The Design of Optimal Collateralized Contracts*

Dan Cao†
Georgetown University

Roger Lagunoff‡
Georgetown University

March, 2016

Abstract

This paper examines the nature of collateral in loan contracts. We study a standard two-period contracting problem in which a risk averse borrower contracts with a risk neutral lender to obtain a loan under asymmetric information. The borrower values a capital good and a composite non-capital good. He privately observes an income shock in the composite good in the second period. Both participants can fully commit to the contract ex ante. There are no contractual frictions other than asymmetric information. Under general assumptions, collateralization of both goods emerges as an endogenous feature of the incentive constrained optimal contract, whereas it does not under full information.

The optimal contract exhibits a distortion in which consumption is excessive in the initial loan period and the capital good is under-consumed in the repayment period. The contract is strictly separating for high income types, but requires forfeiture of collateral and possibly pooling at the bottom. Forfeiture of the collateralized assets are locally (weakly) decreasing in income for very high and very low income earners. We give a precise characterization of the optimal contract in a parameterized model, and a straightforward algorithm for computing optimal contracts more generally.

When embedded in general equilibrium, the optimal contract leads to lower relative price of capital good compared to a GE model of full information.

*First Draft: March 24, 2016. We thank Axel Anderson, Florian Scheuer, and numerous seminar participants for useful comments and suggestions.

†Department of Economics, Georgetown University, Washington DC 20057, *dc448@georgetown.edu* .

‡Department of Economics, Georgetown University, Washington DC 20057, *lagunofr@georgetown.edu* .

1 Introduction

Collateralized contracts arise in many types of credit transactions. These include durable goods purchases, consumption loans, and equity stakes. Typically, such contracts specify a tangible pledge of property to be forfeited in the event the borrower cannot or will not repay the loan.

The credibility of the forfeiture threat is thought to be a critical element in determining whether and what form of collateral is used. Collateral requirements can, for instance, prevent borrowers who are otherwise able to repay from strategically defaulting and/or misusing the loaned funds.[1] Collateral is also shown to be useful when lenders have imperfect information about borrowers' abilities to repay. The credible threat of forfeiture in low states can indemnify the lender against the borrower's incentives to lie or withhold critical information.[2]

Generally, these rationales are a part of a more general inquiry into the nature of credit markets. Collateral is a key ingredient, for instance, in explanations of credit rationing (Stiglitz and Weiss (1981)), debt contracts (Lacker (2001)), inefficient screening (Manove et al. (2001)), and capital structure (Rampini and Viswanathan (2013) and Parlatore (2016)). The contract in these explanations is an optimal solution to a constrained mechanism design problem in limited liability, limited commitment, or course states of nature are used to produce instruments that intuitively resemble collateral.

This paper proposes a mechanism design model of endogenous collateral drops all constraints on liability, enforcement, or commitment. There are no arbitrary bounds on the types of loan contracts that can be offered. We separate out capital structure and credit market concerns. Instead, we focus solely on the role of private information in explaining why certain types of goods are included and acquire collateral-like features, and how the forfeiture requirement varies over the borrower's realized income. We characterize an *optimal contract* as solution to a general mechanism design problem subject to information and resource constraints. We then show that under general assumptions such a contract will entail pledges and forfeiture of collateral in certain states of nature.

We posit a familiar contracting scenario with two agents, a lender and a borrower. The lender is risk neutral, the borrower risk averse. There are two periods. In each pe-

[1] Explanations of this type have a long tradition, dating back at least to Barro (1976).

[2] Chan and Kanatas (1985), Bester (1985), and Besanko and Thakor (1987), are early examples of this second rationale.

riod, the borrower values composite consumption and an additional capital good that generates use value for the borrower. The borrower is endowed with little or nothing in the first period. In the second period he receives a random income/return in composite consumption units. The realized income is determined from a continuous distribution; low realizations correspond to a negative shocks such as job loss or poor returns on investments. Consequently, the borrower seeks a loan contract that allows him to smooth across time and across his uncertain income stream.

The capital good is non-stochastic. It may or may not be liquid, and may or may not depreciate rapidly. When all contracting parties observe the same information, the resulting contract results in constant consumption of both the composite and the capital good, independently of the state. In other words, the lender fully insures the borrower in the stochastic composite good, and the consumption of the capital good does not vary with the income shock. Thus, the capital good is not used in the contract in any meaningful way. It can be shown in this case, that the amount of wealth of the borrower that can be contractually seized by the lender — the contracted *collateral* — is zero. In this sense, the full information optimal contract is not collateralized.

Under incomplete information the borrower's realized income in the composite good in the second period is privately observed by the borrower. A contract specifies consumption of the two goods in period 1 and contingent consumption of the goods in period 2 given the reported income of the borrower. An *optimal contract* maximizes the borrower's expected two-period payoff subject to a standard IC constraint and a 0-profit constraint of the lender.

The capital good in this case is essential in maintaining truthful reporting incentives while allowing some (imperfect) risk-smoothing in the composite good. We show, in fact, that the optimal contract is collateralized, and the capital good is instrumental in making this so. Our findings, based on private information, complement those of Rampini and Viswanathan (2013) who shows that tangible collateral emerges under a viable repossession technology.

Optimal contracts exhibit a distortion in both the intra-temporal marginal rate of substitution between capital and composite consumption and the inter-temporal rate of substitution between first and second period consumption. The distortion is such when capital and non-capital goods are complements, the borrower under-accumulates collateralizable assets in the initial period. In the repayment period all income types of the borrower except the highest, and possibly the lowest type under-consumes the capital good and over-consumes the composite good. Consequently, the composition of seized collateral is excessively tilted toward the capital good.

We also establish some distributional properties of the contract. Optimal contracts in the repayment period are shown to be strictly separating at the top of the income distribution, but may admit pooling and full forfeiture of collateral at the bottom Forfeiture, the value of seized collateral, is shown to be locally decreasing in income at the very top and very bottom of the income distribution.

More generally, we derive a formula relating the distortion to the marginal propensity for forfeiture in the optimal contract. Given an extra dollar of income, higher distortions are associated with larger increases in forfeiture, and they occur toward the middle of the income distribution. The nature of these distortions is similar to the tax wedges in Mirleesian tax models (Mirrlees (1971, 1976)).

Along these lines, we give a precise characterization of the optimal contract in a parameterized model with log preferences and a uniform distribution. There, forfeiture is strictly decreasing in income, and consumption of the capital good is shown to be convex in realized income. Combining this with the inter-temporal distortion, this means that, compared with the full information optimum, a low income borrower's asset accumulation is too small. At the same time, too much of this accumulation is forfeited.

To our knowledge these results are not found in the literature. These findings are, moreover, consistent with mortgage contracts in housing. Data from the Consumer Expenditure Surveys shows that among households with mortgages, housing consumption is roughly convex in income, and this is not generally true of households without mortgages.[3]

The next section, Section 2, gives a brief review of the literature. Section 3 introduces the baseline model. We introduce the planner's problem and and define what it means for the resulting optimal contract (OC) to be "collateralized." Section 4 contains the main characterization results. Section 5 presents a closed form solution for the OC in a parametric model. Section 6 describes methods for computing OCs numerically and tackles robustness issues and describes data consistent with the collateralized OC in a parametric special case of the model. Section 7 embeds are framework in a GE model. The proofs are contained in an Appendix at the end.

[3]See Subsection 6.3 for details.

2 The Literature

The present paper builds upon two distinct, though overlapping, literatures. The first is the large literature on Mirleesian contracts with private information. Our model is in many ways a throwback to first-generation models in that literature, dating back to Mirrlees (1971, 1976) himself. Following in the tradition of these models, we specify general preferences for the borrower that are non-separable in the two goods.[4]

Because our focus is on credit rather than on taxation and labor effort, the random shocks here are additive income shocks rather than multiplicative productivity shocks in the optimal taxation literature.[5] This seemingly minor restriction leads to a sharp characterization of the optimal contract, particularly as it relates to collateral.

A second, related literature concerns the specific role of collateral in credit contracts. These studies characterize contracts under various constraints in order to produce instruments that intuitively resemble collateral. These constraints may be motivated by legal restrictions such as limited liability or limited commitment as with Stiglitz and Weiss (1981), Wette (1983), Lacker (2001), Rampini and Viswanathan (2013).[6] In other cases, the models restrict attention on basic debt instruments in order to focus on particular attributes of the collateralized contract. Examples include Chan and Kanatas (1985), Bester (1985), Besanko and Thakor (1987), and Rampini (2005) who derive discontinuous threshold requirements in debt contracts, Manove et al. (2001) who examine banks' trade off between screening or collateralizing loans to risky projects. Eisfeldt and Rampini (2009) who evaluate the borrower's choice of whether to lease or purchase the collateralized asset, and Campbell and Cocco (2015) who study strategic default by borrowers.

Among these, the closest models to our own are Rampini (2005) and Lacker (2001). Both consider a private information contracting model with an additive shock structure. Rampini (2005) incorporates additive shocks to analyze the role of collateral in aggregate default. Incorporating non-monetary default penalties that do not affect lenders' profits, he derives a cutoff rule in which the default rate can be discontinuous in income. Lacker (2001) examines bilateral credit contracts with private information and a limited commitment assumption. The limited commitment introduces the potential for renegotiation.

[4] For tractability, later generations of models often assume separability and even quasi-linearity in preferences. Non-separability is an important generalization for certain types of collateral since increased consumption of durables such as housing, may increase the marginal utility of nondurable consumption.

[5] The literature is vast, though Diamond (1998) gives canonical treatment used in later models.

[6] See references contained therein.

This bounds both the size of the initial loan and the amount of the capital good that can be credibly collateralized and consumed in his model. As a result, the optimal contract pools types at the upper, rather than lower, end of the distribution.

These constraints no doubt play a significant role in collateralized contracts. Our goal is to see how much mileage can be obtained by private information alone. Thus we allow for full commitment, no constraints on the loan size or collateral, and are still able to derive contracts with collateral. The substantive differences in these contracts highlight the role of information as apart from other constraints in determining the role of collateral.

A purely private information-based notion of liquid collateral appears in an influential paper by Ed Green (1987).[7] Green formulated a dynamic contracting model in which the borrower (agent) possesses private information each period about an income shock to a single composite good. A credit contract emerges with collateral-like features. The availability of multi-period consumption trade offs plays roughly a similar role as our trade offs between composite and capital consumption. In either case, collateralization requires bootstrapping the contract on an "extra" good not subject to shocks in that same period. Our collateralized contract will have very different features than Green's credit contract, owing to the distinct attributes of capital good consumption and to the complementarities between the capital and non-capital good consumption.

3 A Baseline Model

3.1 Overview

The most basic scenario involving collateral is a two-period contracting problem with two agents, a lender and a borrower. In the first period, the borrower needs external funding. Depending on interpretation, external funding may be required for a variety of reasons. Businesses have insufficient internal resources to fund an expansion. Entrepreneurs lack funds for investment opportunities. Households lack accumulated wealth to fund durable goods purchases.

The simplest specification of payoff that accommodates the framework is of the form $U_0(c_0, k_0) + \beta\, U(c, k)$ for the borrower. Composite consumption good c_0 is consumed at

[7] This is also true of many successors that his paper brought forth.

$t = 0$ and c is consumed $t = 1$. Goods k_0 and k are service-generating capital goods that will, eventually, play a critical part in the collateralization of a loan. The capital good can be stored subject to partial depreciation.

The flow payoff function U is assumed concave and strictly increasing in each good. When the borrower is a household, strict concavity is a natural assumption. We assume further that (i) U is twice continuously differentiable, and (ii) satisfies a weak complementarity condition given by

$$\frac{\partial^2 U}{\partial c \partial k} \; > \; \max\left\{ \frac{\partial U/\partial c}{\partial U/\partial k}\frac{\partial U^2}{\partial k^2}, \; \frac{\partial U/\partial k}{\partial U/\partial c}\frac{\partial U^2}{\partial c^2} \right\}, \tag{1}$$

for all c, k. Because the right-hand side is negative, the inequality in (1) is clearly satisfied when c and k are complements, i.e. $\frac{\partial^2 U}{\partial c \partial k} \geq 0$. The inequality, however, also allows for c and k to be substitutes, provided that the interaction is not large.[8]

The borrower starts the initial period with composite wealth A. In the case of interest, A is low enough that the borrower needs a loan to smooth consumption across the two dates. At the beginning of the second period, the borrower realizes a random income/return θ of the composite good, distributed according to F on support $[\underline{\theta}, \overline{\theta}]$. The distribution F admits a continuous density f. Though F is common knowledge, the realized value θ is privately observed only by the borrower. (An alternative, "Mirleesian," interpretation is that a lender in a competitive credit market draws a borrower from population distribution F.)

Goods y_0, k_0, and k are non-stochastic. The price of the composite good is normalized to 1 each period, and the prices of the capital good are q_0 and q, resp. For its part, the lender is risk neutral and belongs to a perfectly competitive set of intermediaries, all of whom offer loans with a market return of R.

3.2 A Contracting Problem

A *loan contract* or simply a *contract* between the lender and borrower consists of a list $(y_0, k_0, \; y, k)$ such that y_0 and k_0 units of composite and capital goods, resp., are offered to the borrower in the initial period, resulting in consumption $c_0 = y_0$ and k_0 for the

[8]In Lemma 3 in the Appendix, we show that condition (1) implies that both goods, c and k, are strictly normal goods.

borrower. In the last (repayment) period, consumption is contingent on realized income, and so y and k are functions mapping types θ to units $y(\theta)$ and $k(\theta)$ of the composite and capital good, respectively. This, in turn, results in second period consumption $c(\theta) = \theta + y(\theta)$ and $k(\theta)$.

Taking the perspective of the borrower ex ante, we characterize the optimal contract taking perfect competition in the lending sector as given.

Definition. An *optimal contract (OC)* is a list $(y_0, k_0,\ y, k)$ that solves:

$$V(A, q_0, q, R) \ = \max_{y_0, k_0, y(\cdot), k(\cdot)} U_0(y_0, k_0) + \beta \int_{\underline{\theta}}^{\overline{\theta}} U\left(\theta + y(\theta), k(\theta)\right) dF(\theta) \tag{2}$$

subject to

$$y_0 \ + \ q_0 k_0 \ + \ \frac{1}{R} \int_{\underline{\theta}}^{\overline{\theta}} \left(y(\theta) + qk(\theta)\right) dF(\theta) \ \leq \ \frac{1}{R} q k_0 + A \tag{3}$$

and

$$U\left(\theta + y(\theta), k(\theta)\right) \ \geq \ U\left(\theta + y(\hat{\theta}), k(\hat{\theta})\right) \quad \forall\, \theta,\ \hat{\theta}. \tag{4}$$

The resource constraint is displayed in (3), the incentive constraint in (4). In the resource constraint A is the borrower's initial wealth, $R > 1$ is the gross return on capital, and q_0 and q are the prices of first and second period capital goods, respectively.[9] A no-arbitrage condition requires that $q_0 > \frac{q}{R}$. The constraint in Equation (3) may be interpreted as a zero profit or participation constraint for a competitive lender.

If A is large enough, it is possible that the borrower does not borrow at all. The interesting case is when A is sufficiently small so that in the optimal contract the borrower obtains a loan for first period consumption in the amount $y_0 \ + \ k_0(q_0 - \frac{q}{R}) - A > 0$. The resource constraint then guarantees that the first period loan is fully repaid in expectation.

The incentive constraint in (4) involves only the *continuation contract* (y, k) received in period 2. The following characterization of IC s standard.

Lemma 1. *A continuation contract (y, k) satisfies the incentive constraint if and only if for $\theta >$*

[9] Depreciation is implicitly captured in q.

θ', we have $y(\theta) \leq y(\theta')$ and $k(\theta) \geq k(\theta')$ and

$$\begin{aligned} \frac{dU^-}{d\theta}(\theta + y(\theta), k(\theta)) &\geq \frac{\partial U}{\partial c}(\theta + y(\theta), k(\theta)) \\ \frac{dU^+}{d\theta}(\theta + y(\theta), k(\theta)) &\leq \frac{\partial U}{\partial c}(\theta + y(\theta), k(\theta)). \end{aligned} \tag{5}$$

where $\frac{dU^-}{d\theta}$ and $\frac{dU^+}{d\theta}$ are the left and right one-sided derivatives respectively.

