ if $\theta_1 \in (\theta^{(1)}, \theta^{(2)})$ and otherwise following the optimal allocation rule from **CASE 1**. This change results in an allocation rule $\tilde{q}$, which is again incentive compatible, as for $\theta_1 > \theta^{(2)}$ one has $W_i(\theta) > 0$ under $\tilde{q}$. Under $\tilde{q}$ it holds that $V_i(\theta) = 0$ and thus expected payments are c.p. higher.\(^\text{12}\) Moreover, from (16) it follows that the pointwise integrand of the maximization problem in (15) is higher under $\tilde{q}$. By the same logic as in **CASE 1** it is clear that there is no possibility to raise revenue by changing the rule, whenever $\theta_1 > \max\{\theta^{(1)}, \theta^{(2)}\}$.

This, the arguments above, coupled with Observations 1 and 2 yield that the optimal allocation rule is to sell the good to the highest bidder, whenever its type is above $\max\{\theta^{(1)}, \theta^{(2)}\}$ and otherwise to not sell the good.

We summarize: the optimal auction is an all-pay auction where types with $\theta < \theta^* = \max\{\theta^{(1)}, \theta^{(2)}\}$ don’t get served and otherwise the highest bidder wins. This allocation rule can be realized in an indirect mechanism by setting a reservation prize of $\beta_i(\theta^*)\theta^*$. \(\Box\)

\(^{11}\)Recall here, that we assume Regularity of virtual valuations, that is, $\gamma(\theta)$ is a strictly increasing function of $\theta$.

\(^{12}\)The designer saves on the individual rationality constraints.
References