In the case where the one-sided derivatives coincide, (5) is the familiar envelope condition. An *incentive compatible* continuation contract is any pair (y, k) satisfying the monotonicity and envelope conditions of Lemma 1.[10]

Thus far, the theory resembles a standard Bayesian mechanism design problem. Unlike many design problems, payoffs are not quasi-linear, nor even necessarily separable in c and k (later, we do examine the special case of separable payoffs). Additionally, there are no interim participation constraints, making our problem somewhat closer to the Mirleesian tax problem than a Myersonian bargaining problem.

3.3 Collateral: A Definition

In order to ascertain when a contract involves "collateral" and how much is required, one can calculate what the borrower will actually forfeit in the contract. To do this, we first identify the borrower's expected net worth heading into the second period. This is given by $E[\theta] + B$ where $B \equiv qk_0 + R(A - q_0k_0 - y_0)$ is the asset value (if positive) or debt obligation (if negative) carried in to the second period. The inclusion of expected future income $E[\theta]$ in his net worth reflects the fact that the borrower has access to the credit market.[11] Next, we identify the type-contingent value of his realized consumption from the contract in that period: $qk(\theta) + \theta + y(\theta)$. Hence, the value of collateral seized by the lender is

$$\Gamma(\theta) \equiv E[\theta] + B - qk(\theta) - \theta - y(\theta). \tag{6}$$

In other words, the seized or forfeited collateral is the difference between the borrower's expected net worth and the value of his total consumption in the last period. By this

[10] Though fairly standard, we provide a proof of Lemma 1 in the Appendix.

[11] A perfectly functioning credit market ensures that the expected value $E[\theta]$ of one's promised compensation from employment is received by the borrower.

definition, forfeiture collateral can be positive or negative.[12] When seizure is smaller than the pledged amount, the difference is, in effect, returned to the borrower. This is the case in mortgage contracts when, for instance, the value of the house exceeds the outstanding debt on the mortgage, or when the self-generated component of one's income is higher than expected. The *value of the collateral pledged by the borrower* is therefore that maximal amount that the lender can seize in the last period:

$$\overline{\Gamma} \equiv \max_{\theta \in [\underline{\theta}, \overline{\theta}]} \Gamma(\theta) \tag{7}$$

Definition. A contract is *collateralized* if $\overline{\Gamma} > 0$.

No other constraints are added to derive a collateralized solution. Both participants can fully commit to the contract ex ante, and there is no renegotiation. There are no liability or enforcement restrictions. There are no borrowing or liquidity constraints, and no moral hazard.[13] Asymmetric information alone is the key.

Indeed, without either asymmetric information or functional credit markets, the optimal contract is *not* collateralized. To see why, consider the following three benchmarks.

1. The full information optimal contract. In the absence of the IC constraint, the solution to (2) corresponds to a *full information optimal contract* $(y_0^\circ, k_0^\circ, y^\circ, k^\circ)$. One can easily verify that under full information, the contract provides provides full insurance. That is, $y(\theta) = c^\circ - \theta$ for some constant c° fully insuring the borrower in the composite good, and $k^\circ(\theta) = k^\circ$ is constant so that the borrower incurs no risk in his consumption of the capital good. The usual optimality conditions equating marginal rates of substitutions to relative prices are satisfied:

$$\frac{\partial U/\partial k}{\partial U/\partial c}(\theta) = q, \quad \frac{\partial U_0/\partial k_0}{\partial U_0/\partial c_0} = q_0 - \frac{q}{R}, \quad \frac{\partial U_0}{\partial k_0} = \left(\frac{Rq_0 - q}{q}\right)\frac{\partial U}{\partial k}(\theta), \quad etc.$$

2. The self-funded allocation. In the absence of *any* contracting option, the borrower faces autarky. The second period consumption in that case satisfies the type-by-type budget constraint $qk^{aut}(\theta) + y^{aut}(\theta) = B^{aut} \equiv qk_0^{aut} + R(A_0 - q_0 k_0^{aut} - y_0^{aut})$. The type-by-type optimality condition also holds under autarky, but without the benefit of complete insur-

[12]This must be so since $\int \Gamma(\theta) dF = 0$.

[13]Without moral hazard or market incompleteness, expected future income can be used as collateral.

ance. While each type's chosen mix of consumption between the two goods is optimal, consumption differences across types exposes the borrower to ex ante risk.

The key take-away from these two benchmarks is that neither of them is collateralized. The self-funded contract is not collateralized by definition. Without access to the credit market, the relevant net worth of the would-be borrower is his realized net worth in the second period. That is, $\theta + B$. Hence, the collateral "forfeited" is $\theta + B^{aut} - qk^{aut}(\theta) - \theta - y^{aut}(\theta) = 0$ for all θ.

The full information contract is not collateralized either. To see why, notice that the capital good is not needed to achieve full insurance. The contractual allocation of y fully suffices to fully insure the borrower. Since $k^\circ(\theta) = k^\circ$ is constant and $y^\circ(\theta) = c^\circ - \theta$ for some constant c°, it follows that $\Gamma^\circ(\theta)$ is constant: $\Gamma^\circ(\theta) = \overline{\Gamma}^\circ$. Because the resource constraint will be satisfied with equality, we obtain $\overline{\Gamma}^\circ = 0$. In other words, value of seized collateral for every type θ is zero. No collateral is pledged. None is seized.

Now consider a third benchmark:

3. The model without capital. In the absence of a capital good k, an optimal contract is a pair (y_0^*, y^*). It's easy to see that the only incentive compatible contract is full pooling. If it were not so, then by Lemma 1, any high type obtains a higher transfer by mimicking a lower type. Thus $y^*(\theta) = \bar{y}$, and so consumption is perfectly correlated with the income shock θ. Under full pooling $\Gamma(\theta) = E[\theta] + B - \bar{y} - \theta$. Notice that $\bar{y} = B$. If $B < 0$ representing debt, then $\bar{y}$ is the non-contingent repayment of the loan. Forfeiture thus reduces to $\Gamma(\theta) = E[\theta] - \theta$, the loss in future income relative to its predicted value. So, while the contract is collateralized (since $\overline{\Gamma} > 0$), collateral is comprised solely of future income. No part of the borrower's net worth is included.

3.4 The Saddle Point Problem

The entire contract (y_0^*, k_0^*, y^*, k^*) is chosen by the planner *before* the borrower self-selects his type. Consequently optimal contracting problem can be restated as a Lagrangian saddle problem. The solution coincides with the HBJ formulation commonly used in Mirleesian mechanism design problems.

Let $\lambda \geq 0$ be the multiplier associated with inequality constraint (3) for the budget and let $\xi(\theta) \in \mathbb{R}$ the multiplier associated with envelope constraint in (5). As for the monotonicity constraint, one can find a function $j(\theta)$ satisfying

$$k(\theta) = k(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} j(\tilde{\theta}) d\tilde{\theta}. \tag{8}$$

The monotonicity condition on $k(\theta)$ then implies

$$j(\theta) \geq 0. \tag{9}$$

Now let $\eta(\theta)$ denote the multiplier on the equality constraint in (8), and $\gamma(\theta)$ the multiplier on the non-negativity constraint in (9).

Because the envelope constraints associated with multipliers $\xi(\theta)$ and $\eta(\theta)$ are equality constraints, both $\xi(\theta)$ and $\eta(\theta)$ can be either positive or negative. On the other hand, the constraints associated with λ and $\gamma(\theta)$ are inequality constraints and so $\lambda \geq 0$ and $\gamma(\theta) \geq 0$ are required for all income types θ.

With the multipliers, the Lagrangian function can be stated as,

$$\begin{aligned} \mathcal{L}\left(y_0, k_0, y, k, j, k(\underline{\theta}),\ \xi, \eta, \gamma, \lambda\right) \ &\equiv \ U_0(y_0, k_0) + \\ \beta \Bigg\{ \int_{\underline{\theta}}^{\bar{\theta}} &\left(U\left(\theta + y(\theta), k(\theta)\right) \ + \ \lambda \left(B - y(\theta) - qk(\theta)\right) \right) f(\theta) d\theta \\ + \int_{\underline{\theta}}^{\bar{\theta}} &\xi(\theta) \left(\frac{\partial U}{\partial \theta}\left(\theta + y(\theta), k\left(\theta\right)\right) \ - \ \frac{\partial U}{\partial c}\left(\theta + y(\theta), k\left(\theta\right)\right) \right) d\theta \\ + \int_{\underline{\theta}}^{\bar{\theta}} &\left(\eta(\theta) \left(k(\theta) - k(\underline{\theta}) \ - \ \int_{\underline{\theta}}^{\theta} j\left(\tilde{\theta}\right) d\tilde{\theta} \right) \ + \ \gamma(\theta) j\left(\theta\right) \right) f(\theta) d\theta \Bigg\} \end{aligned} \tag{10}$$

where, recall, $B \equiv qk_0 + R(A - q_0 k_0 - y_0)$.

A standard application of the Karush-Kuhn-Tucker Theorem shows that any saddle point of $\mathcal{L}$, i..e. any list $(y_0^*, k_0^*, y^*, k^*, j^*, k^*(\underline{\theta});\ \xi^*, \eta^*, \gamma^*, \lambda^*)$ satisfying

$$\begin{aligned} &\mathcal{L}\left(y_0, k_0, y, k, j, k(\underline{\theta}),\ \xi^*, \eta^*, \gamma^*, \lambda^*\right) \ \leq \ \mathcal{L}\left(y_0^*, k_0^*, y^*, k^*, j^*, k^*(\underline{\theta}),\ \xi^*, \eta^*, \gamma^*, \lambda^*\right) \\ &\leq \ \mathcal{L}\left(y_0^*, k_0^*, y^*, k^*, j^*, k^*(\underline{\theta}),\ \xi, \eta, \gamma, \lambda\right) \end{aligned} \tag{11}$$

corresponds to an optimal contract (y_0^*, k_0^*, y^*, k^*).

3.5 The Reformulated Saddle Problem

The saddle problem can be further simplified. As formulated, the controls y, k, and j are complicated functionals. The multipliers, however, can be regarded as type-dependent scalars and can thus be optimized point-by-point. Using integration by parts, we arrive at an alternative formulation of the problem that reverses the roles: in this alternative, controls become type-dependent scalars while multipliers are the more complicated functionals. The alternative formulation will prove more useful. Thus, using integration by parts (see Appendix C), the approach yields a reformulated Lagrangian function:

$$\begin{aligned}
&\mathcal{L}_R\left(y_0, k_0, y, k, j, y(\underline{\theta}), k(\underline{\theta}), y(\bar{\theta}), k(\bar{\theta}),\ \xi, \eta, \gamma, \lambda\right) \quad \equiv \quad U_0(y_0, k_0) \ + \\
&\beta \left\{ \int_{\underline{\theta}}^{\bar{\theta}} U\left(\theta + y(\theta), k(\theta)\right)\left(f(\theta) - \xi'(\theta)\right) d\theta \ + \ \lambda \left(B - \int_{\underline{\theta}}^{\bar{\theta}} \left(y(\theta) + qk(\theta)\right) f(\theta) d\theta \right) \right. \\
&- \int_{\underline{\theta}}^{\bar{\theta}} \xi(\theta) \frac{\partial U}{\partial c}\left(\theta + y(\theta), k\left(\theta\right)\right) d\theta + \xi(\bar{\theta}) U\left(\bar{\theta} + y(\bar{\theta}), k(\bar{\theta})\right) - \xi(\underline{\theta}) U\left(\underline{\theta} + y(\underline{\theta}), k(\underline{\theta})\right) \\
&\left. + \int_{\underline{\theta}}^{\bar{\theta}} \eta(\theta) \left(k(\theta) - k(\underline{\theta}) - \int_{\underline{\theta}}^{\theta} j(\tilde{\theta}) d\tilde{\theta} \right) d\theta + \int_{\underline{\theta}}^{\bar{\theta}} \gamma(\theta) j(\theta) d\theta \right\}.
\end{aligned} \tag{12}$$

The reformulation requires that we separately specify extra controls $y(\underline{\theta}), y(\bar{\theta}), k(\bar{\theta})$ at the boundaries $\underline{\theta}$ and $\bar{\theta}$. The Lagrangian function $\mathcal{L}_R$ admits a saddle point if it is concave in $(y(\theta), k(\theta), j(\theta), y(\underline{\theta}), k(\underline{\theta}), y(\bar{\theta}), k(\bar{\theta}))$ for all types θ.

Any contract that solves $\mathcal{L}_R$ also solves $\mathcal{L}$, and so the solution to the OC problem is unchanged if $\mathcal{L}_R$ replaces $\mathcal{L}$ in the planner's objective function. In the foregoing analysis, we focus attention on optimal contracts that solve this reformulated problem.

4 Characterizing Optimal Contracts

In this section, we present four main findings. First we show that all optimal contracts are collateralized and all will exhibit a distortion — a wedge between the marginal rate of substitution and the relative prices of the two goods in the second period for almost all types θ. Moreover, the distortion will always be in the direction of under-consumption of capital good k in the second period, and will be largest for types in the middle of the income distribution. We relate the distortion to the marginal propensity for forfeiture in the optimal contract. Given an extra dollar of income, higher distortions are associated with larger increases, or smaller decreases, in forfeiture of collateral. The formula can be used to show that forfeiture is locally decreasing in income at the very top and very bottom of the income distribution.

Second, the continuation contract will be strictly separating at the top of income distribution. For some open neighborhood of $\overline{\theta}$, k will be strictly increasing, and y strictly decreasing in θ. If pooling occurs, it is bounded away from the top of the distribution. This implies that, de facto, an OC will not function like a standard debt contract.

Third, we analyze the inter-temporal distortion in the OC between first and second periods. When capital and non-capital goods are complements, the borrower will generally over-consume or under-accumulate assets in the first period due the anticipated distortion in the repayment period. For low income borrowers in particular, this means that while too little of one's assets are accumulated, too much whatever's there is forfeited.

Fourth, in a parametric "log-uniform" model, the contract can be solved in closed form, and capital good consumption is convex in income. We show that if the support of the distribution is not too broad then the OC is strictly monotone — pooling does not occur. In this case, consumption, imputed collateral, and forfeiture all vary continuously. Full default is not a discrete event. When the support is broad enough, however, then pooling occurs at the bottom of the income distribution. Default occurs below some income threshold. In both cases, forfeiture is strictly decreasing in income.

4.1 Collateralization and Intra-temporal Distortion

The first order conditions of the saddle point problem associated with $\mathcal{L}_R$ are as follows. The first order conditions for initial period consumptions $y_0 \equiv c_0$ and k_0 are

$$\frac{\partial U_0}{\partial c_0} = \beta\lambda R \quad \text{and} \quad \frac{\partial U_0}{\partial k_0} = \beta\lambda(q_0 - \frac{q}{R}) \tag{13}$$

The first order condition in $y(\theta)$ is

$$\frac{\partial U}{\partial c}(\theta)\left(f(\theta)-\xi'(\theta)\right)-\lambda f(\theta)=\frac{\partial^2 U}{\partial c^2}(\theta)\,\xi(\theta) \tag{14}$$

The first order condition in $k(\theta)$ is

$$\frac{\partial U}{\partial k}(\theta)\left(f(\theta)-\xi'(\theta)\right)+\eta\left(\theta\right)-\lambda q f(\theta)=\frac{\partial^2 U}{\partial c \partial k}(\theta)\,\xi(\theta) \tag{15}$$

The first order conditions in controls $y(\underline{\theta})$, $y(\bar{\theta})$, and $k(\bar{\theta})$ are

$$\frac{\partial U}{\partial c}\left(\underline{\theta}\right)\xi(\underline{\theta})=\frac{\partial U}{\partial c}\left(\bar{\theta}\right)\xi(\bar{\theta})=\frac{\partial U}{\partial k}\left(\bar{\theta}\right)\xi(\bar{\theta})=0. \tag{16}$$

The FOC in $k(\underline{\theta})$ is

$$\int_{\underline{\theta}}^{\bar{\theta}}\eta\left(\tilde{\theta}\right)d\tilde{\theta}-\xi\left(\underline{\theta}\right)=0. \tag{17}$$

Finally, the first order condition in $j(\theta)$ is

$$-\int_{\theta}^{\bar{\theta}}\eta\left(\tilde{\theta}\right)d\tilde{\theta}+\gamma\left(\theta\right)=0. \tag{18}$$

The Koresh-Kuhn Tucker equations in the saddle problem yield a system of thirteen equations and thirteen unknowns. The first order conditions (13)-(18) on the controls yield nine equations (there are two equations in (13) and three in (16)). The constraints (3), (5), (8), and (9) yield the other four.

Combining first order conditions (14) and (15), one arrives at

$$\frac{\partial U/\partial k}{\partial U/\partial c}(\theta)\;=\;q\left[\frac{\frac{\partial^2 U}{\partial c\partial k}(\theta)\,\xi(\theta)/q\;-\;\eta(\theta)/q\;+\;\lambda f(\theta)}{\frac{\partial^2 U}{\partial c^2}(\theta)\,\xi(\theta)\;+\;\lambda f(\theta)}\right]\;\equiv\;q\Delta(\theta) \tag{19}$$

Equation (19) displays a distortion in second period consumption mix between collateralized and non-collateralized goods. The bracketed term "$[\cdot]$" is the distortion, denoted here by $\Delta(\theta)$. A distortion exists whenever $\Delta(\theta)$ differs from one. When $\Delta(\theta) > (<)1$, there is relative under-consumption (over-consumption) of the capital good k, and over-consumption (under-consumption) of the composite good c in the second (repayment) period.

Our first result asserts that every optimal contract is collateralized, and collateralization results from a distortion of the optimal contract in the interior of the support $[\underline{\theta}, \overline{\theta}]$ in which the second period capital good is under-consumed.

Proposition 1. *Let (y_0^*, k_0^*, y^*, k^*) be an optimal contract. Then:*

(i) A distortion exists at each income type $\theta \in (\underline{\theta}, \overline{\theta})$, and in all such distortions the borrower consumes too little of the capital good relative to that of the composite good in the second period, i.e.,

$$\frac{\partial U/\partial k}{\partial U/\partial c}(\theta) \ > \ q, \quad \text{and so} \quad \Delta^*(\theta) > 1.$$

(ii) No distortion exists for the highest income type $\overline{\theta}$, i.e., $\frac{\partial U/\partial k}{\partial U/\partial c}(\overline{\theta}) \ = \ q$, and if the contracts is strictly separating in a neighborhood of the lowest income type $\underline{\theta}$, then no distortion exists for that type as well.

(iii) The optimal contract is collateralized: $\overline{\Gamma}^* > 0$.

(iv) There exists $\epsilon > 0$ such that both y^* and k^* must be strictly monotone on the interval $(\overline{\theta} - \epsilon, \overline{\theta}]$.

The Proposition focuses on the continuation contract (y^*, k^*) and holds for any first period consumption y_0 and k_0 on or off-path. According to Part (i), distortions exist in $(\underline{\theta}, \overline{\theta})$, and will always be in the direction of relative under-consumption of k, and over-consumption of y. The fact that no distortion exists at the top (Part (ii)) mirrors well known results in optimal taxation models whereby marginal rates should be zero for top earners, and positive for others.

An immediate implication of (i) is that the the borrower's consumption expenditure provided by the lender, $y^*(\theta) + qk^*(\theta)$ is decreasing in θ. Hence, the optimal contract provides at least partial (but not full) insurance to the borrower.

The proofs of distortion results (i) and (ii) in the Appendix are detailed. The basic idea is to formulate a "relaxed" saddle problem in which the incentive constraint is one-sided. Higher types are deterred from mimicking lower types, but not vise-versa. This translates into an inequality constraint on the envelope condition, in which case the multiplier $\xi(\theta)$ is non-negative. With $\xi(\theta) \geq 0$, the saddle problem can be verified to be globally concave, and the distortion produces under-consumption of the capital good in the second period. We show that $\xi(\theta) > 0$ and the envelope constraint binds on the interior of the support.

Therefore the solution under one-sided IC also solves the general saddle problem with two-sided IC.[14]

Part (iii) shows that collateralization must be true in all optimal contracts under private information. The contract must allows for seizure of the capital good in some states to satisfy both IC and the need for at least partial insurance. Part (iv) asserts that the OC must be separating at the upper end of the realized income distribution. This means that the OC is not pooling at the top and therefore not a standard debt contract.

Collateralization (Part (iii)) is, of course, the central feature of the optimal contract. In order to show it, one must find at least two types θ' and θ'' for whom the values of seized collateral (as defined in (6)) are different. This is easy if the contract contains a pooling region since $\Gamma^*(\theta)$ is strictly decreasing in this region. If the contract is strictly separating everywhere, then the no-distortion result (Part (ii)) applies. In this case one can show $\Gamma^*(\underline{\theta}) > \Gamma^*(\overline{\theta})$. If $\Gamma^*(\underline{\theta}) \leq \Gamma^*(\overline{\theta})$, i.e. if the value of forfeiture of low income types was smaller than for high income types, then the no-distortion result would imply $U(c^*(\underline{\theta}), k^*(\underline{\theta})) \geq U(c^*(\overline{\theta}), k^*(\overline{\theta}))$ contradicting the strictly increasing payoffs in θ implied by the envelope constraint.

The last part of the argument can be summarized as: when the optimal contract is strictly separating, then $\Gamma^*(\underline{\theta}) > \Gamma^*(\overline{\theta})$, i.e., the value of the collateral seized from the lowest income type exceeds that of the highest income type.

Unfortunately, we have been unable to verify the inequality elsewhere in the distribution. We suspect that it is not generally true that the level of seized collateral is always strictly decreasing in realized income. However, we do prove it in the parametric model in Section 5, and in the result below, we show that it's true in local neighborhoods of the highest and lowest type.

Proposition 2. *In any optimal contract,*

(i) For any two types θ, θ' with $\theta > \theta'$, if $\Gamma^*(\theta) > \Gamma^*(\theta')$ then $\Delta^*(\theta') > \Delta^*(\theta)$.

(ii) For each type θ,

$$q\frac{dk^*}{d\theta}(\Delta(\theta) - 1) \; = \; 1 + \frac{d\Gamma^*(\theta)}{d\theta} \tag{20}$$

Part (i) of this proposition shows that if a higher income type forfeits more, the distortion associated with the lower type must be larger in order to deter mimicry by the higher

[14] Hellwig (2007) employs a similar proof strategy in a Mirleesian tax model.

income type. Parts (ii) has an appealing interpretation and yields a number of important implications.

One left-hand side of (20), $\frac{dk^*}{d\theta}$ is the marginal propensity of consumption (MPC) of the capital good. $q\Delta(\theta)$ may be regarded as the distorted price of k, with the distortion owing to asymmetric information. Clearly, q is the undistorted price of k. From Proposition 1, $q\Delta(\theta) > q$. Hence, the left hand side is

$$\text{Distorted value of MPC} \quad - \quad \text{Undistorted value of MPC}$$

The right-hand side of (20) represents the value of the loss, due to forfeiture, when an additional "dollar" of income is realized by the borrower. If $\frac{d\Gamma^*(\theta)}{d\theta} < 0$ then an extra dollar of income yields less than a dollar consumed after forfeiture to the bank is factored in. If $\frac{d\Gamma^*(\theta)}{d\theta} > 0$ then an extra dollar yields more than a dollar consumed after forfeiture to the bank. Putting everything together, the difference in the value of the MPC of capital due to distortion equals the incremental forfeiture seized by the lender.

It is easy to verify that $\frac{d\Gamma^*(\theta)}{d\theta} \geq -1$. Combining this result with Proposition 1, it follows that

$$\frac{d\Gamma^*(\overline{\theta})}{d\theta} = -1 \quad \text{and} \quad \frac{d^2\Gamma^*(\overline{\theta})}{d\theta^2} < 0,$$

and

$$\frac{d\Gamma^*(\underline{\theta})}{d\theta} = -1 \quad \text{and} \quad \frac{d^2\Gamma^*(\underline{\theta})}{d\theta^2} \geq 0$$

with strict inequality if the contract is separating in a neighborhood of $\underline{\theta}$.

This means that forfeiture is decreasing locally around $\underline{\theta}$ and $\overline{\theta}$. In these portions of the income distribution, seized collateral is locally regressive in income. Moreover, because forfeiture must average out to zero over the entire distribution, it must be negative somewhere yielding a windfall to certain income types. The implicit subsidy to these types is required in order to satisfy IC constraints.

The second derivatives address what happens locally to the marginal forfeiture of an extra dollar of income. For low income levels very close to $\underline{\theta}$, the marginal forfeiture of an extra dollar of income is weakly increasing (strictly so if there is no pooling at the bottom). Incrementally larger portions of extra income go to the bank. For high income levels close to $\overline{\theta}$, the marginal forfeiture of an extra dollar of income is positive but strictly decreasing. Incrementally smaller portions of extra income go to the lender.

4.2 Collateralization and Inter-temporal Distortion

From the first order equations (13), one can verify that *intra*-temporal consumption in the initial period $t = 0$ is undistorted: $\frac{\partial U_0/\partial k_0}{\partial U_0/\partial y_0} = q_0 - \frac{q}{R}$. The marginal rate of substitution in date 0 equals the user cost of capital. However, an *inter*-temporal distortion comes from the informational distortion (19) in the second period. Combining first order conditions in c_0 and k_0 with those in the continuation problem we obtain,

$$\frac{1}{R\beta}\int \frac{1}{\partial U/\partial c}(\theta)dF(\theta) \;=\; \frac{1}{\partial U_0/\partial c_0}\left(1-\int \frac{\partial^2 U/\partial c^2}{\partial U/\partial c}(\theta)\xi(\theta)d\theta\right) \tag{21}$$

$$\frac{1}{\beta(q_0-\frac{q}{R})}\int \frac{1}{\partial U/\partial c}(\theta)dF(\theta) \;=\; \frac{1}{\partial U_0/\partial k_0}\left(1-\int \frac{\partial^2 U/\partial c^2}{\partial U/\partial c}(\theta)\xi(\theta)d\theta\right) \tag{22}$$

$$\frac{1}{R\beta}\int \frac{1}{\partial U/\partial k}(\theta)dF(\theta) \;=\; \frac{1}{\partial U_0/\partial c_0}\left(1-\int \frac{\partial^2 U/\partial c\partial k}{\partial U/\partial k)}(\theta)\xi(\theta)d\theta\right) \tag{23}$$

$$\frac{q}{\beta(q_0-\frac{q}{R})}\int \frac{1}{\partial U/\partial k}(\theta)dF(\theta) \;=\; \frac{1}{\partial U_0/\partial k_0}\left(1-\int \frac{\partial^2 U/\partial c\partial k}{\partial U/\partial k}(\theta)\xi(\theta)d\theta\right) \tag{24}$$

When U is additively separable between c and k, one obtains $\frac{\partial^2 U}{\partial c \partial k} = 0$ in which case (23) and (24) reduce to

$$\frac{1}{R\beta}\int \frac{1}{\partial U/\partial k}(\theta)dF(\theta) \;=\; \frac{1}{\partial U_0/\partial c_0} \tag{25}$$

and

$$\frac{1}{\beta(q_0-\frac{q}{R})}\int \frac{1}{\partial U/\partial k}(\theta)dF(\theta) \;=\; \frac{1}{\partial U_0/\partial k_0} \tag{26}$$

which are recognizable as inverse Euler equations, common in dynamic mechanism design.[15]

[15]See, for instance, Mirleesian models of capital income taxation such as Farhi and Werning (2012) and references therein. Note that the capital good in our model plays a role similar to the consumption good in the Mirlees model. In either case, the unobserved shock does not directly apply to it. This is the reason that inverse Euler equations hold in both models when U is separable. By contrast, the non-capital consumption in our model is substantively closer to the labor/leisure choice in the Mirlees models since the shock enters consumption of those goods directly.

The application of Jensen's Inequality to (25) and (26) implies $\partial U_0/\partial c_0 < R\beta \int \frac{\partial U}{\partial k}(\theta)dF(\theta)$ and $\partial U_0/\partial k_0 < \beta \frac{q_0 - \frac{q}{R}}{q} \int \frac{\partial U}{\partial k}(\theta)dF(\theta)$. In other words, compared to the full information optimum, the planner puts too much weight on current consumption k_0 and y_0 relative to future consumption of capital. This is primarily due to the imperfect insurance from second period incentive constraints.

Here, because U may not be separable in our model, an additional distortion arises due to income effects of a change in the distorted relative price $q\Delta(\theta)$ for type θ when initial consumption is varied. Taking, as an example, the inter-temporal trade off between k_0 and k, we rewrite (24) to obtain

$$\frac{\partial U_0}{\partial k_0} \;=\; \beta(\frac{q_0 - q/R}{q}) \int \frac{\partial U}{\partial k}(\theta)dF(\theta) \;-\; d_{1,k} \;-\; d_{2,k} \tag{27}$$

where

$$d_{1,k} \;=\; \beta(\frac{q_0 - q/R}{q}) \left[\int \frac{\partial U}{\partial k}(\theta)dF(\theta) - \left(\int \frac{1}{\partial U/\partial k}(\theta)dF(\theta) \right)^{-1} \right] \;>\; 0$$

and

$$d_{2,k} \;=\; \beta(\frac{q_0 - q/R}{q}) \left(\int \frac{1}{\partial U/\partial k}(\theta)dF(\theta) \right)^{-1} \left(\int \frac{\partial^2 U/\partial c\partial k}{\partial U/\partial k}(\theta)\xi(\theta)d\theta \right) \gtrless 0$$

Equation (27) displays the double distortion explicitly. Here, $d_{1,k}$ represents the distortion mentioned above. The planner commits less savings to a period in which consumption is volatile. The second distortion $d_{2,k}$ comes from the income effect off the relative price distortion $q\Delta(\theta)$. $d_{2,k} > 0$ if c and k are complements and $d_{2,k} < 0$ if c and k are substitutes as Assumption 1 allows. The distortion $d_{2,k}$ is zero only if there are no income effects (i.e., U is separable in c and k). A similar expression appears in the trade off between c_0 and k. Since $d_{2,k} > 0$ would be expected when, say, k is housing, the "double distortion" leads to a more severe under-consumption of the second period's capital good than would be the case in the standard incentive contract with separable utility.

The trade offs represented by (21) and (22) also exhibit a double distortion, even with separable U. Unlike the future capital good, the two distortions may have different signs. Rewriting (22) for instance, one obtains

$$\frac{\partial U_0}{\partial k_0} \;=\; \beta(q_0 - \frac{q}{R}) \int \frac{\partial U}{\partial c}(\theta) dF(\theta) \;-\; d_{1,c} \;-\; d_{2,c} \tag{28}$$

where

$$d_{1,c} \;=\; \beta(q_0 - q/R)\left[\int \frac{\partial U}{\partial c}(\theta)dF(\theta) - \left(\int \frac{1}{\partial U/\partial c}(\theta)dF(\theta)\right)^{-1}\right] \;>\; 0$$

and

$$d_{2,c} \;=\; \beta(q_0 - q/R)\left(\int \frac{1}{\partial U/\partial c}(\theta)dF(\theta)\right)^{-1}\left(\int \frac{\partial^2 U/\partial c^2}{\partial U/\partial c}(\theta)\xi(\theta)d\theta\right) \;<\; 0.$$

The distortion $d_{1,c}$ pushes the borrower in the direction of over-consumption in the first period just as before. However, a second distortion $d_{2,c}$ pushes toward *under*-consumption in the first period. The intuition is that the planner wish to increase savings in order to improve insurance options. This is done indirectly since increased savings relaxes incentive constraints. The degree to which this can work is determined by weighted average value of risk aversion (see last term). The larger the average risk aversion, the greater is the effect of relaxing IC on insurance. Which distortion $d_{1,c}$ or $d_{2,c}$ dominates is unknown. The net effect is therefore ambiguous. To summarize,

Proposition 3. *Any optimal contract* (y_0^*, k_0^*, y^*, k^*) *exhibits an inter-temporal distortion summarized by the Euler equations (27) and (28). In (27), the distortions* $d_{1,k}$ *and* $d_{2,k}$ *are positive and nonnegative, respectively, when* c *and* k *are weak complements. In that case both distortions lead to over-consumption of initial capital* k_0 *relative to future capital consumption* k*. In (28), the distortions have opposite signs:* $d_{1,c} > 0$ *and* $d_{2,c} < 0$*. The net effect of these distortions lead to over (under)-consumption of initial* k_0 *relative to future non-capital consumption* c *if* $|d_{1,c}| > (<)|d_{2,c}|$.

5 Closed Form Solutions in a Log-Uniform Model

A parametrization of the model nicely illustrates what form collateralized contract takes and how the distortion varies across income types.

Consider a *log-uniform* contracting problem in which $f(\theta) = \bar{f} = \frac{1}{\bar{\theta}-\underline{\theta}}$ and $U = \frac{1}{2}\log c + \frac{1}{2}\log k$. The log-uniform case simplifies the problem by assuming no complementarities between the two goods. By allowing for curvature of payoffs in the two goods, however, it provides a more nuanced contact than under the standard quasi-linear setup.

Once again, we abstract from the first period and focus attention on the continuation OC as a function of the multiplier λ, itself determined by first period consumption.

The result below characterizes a closed form solution to the OC that, depending on parameter values, is either strictly separating, or pools types only at the lower end of the distribution. In the latter case, full default or full forfeiture of the capital good is is a well defined concept. In either case, we show that that consumption of the capital good is convex in the income type, so that the individual's consumption of capital good k relative to his realized income θ is increasing.

Given the log utility function, the first order conditions, (15) and (14) become

$$\frac{1}{2k(\theta)}\left(\bar{f}-\xi'(\theta)\right) \;+\; \eta(\theta) \;-\; \lambda q\bar{f} \;=\; 0.$$

and

$$\frac{1}{2c(\theta)}\left(\bar{f}-\xi'(\theta)\right)-\lambda\bar{f} \;=\; -\frac{1}{c(\theta)^2}\xi(\theta)$$

We first verify a conjecture that a separating equilibrium exists at some set of parameter values. Under full separation, the monotonicity constraint does not bind. The multipliers associated with the monotonicity constant therefore vanish: $\gamma(\theta)\equiv 0$ and $\eta(\theta)\equiv 0$. Using this restriction, we combined the two first order conditions, and then differentiate with respect to θ to obtain,

$$2\lambda q\left(k'(\theta)c(\theta)+k(\theta)c'(\theta)\right)-2\lambda\frac{d}{d\theta}\left\{\left(c(\theta)\right)^2\right\}=2\lambda qk(\theta)-1. \tag{29}$$

Now, the incentive constraint implies

$$\frac{1}{2c(\theta)}\left(c'(\theta)-1\right)+\frac{1}{2k(\theta)}k'(\theta)=0. \tag{30}$$

Combining this restriction with (29) and (30), we obtain $2\lambda\frac{d}{d\theta}\left\{\left(c(\theta)\right)^2\right\}=1$, implying

$$c(\theta) \;=\; \sqrt{\left(c(\underline{\theta})\right)^2+\frac{\theta-\underline{\theta}}{2\lambda}}. \tag{31}$$

Consequently,

$$y(\theta) \;=\; \sqrt{(c(\underline{\theta}))^2 + \frac{\theta - \underline{\theta}}{2\lambda}} \;-\; \theta \tag{32}$$

Notice that y is weakly decreasing if $c(\underline{\theta}) \geq \frac{1}{4\lambda}$.

Integrating (30) and substituting into (31), we obtain

$$\log(k(\theta)) \;=\; \log(k(\underline{\theta})) + 4\lambda c(\theta) - 4\lambda c(\underline{\theta}) - \frac{1}{2}\log\left(c(\theta)\right) + \log\left(c(\underline{\theta})\right). \tag{33}$$

We now establish conditions under which (y, k) is a separating OCC, i.e., y and k are strictly monotone. Given our earlier results, if k is strictly increasing, there is no distortion at either boundary $\underline{\theta}$ or $\overline{\theta}$. We therefore have

$$c(\underline{\theta}) = qk(\underline{\theta}) \quad \text{and} \quad c(\overline{\theta}) = qk(\overline{\theta}) \tag{34}$$

These no-distortion-at-the-boundary conditions are combined with (32) to produce

$$k(\theta) \;=\; A\frac{\exp(4\lambda c(\theta))}{c(\theta)} \tag{35}$$

where A, a constant, is equal to $q(c(\underline{\theta}))^2 \exp(-4\lambda c(\underline{\theta}))$.

It is not difficult to show that $k(\theta)$ is strictly globally convex in θ. In other words, *the ratio of capital good consumption to realized income is increasing.*

Notice that k and y are ultimately functions of $c(\underline{\theta})$. One can solve for $c(\underline{\theta})$ by evaluating (33) at $\theta = \overline{\theta}$ and using the no-distortion-at-the-boundary conditions in (34) to obtain

$$\log\left(c(\underline{\theta})^2 + \frac{\overline{\theta} - \underline{\theta}}{2\lambda}\right) \;=\; \log(c(\underline{\theta})^2) \;+ 4\lambda\left((c(\underline{\theta}))^2 + \frac{\overline{\theta} - \underline{\theta}}{2\lambda}\right)^{1/2} - 4\lambda c(\underline{\theta}). \tag{36}$$

The result below establishes parametric restrictions under which this construction is valid; k and y are, in fact, strictly separating. It also asserts that when these parametric restrictions are not satisfied, the optimal collateralized contract has a pooling region at the bottom of the income distribution, and is strictly separating everywhere else.

Proposition 4. *Let δ be defined as the unique scalar that satisfying* $\log(\delta) = \delta^{1/2} - 1$. *($\delta \approx$*

12.34*). Then:*

1. *If* $8\lambda(\overline{\theta}-\underline{\theta}) \leq \delta - 1$, *then there exists an optimal continuation contract* (y^*, k^*) *that is strictly separating and satisfies (31), (32), (35), and (36). Notably,*
   (a) k^* *is increasing strictly convex in* θ, y^* *is decreasing and strictly concave, and*
   (b) $\Gamma^*(\theta)$, *the collateral seized from income type* θ, *is strictly decreasing in* θ.
2. *If* $8\lambda(\overline{\theta}-\underline{\theta}) > \delta - 1$, *then there exists a cutoff type* $\theta^* \in (\underline{\theta}, \overline{\theta})$ *and an optimal continuation contract* (y^*, k^*) *such that* (y, k) *is strictly separating in the interval* $[\theta^*, \overline{\theta}]$ *and is pooling in the interval* $[\underline{\theta}, \theta^*)$. *Notably,*
   (a) *above the cutoff* θ^*, k^* *is strictly convex in* θ, y^* *is strictly concave,*
   (b) *above the cutoff* θ^*, *the collateral seized is strictly decreasing in income* θ.

We refer to OCs that satisfy Part (2) of the Proposition as *semi-pooling*. The Proposition identifies a partition of the parameter set into those that generate separating OCs, and those that generate semi-pooling ones. Both the support and the multiplier λ play a critical role. A larger support increases the incentive to misreport in a separating equilibrium since the consumption of the high income types cannot be increased too much without violating the resource constraint. As a result, incentives can be only be brought into line by a pooling contract at the lower end. In that case there is no further gain from mimicry below the threshold θ^*.

An increase in λ also diminishes the range in which the separating contract exists. In the Proposition, the multiplier λ the resource constraint is treated as a parameter. However, λ is an implicit solution to the resource constraint evaluated at y and k, given budget B. As a shadow price of the constraint, λ is decreasing in B. Consequently, in Part (1) an increase in B, the date 0 purchase k_0 of collateralized capital good, results in an increase in consumption of the capital good k. An increase in k_0 is, in effect, an increase in the collateral requirement. Since the capital good has consumption value, an increase in its required purchase at date $t = 0$ increases its consumption in $t = 1$.

Interestingly enough, the increased collateral requirement has an ambiguous effect on c. Consumption at the bottom, $c(\underline{\theta})$, decreases even though c is a normal good. Consumption at the top may or may not increase. This entirely due to information incentives. A uniform increase in the loan amount y would magnify incentives to misreport. To discourage high types from mimicry, the loan to poorer types must be reduced.

In the Appendix we extend the analysis above to show that the separating part of the semi-pooling contract in Part 2 satisfies

$$y(\theta) \;=\; \sqrt{\left(c(\theta^*)\right)^2+\frac{\theta-\underline{\theta}}{2\lambda}}\;-\;\theta \tag{37}$$

and

$$k(\theta) \;=\; A^*\frac{\exp(4\lambda c(\theta))}{c(\theta)^{1/2}} \tag{38}$$

where A^*, in this case, is equal to $k(\theta^*)c(\theta^*)\exp(-4\lambda c(\theta^*))$.

Notice that (37) and (38) are analogues of (32) and (35). They have the same functional form with the pooling threshold θ^* replacing the lower bound $\underline{\theta}$ of the support in the functions. The pooling consumptions $k(\theta^*)$ and $y(\theta^*)$ can be found by evaluating (37) and (38) at $\theta = \overline{\theta}$ and using the no-distortion condition $qc(\overline{\theta}) = k(\overline{\theta})$ to define three equations with three unknowns $c(\overline{\theta})$, $k(\overline{\theta})$, and θ^* (see Appendix for details).

In the semi-pooling contract, k consumption is globally convex in θ; it is constant below the threshold θ^* and strictly convex above the θ^*. In this sense, the semi-pooling contract more closely resembles the familiar debt contract in home mortgages.

The continuation optimal collateralized contract is illustrated in two parametric special cases. We set parameter values

$$q=1 \qquad \lambda=1 \qquad f'/f=1$$

The allocations, distortions, and forfeitures are then displayed in Figure 1 for two different supports. Each column displays a continuation OC, a distribution of the distortionary wedge, and the distribution of the value of forfeiture/seizure of the asset across the realized incomes.

In the first column of Figure 1 the support is $[\underline{\theta},\overline{\theta}] = [-.5,.5]$. The OC in this case is strictly separating, as per Part 1 of Proposition 4.

In the second column $[\underline{\theta},\overline{\theta}] = [-1.5,1.5]$ so that the second support is twice as broad as the first. In this case, the OC is semi-pooling, as per Part 2 of the Proposition. Poorer income below a cutoff of around -1.3 are pooled, while types above that threshold strictly separate.

In both cases, the distortion is a non-monotonic function of income. The wedge is

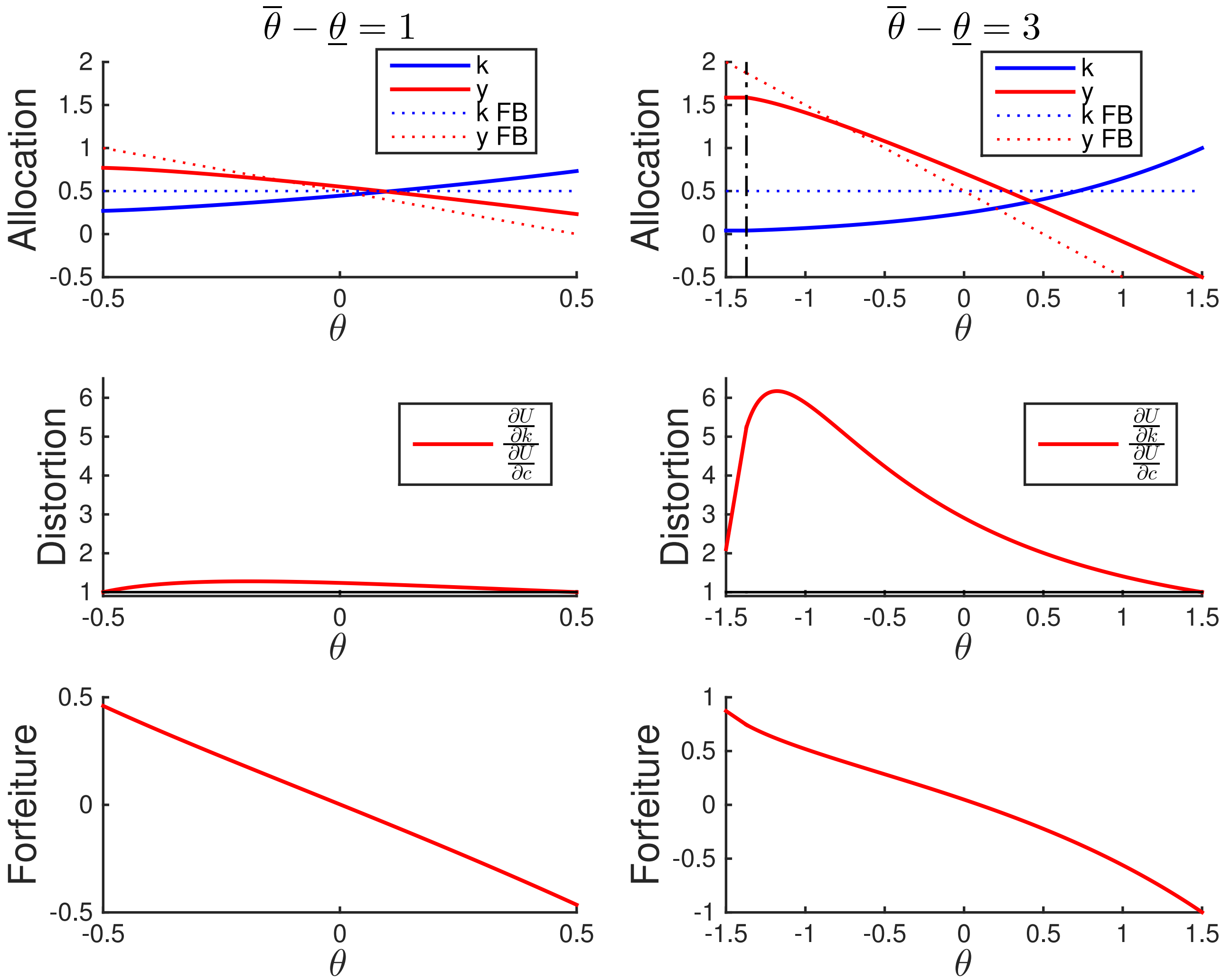

Figure 1: Allocations, Distortions, and Forfeiture

largest for middle income types, and smallest for the very poor and very rich. This is largely due to the differing information incentives and the cost mitigating bad incentives across types.

From the perspective of the full information optimum, the very poor have no incentive to lie. Hence, there is no reason to distort their consumption. The very rich, by contrast have the greatest incentive to lie. To align incentives of this group, one might suppose that a large distortion is required. Unfortunately, it is more costly to distort the consumption of the rich than the poor, since they take up a larger share of the resource constraint. Hence, the largest distortion is assigned to the middle income types — the types who are most likely to mimicked by high income types, a priori. The distribution of forfeiture illustrates, however, that the brunt of incentive costs is borne by low income types.

6 Computing Optimal Contracts

The previous section saw that an explicit closed form solution for the optimal contract exists in the log uniform case. Here we give an algorithm for computing one in the general case following the same guidelines.

6.1 An Algorithm

Given strict concavity of U, Lemma 2 in the Appendix is used to prove there exists a one-to-one function $H : \mathbb{R}^2_+ \to \mathbb{R}^2$ satisfying

$$\begin{bmatrix} c \\ \\ k \end{bmatrix} = H\left(\frac{\partial U(c,k)}{\partial c}, \frac{\partial U(c,k)}{\partial k}\right) = \begin{bmatrix} H^c\left(\frac{\partial U(c,k)}{\partial c}, \frac{\partial U(c,k)}{\partial k}\right) \\ \\ H^k\left(\frac{\partial U(c,k)}{\partial c}, \frac{\partial U(c,k)}{\partial k}\right) \end{bmatrix}$$

The function H associates marginal utilities of each of the two goods to those values of the consumption that generated them. Using the change of variables: $x = \frac{\partial U(c,k)}{\partial c}$ and $z = \frac{\partial U(c,k)}{\partial k}$, we obtain

$$
\begin{bmatrix} \frac{\partial^2 U(c,k)}{\partial c^2} & \frac{\partial^2 U(c,k)}{\partial c \partial k} \\ \\ \frac{\partial^2 U(c,k)}{\partial c \partial k} & \frac{\partial^2 U(c,k)}{\partial k^2} \end{bmatrix} = \begin{bmatrix} \frac{\partial H^c(x,z)}{\partial x} & \frac{\partial H^c(x,z)}{\partial z} \\ \\ \frac{\partial H^k(x,z)}{\partial x} & \frac{\partial H^k(x,z)}{\partial z} \end{bmatrix}^{-1}
$$

$$
= \frac{1}{\frac{\partial H^c(x,z)}{\partial x}\frac{\partial H^k(x,z)}{\partial z} - \frac{\partial H^c(x,z)}{\partial z}\frac{\partial H^k(x,z)}{\partial x}} \begin{bmatrix} \frac{\partial H^k(x,z)}{\partial z} & -\frac{\partial H^c(x,z)}{\partial z} \\ \\ -\frac{\partial H^k(x,z)}{\partial x} & \frac{\partial H^c(x,z)}{\partial x} \end{bmatrix}.
$$

$$
\equiv \begin{bmatrix} J_z^k(x,z) & -J_z^c(x,z) \\ \\ -J_x^k(x,z) & J_x^c(x,z) \end{bmatrix}.
$$

Assuming that the monotonicity constraint is not binding locally, the first order conditions in y and k, (14) and (15), can now be expressed as

$$
x(\theta)(f(\theta) - \xi'(\theta)) - \lambda f(\theta) \;=\; J_z^k(x,z)\xi(\theta), \tag{39}
$$

and

$$
z(\theta)\left(f(\theta) - \xi'(\theta)\right) - \lambda q f(\theta) \;=\; -J_z^c(x,z)\xi(\theta). \tag{40}
$$

From these two equations, we can solve for $\xi(\theta)$ as

$$
\xi(\theta) = \lambda f(\theta) K\left(x(\theta), z(\theta)\right), \tag{41}
$$

where

$$
K(x,y) = \frac{qx - z}{zJ_z^k(x,z) + xJ_z^c(x,z)}.
$$

Plugging in this solution of $\xi(\theta)$ back to (39), and divide both side by $f(\theta)$, we obtain

$$x(\theta)\left(1-\lambda\frac{f'(\theta)}{f(\theta)}K\left(x(\theta),z(\theta)\right)-\lambda\left(\frac{\partial K}{\partial x}x'(\theta)+\frac{\partial K}{\partial z}z'(\theta)\right)\right)-\lambda$$
$$=J_z^k(x(\theta),z(\theta))\lambda K\left(x(\theta),z(\theta)\right). \tag{42}$$

From the incentive constraint, we obtain another equation

$$\frac{\partial U}{\partial c}\left(\frac{dc}{d\theta}-1\right)+\frac{\partial U}{\partial k}\frac{dk}{d\theta}=0,$$

or

$$x(\theta)\left(\frac{\partial H^c}{\partial x}x'(\theta)+\frac{\partial H^c}{\partial z}z'(\theta)-1\right)+z(\theta)\left(\frac{\partial H^k}{\partial x}x'(\theta)+\frac{\partial H^k}{\partial z}z'(\theta)\right)=0. \tag{43}$$

(42) and (43) form a system of differential equations in $x(\theta)$ and $z(\theta)$. Under the full separating assumption, Proposition 1 shows that there is no distortion at the top and the bottom. Therefore,

$$\frac{\frac{\partial U}{\partial k}\left(\underline{\theta}\right)}{\frac{\partial U}{\partial c}\left(\underline{\theta}\right)}=q=\frac{\frac{\partial U}{\partial k}\left(\overline{\theta}\right)}{\frac{\partial U}{\partial c}\left(\overline{\theta}\right)}.$$

Equivalently,

$$\frac{z\left(\underline{\theta}\right)}{x\left(\underline{\theta}\right)}=q=\frac{z\left(\overline{\theta}\right)}{x\left(\overline{\theta}\right)}. \tag{44}$$

We can use standard boundary value problem, e.g. MATLAB's BVP functions, to solve numerically for $x(\theta), z(\theta)$ as a solution to the ODE system (42)-(43) and the boundary condition (44).

6.2 The Phase Diagram

When the density f belong to the exponential family, i.e. satisfies $\frac{f'}{f}=const$ and equivalently,

$$f(\theta)=\phi\exp(-\psi\theta), \tag{45}$$

for some constants ϕ and ψ, the ODE system (42)-(43) becomes autonomous in x and z, i.e independent of θ. We can then use the phase diagram to analyze this system for characterizing optimal contracts.

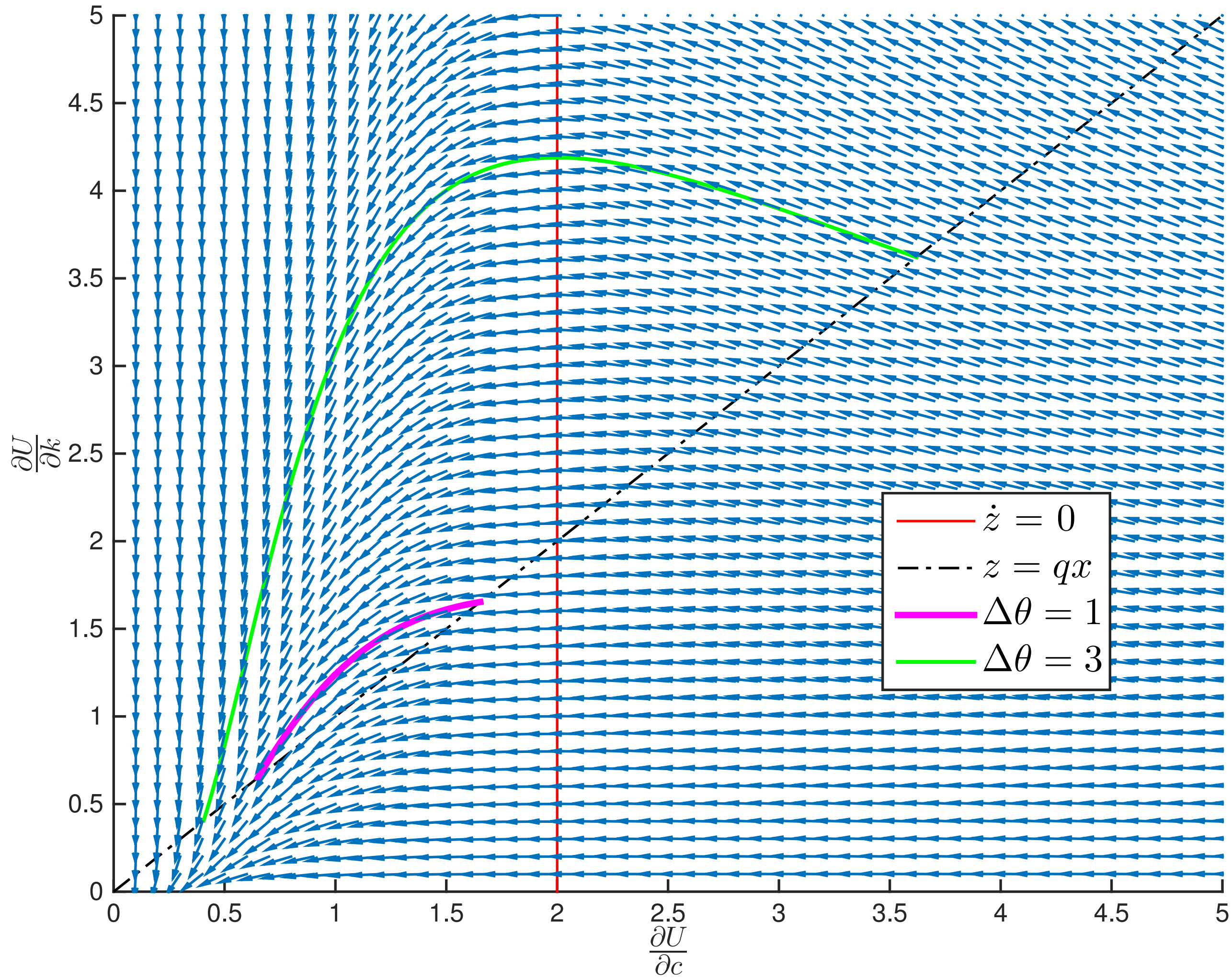

Figure 2: Phase Diagram and Two Solutions

The log-uniform model fits in this case. The phase diagram for the solutions in Figure 1 are shown in Figures 2 and 3.

For more general utility functions, we explore further the phase diagram in the numerical analysis below.

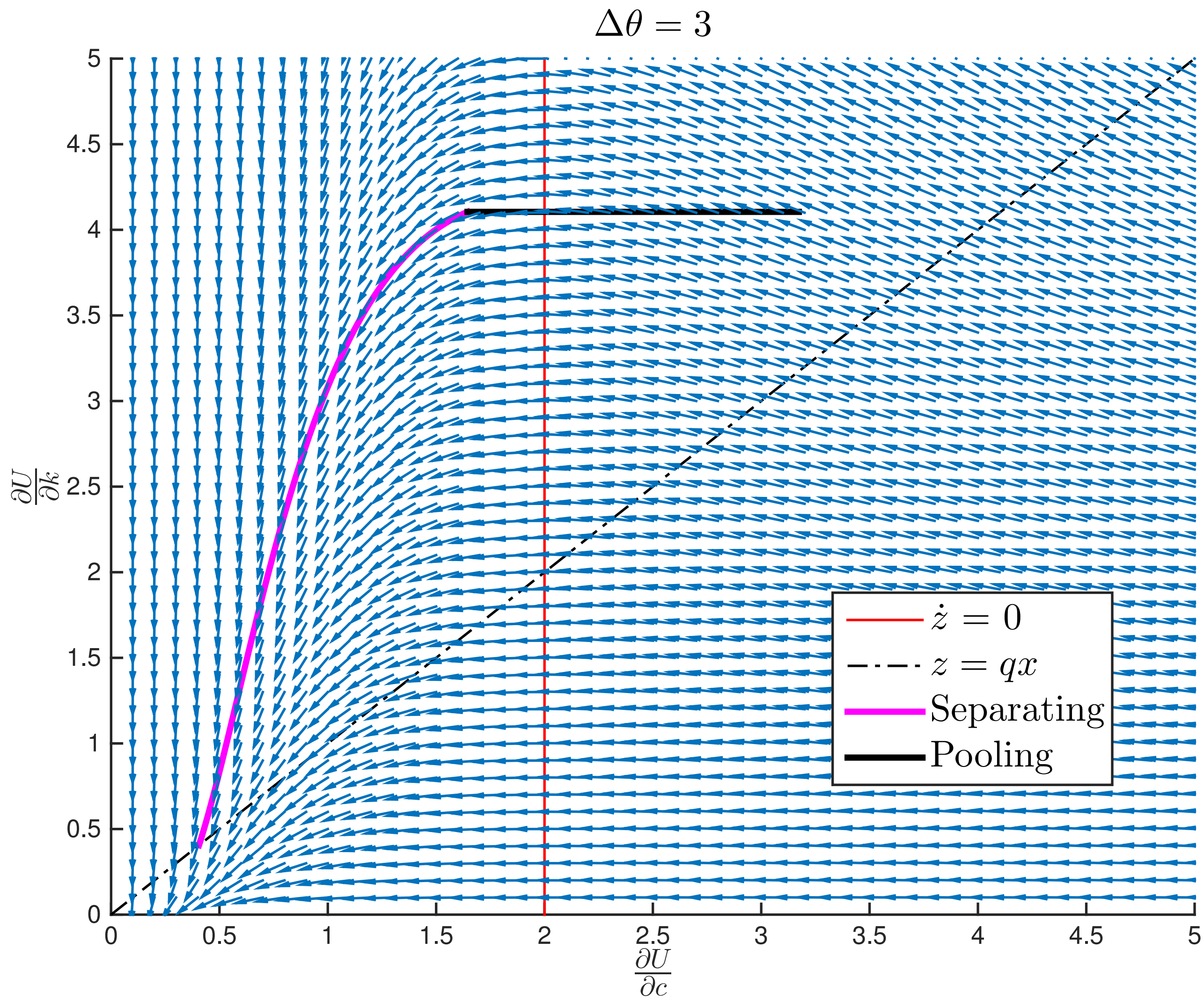

Figure 3: Phase Diagram for Semi-Pooling Contract

6.3 Robustness

This section makes use of the algorithm developed in the previous section to examine the properties of optimal contracts under several specifications of the utility function and type distribution.

We focus on the case of exponential type distribution, (45), to take advantage of the phase diagram in analyzing the ODE system (42)-(43) . We consider the general specification of the utility function with constant elasticity of substitution (CES) between non-housing and housing consumption:

$$U(c,k) = \frac{\left(\alpha c^{\frac{\epsilon-1}{\epsilon}} + (1-\alpha)k^{\frac{\epsilon-1}{\epsilon}}\right)^{\frac{\epsilon(1-\sigma)}{\epsilon-1}} - 1}{1-\sigma}, \tag{46}$$

where ϵ is the CES coefficient and σ is the constant relative risk-aversion coefficient. When $\epsilon \to 1$ and $\sigma \to 1$,

$$U(c,k) = \alpha \log(c) + (1-\alpha)\log(k). \tag{47}$$

Example 1. *In this example, we consider the case with utility function (47) with $\alpha = \frac{1}{2}$ as in the log-uniform case considered in Proposition 4. However, $\phi = 1$ instead of $\phi = 0$ as in the proposition. Figures 5 and 6 depict the phase diagram, allocations, distortion, and forfeiture for this case.*

Example 2. *In this example, we consider the case with utility function (47) with more weight on non-housing consumption, $\alpha = 0.8$, as estimated in Gorea and Midrigan (2016). We keep the uniform distribution of types. Figures 7 and 8 depict the phase diagram, allocations, distortion, and forfeiture for this case.*

Example 3. *In this example, we consider the case with CES function (46) with $\alpha = 0.5$, and $\epsilon = 0.5$, $\sigma = 5$ as used in Piazzesi et al. (2007). We keep the uniform distribution of types. Figures 9 and 10 depict the phase diagram, allocations, distortion, and forfeiture for this case.*

One property stands out from Examples 1-3 is that housing consumption $k(\theta)$ is convex in θ. This property is shown analytically in Proposition 4. When utility function is separable and has CRRA in c, k, i.e. when $\epsilon = \frac{1}{\sigma}$ in (46), we show a similar result. We call an contract exhibiting *end-point convexity* if $\frac{\partial k}{\partial \theta}(\underline{\theta}) < \frac{\partial k}{\partial \theta}(\overline{\theta})$.

Proposition 5. *In the class of collateralized contracting problems with CRRA payoffs $U = \alpha \frac{c^{1-\sigma}}{1-\sigma} + (1-\alpha)\frac{k^{1-\sigma}}{1-\sigma}$, any continuation OC exhibits end-point convexity.*

The result can be extended without much trouble to CRRA payoffs with different exponential weights σ_1 and σ_2. Without further restriction on payoffs and distributions, little can be said about the behavior of k^* in the middle of the distribution.

The convexity of housing consumption is, moreover, consistent with mortgage contracts in housing. Data from the Consumer Expenditure Surveys shows that among households with mortgages, housing consumption is roughly convex in income, and this is not generally true of households without mortgages.

7 General Equilibrium

As argue in the previous sections, in general, because of asymmetric information problems, housing good is relative under-consumed. Then, one would suspect that the general equilibrium version of this model in which housing price is endogenously determined will have housing price too low relative to the first best level. Numerical simulations for the log-log uniform model in Subsection 5 confirms this intuition.

To simplify the analysis we focus on the second period. We assume that the consumers are endowed with one unit of housing in addition to the stochastic endowment of composite goods. The banking sector is perfectly competitive. The banks offer optimal incentive compatible contracts to the consumers, in exchange for the consumers' deposit of wealth before income shocks are realized, to maximize profit subject to an endogenous outside options of the consumers. Both consumers and banks have access to a market for housing at unit price q.

In particular, each bank solves

$$\Pi(q, \underline{U}) = \max_{\{y,k\}} \left\{ q \cdot 1 - \int_{\underline{\theta}}^{\overline{\theta}} (y(\theta) + qk(\theta))\, dF(\theta) \right\}$$

subject to

$$\int_{\underline{\theta}}^{\overline{\theta}} U(\theta + y(\theta), k(\theta)) dF(\theta) \geq \underline{U},$$

and

$$U(\theta + y(\theta), k(\theta)) \geq U(\theta + y(\theta'), k(\theta')),$$

for all $\theta, \theta' \in [\underline{\theta}, \overline{\theta}]$.

We define a competitive equilibrium as following.

Definition. A competitive equilibrium is a price q, an endogenous outside option $\underline{U}$ and an optimal contract $\{y^*(\theta), k^*(\theta)\}_{\theta \in [\underline{\theta}, \overline{\theta}]}$ such that at this optimal contract, banks make zero profit

$$\Pi(q, \underline{U}) = 0$$

and the housing and composite good markets clear

$$\int_{\underline{\theta}}^{\overline{\theta}} k^*(\theta) dF(\theta) = 1,$$

and

$$\int_{\underline{\theta}}^{\overline{\theta}} y^*(\theta) dF(\theta) = 0.$$

It is easy to show that the outside option and the equilibrium optimal contracts solves

$$\underline{U} = \max_{\{y,k\}} \int_{\underline{\theta}}^{\overline{\theta}} U(\theta + y(\theta), k(\theta)) dF(\theta)$$

subject to

$$q = \int_{\underline{\theta}}^{\overline{\theta}} \left(y(\theta) + qk(\theta) \right) dF(\theta)$$

and

$$U(\theta + y(\theta), k(\theta)) \geq U(\theta + y(\theta'), k(\theta')).$$

We compare this competitive equilibrium to the one under perfect information (first-best competitive equilibrium) in which the contracts are not subject to the incentive constraints.

Example 4. *Consider the case with log-log utility function and uniform distribution of composite goods endowment as in Subsection 5. In the first best competitive equilibrium*

$$\theta + y^{FB}(\theta) = q^{FB} k^{FB}(\theta)$$

for all θ. Integrating this equality from $\underline{\theta}$ to $\overline{\theta}$ and using the market clearing conditions, we arrive at

$$q^{FB} = \frac{\underline{\theta} + \overline{\theta}}{2}.$$

However, under asymmetric information, by the distortion result,

$$\theta + y^*(\theta) > q^* k^*(\theta)$$

for all $\theta \in (\underline{\theta}, \overline{\theta})$*. After integrating this equality from* $\underline{\theta}$ *to* $\overline{\theta}$ *and using the market clearing conditions, we obtain*

$$q^* < \frac{\underline{\theta} + \overline{\theta}}{2} = q^{FB}.$$

Figure 4 shows the competitive equilibrium price for housing, compared to the first-best competitive equilibrium price, when we vary $\underline{\theta}$ *while keeping* $\overline{\theta} - \underline{\theta} = 2$*. We observe that the competitive equilibrium price always lie below the first-best equilibrium price and the difference between the two prices is larger at lower levels of* $\underline{\theta}$*. This result might speak to the low housing price during the last financial crisis 2007-2009 and the subsequent recession.*

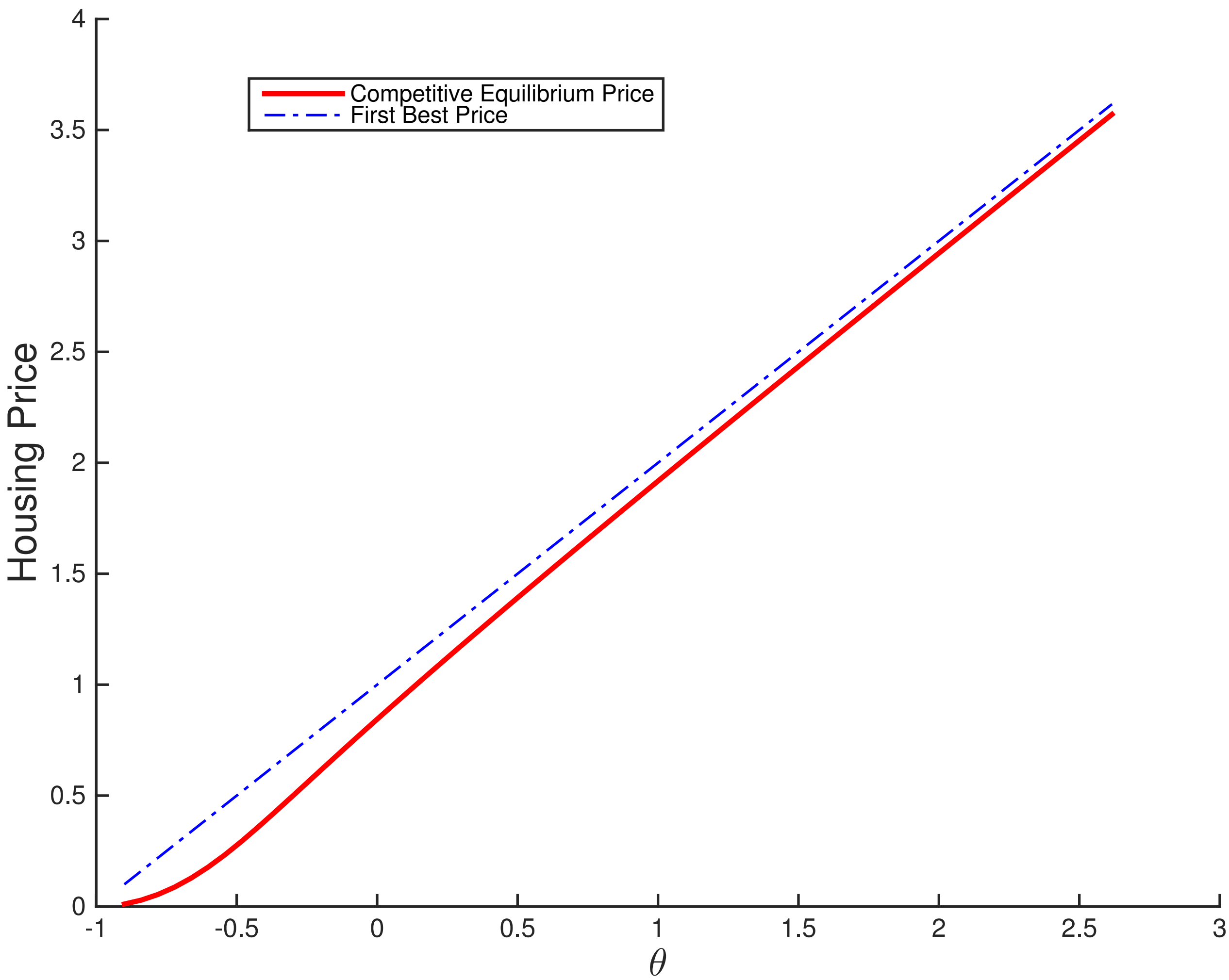

Figure 4: General Equilibrium

References

Barro, R. J. (1976). The loan market, collateral, and rates of interest. *Journal of Money, Credit and Banking*, 8(4):439–456.

Besanko, D. and Thakor, A. V. (1987). Collateral and rationing: Sorting equilibria in monopolistic and competitive credit markets. *International Economic Review*, 28(3):671–689.

Bester, H. (1985). Screening vs. rationing in credit markets with imperfect information. *The American Economic Review*, 75(4):850–855.

Campbell, J. and Cocco, J. (2015). A model of mortgage default. *Journal of Finance*, 70(4):1495–1554.

Chan, Y.-S. and Kanatas, G. (1985). Asymmetric valuations and the role of collateral in loan agreements. *Journal of Money, Credit and Banking*, 17(1):84–95.

Diamond, P. A. (1998). Optimal income taxation: An example with a u-shaped pattern of optimal marginal tax rates. *The American Economic Review*, 88(1):83–95.

Dixit, A. and Seade, J. (1979). Utilitarian versus egalitarian redistributions. *Economics Letters*, 4(2):121–124.

Eisfeldt, A. L. and Rampini, A. A. (2009). Leasing, ability to repossess, and debt capacity. *Review of Financial Studies*, 22(4):1621–1657.

Farhi, E. and Werning, I. (2012). Capital taxation: Quantitative explorations of the inverse euler equation. *Journal of Political Economy*, 120(3):398–445.

Gorea, D. and Midrigan, V. (2016). Liquidity constraints in the u.s. housing market. NYU Working Paper.

Green, E. J. (1987). *Lending and the Smoothing of Uninsurable Income*, pages 3–25. University of Minnesota Press, ned - new edition edition.

Hellwig, M. F. (2007). A contribution to the theory of optimal utilitarian income taxation. *Journal of Public Economics*, 91(7):1449 – 1477.

Lacker, J. M. (2001). Collateralized debt as the optimal contract. *Review of Economic Dynamics*, 4(4):842 – 859.

Manove, M., Padilla, A. J., and Pagano, M. (2001). Collateral versus project screening: A model of lazy banks. *The RAND Journal of Economics*, 32(4):726–744.

Milgrom, P. and Shannon, C. (1994). Monotone comparative statics. *Econometrica*, 62(1):157–180.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *The Review of Economic Studies*, 38(2):175–208.

Mirrlees, J. A. (1976). Optimal tax theory. *Journal of Public Economics*, 6(4):327 – 358.

Parlatore, C. (2016). Collateralizing liquidity. NYU Stern School of Business Working Paper.

Piazzesi, M., Schneider, M., and Tuzel, S. (2007). Housing, consumption and asset pricing. *Journal of Financial Economics*, 83(3):531 – 569.

Rampini, A. A. (2005). Default and aggregate income. *Journal of Economic Theory*, 122(2):225 – 253.

Rampini, A. A. and Viswanathan, S. (2013). Collateral and capital structure. *Journal of Financial Economics*, 109(2):466 – 492.

Stiglitz, J. E. and Weiss, A. (1981). Credit rationing in markets with imperfect information. *The American Economic Review*, 71(3):393–410.

Wette, H. C. (1983). Collateral in credit rationing in markets with imperfect information: Note. *The American Economic Review*, 73(3):442–445.

Appendix

A Preliminaries: Implications of Assumption 1

For the proofs in this section, we re-write condition (1) as

$$\frac{\partial U^2}{\partial c^2}\frac{\partial U}{\partial k}-\frac{\partial^2 U}{\partial c\partial k}\frac{\partial U}{\partial c}<0 \tag{48}$$

and

$$\frac{\partial U^2}{\partial k^2}\frac{\partial U}{\partial c}-\frac{\partial^2 U}{\partial c\partial k}\frac{\partial U}{\partial k}<0 \tag{49}$$

for all c, k.

The first condition (48) is equivalent to the Strict Single Crossing Condition (SSCC) for

$$\hat{U}\left(y,k,\theta\right)\equiv U(\theta+y,k).$$

Indeed,

$$\frac{\partial}{\partial\theta}\left(\frac{\frac{\partial\hat{U}}{\partial k}}{\frac{\partial\hat{U}}{\partial y}}\right)=\frac{\frac{\partial^2 U}{\partial k\partial c}\frac{\partial U}{\partial c}-\frac{\partial U^2}{\partial c^2}\frac{\partial U}{\partial k}}{\left(\frac{\partial U}{\partial c}\right)^2}>0. \tag{50}$$

The following lemma is a generalization of Dixit and Seade (1979) and is important for the change of variables used in Subsection 6.1.

Lemma 2. *If* $\left(c_0,k_0\right),\left(c_1,k_1\right)$ *satisfy*

$$\frac{\partial U(c_0,k_0)}{\partial c}=\frac{\partial U(c_1,k_1)}{\partial c}$$
$$\frac{\partial U(c_0,k_0)}{\partial k}=\frac{\partial U(c_1,k_1)}{\partial k},$$

then

$$\left(c_0,k_0\right)=\left(c_1,k_1\right).$$

Proof. If $(c_0, k_0) \neq (c_1, k_1)$, by strict concavity of U:

$$U(c_0, k_0) - U(c_1, k_1) < \frac{\partial U(c_1, k_1)}{\partial c}(c_0 - c_1) + \frac{\partial U(c_1, k_1)}{\partial c}(k_0 - k_1).$$

Similarly

$$\begin{aligned} U(c_1, k_1) - U(c_0, k_0) <& \frac{\partial U(c_0, k_0)}{\partial c}(c_1 - c_0) + \frac{\partial U(c_0, k_0)}{\partial c}(k_1 - k_0) \\ =& \frac{\partial U(c_1, k_1)}{\partial c}(c_1 - c_0) + \frac{\partial U(c_1, k_1)}{\partial c}(k_1 - k_0). \end{aligned}$$

Therefore

$$U(c_0, k_0) - U(c_1, k_1) > \frac{\partial U(c_1, k_1)}{\partial c}(c_0 - c_1) + \frac{\partial U(c_1, k_1)}{\partial c}(k_0 - k_1),$$

which contradicts the earlier inequality. Thus $(c_0, k_0) = (c_1, k_1)$. □

Lemma 3. *Conditions* (48) *and* (49) *are satisfied for all* c, k *if and only if given any vector of prices* $p > 0$ *an increase in income increases both* c *and* k *consumption, i.e.* c *and* k *are normal goods.*

Proof. First we show the sufficient condition, i.e. if c, k are normal goods then (48) and (49) are satisfied for all c, k. Indeed, given c, k, consider the following income and prices: $m = \frac{\partial U(c,k)}{\partial c}c + \frac{\partial U(c,k)}{\partial k}k$ and $p = \begin{bmatrix} \frac{\partial U(c,k)}{\partial c} \\ \frac{\partial U(c,k)}{\partial k} \end{bmatrix} > 0$. It is easy to verify that (c, k) is the optimal consumption at m and p.

Let $D(m, p)$ denote the demand function associated with the utility function U. The standard algebra derivations for D implies

$$\frac{\partial D}{\partial m} = \frac{\begin{bmatrix} \frac{\partial^2 U(c,k)}{\partial c^2} & \frac{\partial^2 U(c,k)}{\partial c \partial k} \\ \frac{\partial^2 U(c,k)}{\partial c \partial k} & \frac{\partial^2 U(c,k)}{\partial k^2} \end{bmatrix}^{-1} p}{p' \begin{bmatrix} \frac{\partial^2 U(c,k)}{\partial c^2} & \frac{\partial^2 U(c,k)}{\partial c \partial k} \\ \frac{\partial^2 U(c,k)}{\partial c \partial k} & \frac{\partial^2 U(c,k)}{\partial k^2} \end{bmatrix}^{-1} p} > 0,$$

because both housing and consumption are normal goods. Since $p = \begin{bmatrix} \frac{\partial U(c,k)}{\partial c} \\ \frac{\partial U(c,k)}{\partial k} \end{bmatrix}$ and $p' \begin{bmatrix} \frac{\partial^2 U(c,k)}{\partial c^2} & \frac{\partial^2 U(c,k)}{\partial c \partial k} \\ \frac{\partial^2 U(c,k)}{\partial c \partial k} & \frac{\partial^2 U(c,k)}{\partial k^2} \end{bmatrix}^{-1} p < 0$, (from the concavity of U), the last inequality yields

$$\begin{bmatrix} \frac{\partial^2 U(c,k)}{\partial c^2} & \frac{\partial^2 U(c,k)}{\partial c \partial k} \\ \frac{\partial^2 U(c,k)}{\partial c \partial k} & \frac{\partial^2 U(c,k)}{\partial k^2} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial U(c,k)}{\partial c} \\ \frac{\partial U(c,k)}{\partial k} \end{bmatrix} < 0.$$

By direct algebra, the last inequality is equivalent to (48) and (49), given that $\frac{\partial^2 U(c,k)}{\partial c^2} \frac{\partial^2 U(c,k)}{\partial k^2} - \left(\frac{\partial^2 U(c,k)}{\partial c \partial k} \right)^2 > 0$ due to the concavity of U.

The necessary condition is proved in exactly the same way. □

B Proof of Lemma 1

We divide the proof of Lemma 1 in two parts. In the first part, Lemma 4, we show the necessary condition, i.e. any incentive compatible contract must satisfy the monotonicity and the envelope conditions. In the second part, Lemma 5, we show that the two properties guarantee incentive compatibility.

Lemma 4. *Given any incentive compatible contract, for $\theta > \theta'$, we have $y(\theta) \leq y(\theta')$ and $k(\theta) \geq k(\theta')$. In addition, the envelope condition (5) is satisfied for all θ.*

Proof. From the IC constraint for type θ:

$$U\left(\theta + y(\theta), k(\theta)\right) \geq U\left(\theta + y(\theta'), k(\theta')\right),$$

we rule out the possibility that $y(\theta') \geq y(\theta)$ and $k(\theta') \geq k(\theta)$, with at least on strictly inequality. Similarly from the IC constraint for type θ', we rule out the possibility that

$y(\theta') \le y(\theta)$ and $k(\theta') \le k(\theta)$, with at least one strict inequality. Therefore to obtain this lemma, we just need to eliminate the possibility that $y(\theta) \ge y(\theta')$ and $k(\theta) \le k(\theta')$ with at least one strict inequality.

We show this by contradiction. Suppose that it is true. Let

$$\tilde{U}(\tilde{y}, k, \theta) \equiv U(\theta - \tilde{y}, k).$$

Then $\tilde{U}$ satisfies the Strict Single Crossing Condition (SSCC) because it satisfies the Spence-Mirlees condition (see Milgrom and Shannon (1994) for the exact definition of these conditions):

$$\frac{\partial}{\partial\theta}\left(\frac{\partial\tilde{U}/\partial\tilde{y}}{\partial\tilde{U}/\partial k}\right) = \frac{\partial}{\partial\theta}\left(-\frac{\partial U(\theta - \tilde{y}, k)/\partial c}{\partial U(\theta - \tilde{y}, k)/\partial k}\right) > 0,$$

where the last inequality is equivalent to (48). By Milgrom and Shannon (1994, Theorem 3), since

$$(-y(\theta), k(\theta)) < \left(-y(\theta'), k(\theta')\right),$$

and

$$\tilde{U}\left(-y(\theta'), k(\theta'), \theta'\right) \ge \tilde{U}\left(-y(\theta), k(\theta), \theta'\right),$$

we have

$$\tilde{U}\left(-y(\theta'), k(\theta'), \theta\right) > \tilde{U}\left(-y(\theta), k(\theta), \theta\right),$$

or equivalently

$$U\left(\theta + y(\theta'), k(\theta')\right) > U\left(\theta + y(\theta), k(\theta)\right),$$

which contradicts the IC constraint for θ. Therefore by contradiction, we obtain the monotonicity property.

Now we turn to the envelope conditions. For $\theta' < \theta$, we write the IC constraint for type θ as

$$U\left(\theta + y(\theta), k(\theta)\right) - U\left(\theta' + y(\theta'), k(\theta')\right) \ge U\left(\theta + y(\theta'), k(\theta')\right) - U\left(\theta' + y(\theta'), k(\theta')\right).$$

Dividing both-side by $\theta - \theta'$ and take the limit $\theta' \to \theta$, we obtain

$$\frac{dU^-}{d\theta}(\theta + y(\theta), k(\theta)) \ge \frac{\partial U}{\partial c}\left(\theta + y(\theta), k\left(\theta\right)\right).$$

The second half of the envelope condition is obtained similarly by considering the IC constraint for type θ'.

$\square$

Lemma 5. *Any piecewise differentiable allocation that satisfies the Monotonicity and Envelope Conditions satisfies the incentive constraint.*

Proof. We show that if the first part of (5) holds, then the global downward IC property holds, i.e. for $\theta' < \theta$

$$U(\theta + y(\theta), k(\theta)) \geq U(\theta + y(\theta'), k(\theta')).$$

Indeed, the inequality is equivalent to

$$\int_{\theta'}^{\theta} \frac{d^-}{d\tilde{\theta}} \left(U(\theta + y(\tilde{\theta}), k(\tilde{\theta}))\right) d\tilde{\theta} \geq 0. \tag{51}$$

We use the differential form of the (left) Envelope Condition:

$$\frac{\partial U(\tilde{\theta} + y(\tilde{\theta}), k(\tilde{\theta}))}{\partial c} \frac{dy^-}{d\tilde{\theta}} + \frac{\partial U(\tilde{\theta} + y(\tilde{\theta}), k(\tilde{\theta}))}{\partial k} \frac{dk^-}{d\tilde{\theta}} \geq 0$$

to estimate the intergrand above as:

$$\begin{aligned}\frac{d^-}{d\tilde{\theta}} \left(U(\theta + y(\tilde{\theta}), k(\tilde{\theta}))\right) &= \left(\frac{\partial U(\theta + y(\tilde{\theta}), k(\tilde{\theta}))}{\partial c} \frac{dy^-}{d\tilde{\theta}} + \frac{\partial U(\theta + y(\tilde{\theta}), k(\tilde{\theta}))}{\partial k} \frac{dk^-}{d\tilde{\theta}}\right) \\ &\geq \left(\frac{\partial U(\theta + y(\tilde{\theta}), k(\tilde{\theta}))}{\partial c} \left(-\frac{\frac{\partial U(\tilde{\theta}+y(\tilde{\theta}),k(\tilde{\theta}))}{\partial k}}{\frac{\partial U(\tilde{\theta}+y(\tilde{\theta}),k(\tilde{\theta}))}{\partial c}}\right) + \frac{\partial U(\theta + y(\tilde{\theta}), k(\tilde{\theta}))}{\partial k}\right) \frac{dk^-}{d\tilde{\theta}}.\end{aligned}$$

Because of the SSCC, (50),

$$\frac{\frac{\partial U(\theta+y(\tilde{\theta}),k(\tilde{\theta}))}{\partial k}}{\frac{\partial U(\theta+y(\tilde{\theta}),k(\tilde{\theta}))}{\partial c}} > \frac{\frac{\partial U(\tilde{\theta}+y(\tilde{\theta}),k(\tilde{\theta}))}{\partial k}}{\frac{\partial U(\tilde{\theta}+y(\tilde{\theta}),k(\tilde{\theta}))}{\partial c}},$$

and because of the monotonicity property $\frac{dk^-}{d\tilde{\theta}} \geq 0$. Therefore $\frac{d^-}{d\tilde{\theta}} \left(U(\theta + y(\tilde{\theta}), k(\tilde{\theta}))\right)$ is positive, thus the integral (51) is positive. So the global downward IC property holds.

Similarly, we can show that if the second part of (5) holds then the global upward IC property holds.

$\square$

C Integration-by-Parts for the Reformulated Saddle Problem

The integration-by-parts arguments needed to derive the reformulated saddle problem are given by

$$\int_{\underline{\theta}}^{\bar{\theta}} \xi(\theta) \frac{dU}{d\theta} \left(\theta + y(\theta), k\left(\theta\right)\right) d\theta \ =$$

$$\xi(\bar{\theta}) U\left(\overline{\theta} + y(\overline{\theta}), k\left(\overline{\theta}\right)\right) \ - \ \xi(\underline{\theta}) U\left(\underline{\theta} + y(\underline{\theta}), k\left(\underline{\theta}\right)\right) \ - \ \int_{\underline{\theta}}^{\bar{\theta}} U\left(\theta + y(\theta), k\left(\theta\right)\right) \xi'(\theta) d\theta$$

and

$$\int_{\underline{\theta}}^{\bar{\theta}} \eta(\theta) \left(k(\theta) - k(\underline{\theta}) - \int_{\underline{\theta}}^{\theta} j\left(\tilde{\theta}\right) d\tilde{\theta} \right) d\theta =$$

$$\int_{\underline{\theta}}^{\bar{\theta}} \eta(\theta) \left(\, k(\theta) - k(\underline{\theta}) \right) d\theta \ - \ \int_{\underline{\theta}}^{\bar{\theta}} j(\theta) \left(\int_{\theta}^{\bar{\theta}} \eta\left(\tilde{\theta}\right) d\tilde{\theta} \right) d\theta.$$

D Proof of Proposition 1

To prove the results in Proposition 1, we will first need to show that the multiplier $\xi(\theta)$ is positive for all θ. To do so we employ the strategy in Hellwig (2007). In particular, we consider a ”weakly relaxed problem” in which only downward incentive compatibility and monotonicity are imposed. In this case, the multiplier $\xi(\theta)$ on the local downward IC constraint is positive by assumption. Then, we show that the optimal solution to this ”weakly relaxed problem” also satisfies the global incentive constraint, there for it is also the optimal solution to the original problem.

Let

$$v(\theta) = U(\theta + y(\theta), k(\theta)).$$

The local IC constraint can be written as (for simplicity we assume differentiability)

$$v'(\theta) - \frac{\partial U}{\partial c}\left(\theta + y(\theta), k\left(\theta\right)\right) \geq 0. \tag{52}$$

We also require that $k(.)$ is non-decreasing in θ, i.e. (8) and (9) hold. As shown in the proof of Lemma 5, the local downward IC constraint (52) and monotonicity constraint imply the global downward IC constraint.

We write the Lagrangian of this "weakly relaxed problem" as

$$\begin{aligned}\mathcal{L} = & \int_{\underline{\theta}}^{\bar{\theta}} U\left(\theta + y(\theta), k(\theta)\right) dF(\theta) \\ & + \lambda \left(B - \int_{\underline{\theta}}^{\bar{\theta}} \left(y(\theta) + qk(\theta)\right) dF(\theta) \right) \\ & + \int_{\underline{\theta}}^{\bar{\theta}} \xi(\theta) \left(v'(\theta) - \frac{\partial U}{\partial c}\left(\theta + y(\theta), k\left(\theta\right)\right) \right) d\theta \\ & + \int_{\underline{\theta}}^{\bar{\theta}} \eta(\theta) \left(k(\theta) - k(\underline{\theta}) - \int_{\underline{\theta}}^{\theta} j(\tilde{\theta}) d\tilde{\theta} \right) d\theta + \int_{\underline{\theta}}^{\bar{\theta}} \gamma(\theta) j(\theta) d\theta.\end{aligned}$$

Because of the constraints (52) and (9) , $\xi \geq 0$ and $\gamma \geq 0$.

We use the integral by parts

$$\int_{\underline{\theta}}^{\bar{\theta}} \xi(\theta) v'(\theta) d\theta = \xi(\bar{\theta}) v(\bar{\theta}) - \xi(\underline{\theta}) v(\underline{\theta}) - \int_{\underline{\theta}}^{\bar{\theta}} v(\theta) \xi'(\theta) d\theta$$

and let $\mu(\theta) = -\xi'(\theta)$ to write $\mathcal{L}$ as

$$\begin{aligned}\mathcal{L} = & \int_{\underline{\theta}}^{\bar{\theta}} U\left(\theta + y(\theta), k(\theta)\right) \left(f(\theta) + \mu(\theta)\right) d\theta \\ & + \lambda \left(B - \int_{\underline{\theta}}^{\bar{\theta}} \left(y(\theta) + qk(\theta)\right) f(\theta) d\theta \right) \\ & - \int_{\underline{\theta}}^{\bar{\theta}} \xi(\theta) \frac{\partial U}{\partial c}\left(\theta + y(\theta), k\left(\theta\right)\right) d\theta + \xi(\bar{\theta}) v(\bar{\theta}) - \xi(\underline{\theta}) v(\underline{\theta}) \\ & + \int_{\underline{\theta}}^{\bar{\theta}} \eta(\theta) \left(k(\theta) - k(\underline{\theta}) - \int_{\underline{\theta}}^{\theta} j(\tilde{\theta}) d\tilde{\theta} \right) d\theta + \int_{\underline{\theta}}^{\bar{\theta}} \gamma(\theta) j(\theta) d\theta.\end{aligned}$$

We use shortcuts $\frac{\partial U}{\partial c}(\theta)$,$\frac{\partial^2 U}{\partial c^2}(\theta)$, and $\mu(\theta) = -\xi'(\theta)$ to write the F.O.Cs as follow.

F.O.C. in $y(\theta)$

$$\frac{\partial U}{\partial c}(\theta)\left(\mu(\theta) + f(\theta)\right) - \lambda f(\theta) = \frac{\partial^2 U}{\partial c^2}(\theta)\,\xi(\theta). \tag{53}$$

F.O.C. in $k(\theta)$

$$\frac{\partial U}{\partial k}(\theta)\left(\mu(\theta) + f(\theta)\right) + \eta(\theta) - \lambda q f(\theta) = \frac{\partial^2 U}{\partial c \partial k}(\theta)\,\xi(\theta). \tag{54}$$

F.O.C. in $j(\theta)$

$$\gamma(\theta) - \int_{\theta}^{\bar{\theta}} \eta\left(\tilde{\theta}\right) d\tilde{\theta} = 0. \tag{55}$$

F.O.C. in $k(\bar{\theta})$

$$U_k\left(\bar{\theta}\right) \xi(\bar{\theta}) = 0. \tag{56}$$

F.O.C. in $y(\underline{\theta})$:

$$U_c\left(\underline{\theta}\right) \xi(\underline{\theta}) = 0. \tag{57}$$

F.O.C. in $k(\underline{\theta})$:

$$-U_k(\underline{\theta})\xi(\underline{\theta}) + \int_{\underline{\theta}}^{\bar{\theta}} \eta(\theta) d\theta = 0. \tag{58}$$

Because $U_k(\bar{\theta}) > 0$, (56) implies that $\xi(\bar{\theta}) = 0$. Similarly, because $U_c(\underline{\theta}) > 0$, (57) implies$\xi(\underline{\theta}) = 0$.

Combining this result with (58) yields

$$\int_{\underline{\theta}}^{\bar{\theta}} \eta(\theta) d\theta = 0. \tag{59}$$

Therefore, from (18),

$$\gamma(\underline{\theta}) = 0. \tag{60}$$

Also from (18),

$$\gamma(\bar{\theta}) = 0. \tag{61}$$

Armed with these properties, Lemma 8 below show that in the optimal solution of the "weakly relaxed problem," the local incentive constraint is satisfied. Lemma 5 then shows that the global incentive constraint is satisfied. Lemma 8 uses the following two Lemmas.

Lemma 6. *In an optimal solution to the "weakly relaxed problem", if $k(\theta)$ is constant over some interval $[\theta_1, \theta_2] \in [\underline{\theta}, \overline{\theta}]$ then $y(\theta)$ is constant over the same interval.*

Proof. Assume that $k(\theta) = k^*$ over $[\theta_1, \theta_2] \in [\underline{\theta}, \overline{\theta}]$. By downward incentive compatibility, $y(\theta)$ is non-decreasing over $[\theta_1, \theta_2]$.

We show the result in this lemma by contradiction. Assume that $y(\theta)$ is not constant over the same interval because y is continuous, there exists a non-degenerate subinterval $[\theta', \theta''] \in [\theta_1, \theta_2]$ such that $y(.)$ is strictly increasing over this interval.

In this interval

$$
\begin{aligned}
v'(\theta) &= \frac{d}{d\theta}\left(U\left(\theta + y(\theta), k^*\right)\right) \\
&= \frac{\partial U}{\partial c}\left(\theta + y(\theta), k^*\right) + \frac{\partial U}{\partial c}\left(\theta + y(\theta), k^*\right)\frac{dy}{d\theta} \\
&> \frac{\partial U}{\partial c}\left(\theta + y(\theta), k^*\right).
\end{aligned}
$$

Therefore (52) does not bind, i.e. $\xi(\theta) = 0$ for $\theta \in [\theta', \theta'']$. Since $\mu = -\xi'$, $\mu(\theta) = 0$ for $\theta \in [\theta', \theta'']$. (53) then implies

$$\frac{\partial U}{\partial c}(\theta + y(\theta), k^*) = \lambda$$

for $\theta \in [\theta', \theta'']$. Differentiate both sides with respect to θ, we have

$$\frac{\partial^2 U}{\partial c^2}(\theta + y(\theta), k^*)\left(1 + \frac{dy}{d\theta}\right) = 0.$$

This is a contradiction since $\frac{\partial^2 U}{\partial c^2} < 0$ and $\frac{dy}{d\theta} > 0$. □

Lemma 7. *In the optimal solution to the "weakly relaxed problem," for each $\theta^* \in (\underline{\theta}, \overline{\theta})$, if $\xi(\theta^*) = 0$ then $\gamma(\theta^*) > 0$.*

Proof. We show this result by contradiction. Assume that $\gamma(\theta^*) = 0$.

We rewrite (53) as

$$\xi'(\theta) + \frac{\frac{\partial^2 U}{\partial c^2}(\theta)}{\frac{\partial U}{\partial c}(\theta)} \xi(\theta) = f(\theta) \left(1 - \frac{\lambda}{\frac{\partial U}{\partial c}(\theta)}\right)$$

where $\frac{\partial U}{\partial c}(\theta)$, $\frac{\partial^2 U}{\partial c^2}(\theta) \frac{\partial^2 U}{\partial c \partial k}(\theta)$ are shortcuts. Using the fact that $\xi(\underline{\theta}) = 0$, we obtain

$$\xi(\theta) = \exp\left(-\int_{\underline{\theta}}^{\theta} \frac{\frac{\partial^2 U}{\partial c^2}(\tilde{\theta})}{\frac{\partial U}{\partial c}(\tilde{\theta})} d\tilde{\theta}\right) g_1(\theta),$$

where

$$g_1(\theta) = \int_{\underline{\theta}}^{\theta} \exp\left(\int_{\underline{\theta}}^{\theta_1} \frac{\frac{\partial^2 U}{\partial c^2}(\tilde{\theta})}{\frac{\partial U}{\partial c}(\tilde{\theta})} d\tilde{\theta}\right) \left(1 - \frac{\lambda}{\frac{\partial U}{\partial c}(\theta_1)}\right) dF(\theta_1).$$

Because $g \geq 0$, $g_1(\theta) \geq 0$ for all $\theta \in (\underline{\theta}, \overline{\theta})$. In addition $g_1(\theta^*) = 0$. Therefore θ^* is a local minimum of g_1. Therefore, $g_1'(\theta^*) = 0$ and $g_1''(\theta^*) \geq 0$. By the definition of g_1, this is equivalent to,

$$1 - \frac{\lambda}{\frac{\partial U}{\partial c}(\theta^*)} = 0$$

and

$$\frac{d}{d\theta}\left(1 - \frac{\lambda q}{\frac{\partial U}{\partial c}(\theta)}\right) \geq 0$$

and $\theta = \theta^*$. Equivalently, at $\theta = \theta^*$

$$x = \frac{d}{d\theta}\left\{\frac{\partial U}{\partial c}(\theta)\right\} \geq 0. \tag{62}$$

Similarly, we rewrite (54) as

$$\xi'(\theta) + \frac{\frac{\partial^2 U}{\partial c \partial k}(\theta)}{\frac{\partial U}{\partial k}(\theta)} \xi(\theta) = f(\theta) \left(1 - \frac{\lambda q}{\frac{\partial U}{\partial k}(\theta)} + \frac{\eta(\theta)}{\frac{\partial U}{\partial k}(\theta) f(\theta)}\right).$$

Again, since $\xi(\underline{\theta}) = 0$, we have

$$\xi(\theta) = \exp\left(-\int_{\underline{\theta}}^{\theta} \frac{\frac{\partial^2 U}{\partial c \partial k}(\tilde{\theta})}{\frac{\partial U}{\partial k}(\tilde{\theta})} d\tilde{\theta}\right) g_2(\theta),$$

where

$$g_2(\theta) = \int_{\underline{\theta}}^{\theta} \exp\left(\int_{\underline{\theta}}^{\theta_1} \frac{\frac{\partial^2 U}{\partial c^2}(\tilde{\theta})}{\frac{\partial U}{\partial c}(\tilde{\theta})} d\tilde{\theta}\right)\left(1 - \frac{\lambda q}{\frac{\partial U}{\partial k}(\theta_1)} + \frac{\eta(\theta_1)}{\frac{\partial U}{\partial k}(\theta_1) f(\theta_1)}\right) dF(\theta_1).$$

Because $g \geq 0$, $g_2 \geq 0$. In addition, $g_2(\theta^*) = 0$. Therefore θ^* is a local minimum of g_2. Thus, $g_2'(\theta^*) = 0$ and $g_2''(\theta^*) \geq 0$. By the definition of g_2, this is equivalent to

$$1 - \frac{\lambda q}{\frac{\partial U}{\partial k}(\theta^*)} + \frac{\eta(\theta^*)}{\frac{\partial U}{\partial k}(\theta^*) f(\theta^*)} = 0 \tag{63}$$

and at $\theta = \theta^*$:

$$\frac{d}{d\theta}\left(1 - \frac{\lambda q}{\frac{\partial U}{\partial k}(\theta)} + \frac{\eta(\theta)}{\frac{\partial U}{\partial k}(\theta) f(\theta)}\right) \geq 0. \tag{64}$$

Notice also that $\gamma \geq 0$ and $\gamma'(\theta) = -\eta(\theta)$ and $\gamma''(\theta) = -\eta'(\theta)$. In addition $\gamma(\theta^*) = 0$. Therefore θ^* is a local minimum of γ. Thus $\gamma'(\theta^*) = -\eta(\theta^*) = 0$ and $\gamma''(\theta^*) = -\eta'(\theta^*) \geq 0$. Plugging the first equality into (63) implies

$$1 - \frac{\lambda q}{\frac{\partial U}{\partial k}(\theta^*)} = 0.$$

Plugging the second inequality into (64) implies

$$\begin{aligned}
\frac{d}{d\theta}\left(1 - \frac{\lambda q}{\frac{\partial U}{\partial k}(\theta)} + \frac{\eta(\theta)}{\frac{\partial U}{\partial k}(\theta) f(\theta)}\right) &= \frac{d}{d\theta}\left(1 - \frac{\lambda q}{\frac{\partial U}{\partial k}(\theta)}\right) + \frac{d}{d\theta}\left(\frac{\eta(\theta)}{\frac{\partial U}{\partial k}(\theta) f(\theta)}\right) \\
&= \frac{d}{d\theta}\left(1 - \frac{\lambda q}{\frac{\partial U}{\partial k}(\theta)}\right) + \frac{1}{\frac{\partial U}{\partial k}(\theta) f(\theta)} \frac{d}{d\theta}(\eta(\theta)) \\
&\geq 0.
\end{aligned} \tag{65}$$

Since $\frac{d}{d\theta}(\eta(\theta)) \leq 0$, the inequality above implies that $\frac{d}{d\theta}\left(1 - \frac{\lambda q}{\frac{\partial U}{\partial k}(\theta)}\right) \geq 0$ at $\theta = \theta^*$. Therefore,

$$z = \frac{d}{d\theta}\left\{\frac{\partial U}{\partial k}(\theta^*)\right\} \geq 0. \tag{66}$$

However, (62) and (66) contradict the complementarity of housing and non-housing consumption.

Indeed by total differentiation

$$z = \frac{\partial^2 U}{\partial k \partial c}\frac{dc}{d\theta} + \frac{\partial^2 U}{\partial k^2}\frac{dk}{d\theta}$$

and

$$x = \frac{\partial^2 U}{\partial c^2}\frac{dc}{d\theta} + \frac{\partial^2 U}{\partial c \partial k}\frac{dk}{d\theta}.$$

So

$$\begin{bmatrix} \frac{dc}{d\theta} \\ \\ \frac{dk}{d\theta} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 U(c,k)}{\partial c^2} & \frac{\partial^2 U(c,k)}{\partial c \partial k} \\ \\ \frac{\partial^2 U(c,k)}{\partial c \partial k} & \frac{\partial^2 U(c,k)}{\partial k^2} \end{bmatrix}^{-1} \begin{bmatrix} x \\ \\ y \end{bmatrix}.$$

Besides,

$$\frac{d}{d\theta}\{U(c(\theta), k(\theta))\} = \begin{bmatrix} \frac{\partial U}{\partial c} & \frac{\partial U}{\partial k} \end{bmatrix} \begin{bmatrix} \frac{dc}{d\theta} \\ \\ \frac{dk}{d\theta} \end{bmatrix} = \frac{\partial U(c(\theta), k(\theta))}{\partial c} > 0.$$

On the other hand

$$\begin{bmatrix} \frac{\partial U}{\partial c} & \frac{\partial U}{\partial k} \end{bmatrix} \begin{bmatrix} \frac{dc}{d\theta} \\ \\ \frac{dk}{d\theta} \end{bmatrix} = \begin{bmatrix} \frac{\partial U}{\partial c} & \frac{\partial U}{\partial k} \end{bmatrix} \begin{bmatrix} \frac{\partial^2 U(c,k)}{\partial c^2} & \frac{\partial^2 U(c,k)}{\partial c \partial k} \\ \\ \frac{\partial^2 U(c,k)}{\partial c \partial k} & \frac{\partial^2 U(c,k)}{\partial k^2} \end{bmatrix}^{-1} \begin{bmatrix} x \\ \\ y \end{bmatrix} \leq 0,$$

since, by (1)

$$\begin{bmatrix} \frac{\partial U}{\partial c} & \frac{\partial U}{\partial k} \end{bmatrix} \begin{bmatrix} \frac{\partial^2 U(c,k)}{\partial c^2} & \frac{\partial^2 U(c,k)}{\partial c \partial k} \\ \\ \frac{\partial^2 U(c,k)}{\partial c \partial k} & \frac{\partial^2 U(c,k)}{\partial k^2} \end{bmatrix}^{-1} < 0$$

and $x, y \geq 0$. This is the desired contradiction. □

Given Lemma 6 and Lemma 7, it is relatively straightforward to show the main result.

Lemma 8. *In the optimal solution to the "weakly relaxed problem," for all* $\theta \in (\underline{\theta}, \overline{\theta})$,

$$v'(\theta) - \frac{\partial U}{\partial c}(\theta + y(\theta), k(\theta)) = 0.$$

Proof. We show this result by contradiction. If there exists θ^* such that this is not true:

$$v'(\theta^*) > \frac{\partial U}{\partial c}\left(\theta^* + y(\theta^*), k\left(\theta^*\right)\right). \tag{67}$$

Then $\xi(\theta^*) = 0$. By Lemma 7, $\gamma(\theta^*) > 0$, therefore by continuity $\gamma(\theta) > 0$ in some neighborhood of θ^*. So $k(\theta) = k^*$ in this neighborhood. By Lemma 6, $y(\theta) = y^*$ in this neighborhood. This however contradicts (67). □

We have established that the optimal solution to the "weakly relaxed problem" satisfies the incentive compatibility constraint, therefore it is also an optimal solution to the original problem. One immediate consequence is that the multiplier $\xi(\theta)$ in the original problem is positive for all θ. We are now ready to prove the properties in Proposition 1.

Proof of Proposition 1, Parts (i) and (ii). Consider $\theta^* \in \left(\underline{\theta}, \overline{\theta}\right)$. There are two cases:

Case 1: $\gamma(\theta^*) > 0$ then there is pooling at θ^*. Let θ^{**} denote the left most point, such that there is pooling from θ^{**} to θ^*. Formally

$$\theta^{**} = \inf\left\{\theta \in [\underline{\theta}, \theta^*] : \gamma(\theta_1) > 0 \text{ for all } \theta_1 \in (\theta, \theta^*)\right\}.$$

Then $\gamma(\theta^{**}) = 0$ and $\gamma(\theta) > 0$ for all $\theta \in (\theta^{**}, \theta^*)$ (this comes from definition of θ^{**} if $\theta^{**} > \underline{\theta}$. If $\theta^{**} = \underline{\theta}$, then this is also true since $\gamma(\underline{\theta}) = 0$). Consequently,

$$\gamma'(\theta^{**}) = -\eta(\theta^{**}) \geq 0.$$

We show that

$$\frac{\frac{\partial U}{\partial k}\left(\theta^{**} + y(\theta^{**}), k(\theta^{**})\right)}{\frac{\partial U}{\partial c}\left(\theta^{**} + y(\theta^{**}), k(\theta^{**})\right)} \geq q. \tag{68}$$

From (53) and (54) at θ^{**}

$$\begin{aligned}&\xi(\theta^{**})\left(\frac{\partial U}{\partial c}(\theta^{**})\frac{\partial^2 U}{\partial c \partial k}(\theta^{**}) - \frac{\partial U}{\partial k}(\theta^{**})\frac{\partial^2 U}{\partial c^2}(\theta^{**})\right)\\ &= \lambda f(\theta^{**})\left(\frac{\partial U}{\partial k}(\theta^{**}) - q\frac{\partial U}{\partial c}(\theta^{**})\right) + \eta(\theta^{**})\frac{\partial U}{\partial c}(\theta^{**}).\end{aligned}$$

Together with $\xi(\theta^{**}) \geq 0$, $\eta(\theta^{**}) \leq 0$ and (48) at θ^{**}, we obtain (68).

Now, since there is pooling over (θ^{**}, θ^*), and because of (50) we have

$$\frac{\frac{\partial U}{\partial k}\left(\theta^* + y(\theta^*), k(\theta^*)\right)}{\frac{\partial U}{\partial c}\left(\theta^* + y(\theta^*), k(\theta^*)\right)} > \frac{\frac{\partial U}{\partial k}\left(\theta^{**} + y(\theta^{**}), k(\theta^{**})\right)}{\frac{\partial U}{\partial c}\left(\theta^{**} + y(\theta^{**}), k(\theta^{**})\right)} \geq q.$$

Case 2: $\gamma(\theta^*) = 0$. By Lemma 7 $\xi(\theta^*) > 0$.

Since $\gamma(\theta^*) = 0$ and $\gamma(\theta) \geq 0$ for all θ,

$$\gamma'(\theta^*) = -\eta(\theta^*) \geq 0.$$

From (53) and (54) at θ^*

$$\xi(\theta^*)\left(\frac{\partial U}{\partial c}(\theta^*)\frac{\partial^2 U}{\partial c \partial k}(\theta^*) - \frac{\partial U}{\partial k}(\theta^*)\frac{\partial^2 U}{\partial c^2}(\theta^*)\right)$$
$$= \lambda f(\theta^*)\left(\frac{\partial U}{\partial k}(\theta^*) - q\frac{\partial U}{\partial c}(\theta^*)\right) + \eta(\theta^*)\frac{\partial U}{\partial c}(\theta^*).$$

Together with $\xi(\theta^*) > 0$, $\eta(\theta^*) \leq 0$ and (48) at θ^*, we obtain

$$\frac{\frac{\partial U}{\partial k}\left(\theta^* + y(\theta^*), k(\theta^*)\right)}{\frac{\partial U}{\partial c}\left(\theta^* + y(\theta^*), k(\theta^*)\right)} > q.$$

Now, at the top, the proof of Part (iii) below shows that there is no pooling at the top, $\gamma(\theta) = 0$ in some neighborhood to the left of $\bar{\theta}$. So $\eta(\bar{\theta}) = -\gamma'(\bar{\theta}) = 0$. In addition $\xi(\bar{\theta}) = 0$. Equations (53) and (54) at $\bar{\theta}$ then implies that

$$\frac{\frac{\partial U}{\partial k}\left(\bar{\theta} + y(\bar{\theta}), k(\bar{\theta})\right)}{\frac{\partial U}{\partial c}\left(\bar{\theta} + y(\bar{\theta}), k(\bar{\theta})\right)} = q.$$

Similarly, if there is no pooling at the bottom then there is no distortion at the bottom. In general, by the continuity of c and k, there is also (weakly) positive distortion at the bottom. □

Proof of Proposition 1, Part (iii). Assume by contradiction that this property does not hold, i.e. there is pooling at the top. Let $\left[\theta^*, \bar{\theta}\right]$ denote the maximum pooling interval. If $\theta^* > \underline{\theta}$. By the definition of θ^*, $\gamma(\theta^*) = 0$. If $\theta^* = \underline{\theta}$, we also have $\gamma(\theta^*) = \gamma(\underline{\theta}) = 0$ by (60).

Evaluating (53) and (54) at $\bar{\theta}$, we have

$$\frac{\partial U}{\partial c}(\bar{\theta})\left(\mu(\bar{\theta})+f(\bar{\theta})\right)-\lambda f(\bar{\theta})=0.$$

This implies $\left(\mu(\bar{\theta})+f(\bar{\theta})\right)>0$. And

$$\frac{\partial U}{\partial k}(\theta)\left(\mu(\bar{\theta})+f(\bar{\theta})\right)+\eta\left(\bar{\theta}\right)-\lambda q f(\theta)=0.$$

Since $\gamma(\bar{\theta})=0$ and $\gamma(\theta)\geq 0$ for all θ, $\eta(\bar{\theta})=-\gamma'(\bar{\theta})\geq 0$. Therefore

$$\frac{\frac{\partial U}{\partial k}(\bar{\theta})}{\frac{\partial U}{\partial c}(\bar{\theta})}\leq q. \tag{69}$$

From (53) and (54) at θ^*

$$\xi(\theta^*)\left(\frac{\partial U}{\partial c}(\theta^*)\frac{\partial^2 U}{\partial c\partial k}(\theta^*)-\frac{\partial U}{\partial k}(\theta^*)\frac{\partial^2 U}{\partial c^2}(\theta^*)\right)=\lambda f(\theta^*)\left(\frac{\partial U}{\partial k}(\theta^*)-q\frac{\partial U}{\partial c}(\theta^*)\right)+\eta(\theta^*)\frac{\partial U}{\partial c}(\theta^*).$$

Because $\gamma(\theta^*)=0$ and $\gamma(\theta)\geq 0$ for all $\theta\geq\theta^*$, $\eta(\theta^*)=-\gamma'(\theta^*)\leq 0$. In addition we have $\xi(\theta^*)\geq 0$ and (48) at θ^*, so

$$\frac{\frac{\partial U}{\partial k}(\theta^*)}{\frac{\partial U}{\partial c}(\theta^*)}\geq q. \tag{70}$$

Combining (69) and (70), we obtain:

$$\frac{\frac{\partial U}{\partial k}(\theta^*)}{\frac{\partial U}{\partial c}(\theta^*)}\geq\frac{\frac{\partial U}{\partial k}(\bar{\theta})}{\frac{\partial U}{\partial c}(\bar{\theta})}.$$

This is a contradiction since there is pooling over $\left[\theta^*,\bar{\theta}\right]$ and therefore

$$\frac{d}{d\theta}\left(\frac{\frac{\partial U}{\partial k}(\theta)}{\frac{\partial U}{\partial c}(\theta)}\right)=\frac{\partial}{\partial\theta}\left(\frac{\frac{\partial U}{\partial k}(\theta)}{\frac{\partial U}{\partial c}(\theta)}\right)>0,$$

for $\theta\in\left[\theta^*,\overline{\theta}\right]$, which implies

$$\frac{\frac{\partial U}{\partial k}(\theta^*)}{\frac{\partial U}{\partial c}(\theta^*)}<\frac{\frac{\partial U}{\partial k}(\bar{\theta})}{\frac{\partial U}{\partial c}(\bar{\theta})}.$$

□

$$\frac{\partial U}{\partial k}(y^*(\overline{\theta},B),k^*(\overline{\theta},B)) \;=\; \lambda^*(B) \quad \text{and} \quad \frac{\partial U}{\partial y}(y^*(\overline{\theta},B),k^*(\overline{\theta},B)) \;=\; \lambda^*(B).$$

E Proof of Proposition 2

Proof. Part (ii). Observe that

$$\frac{d\Gamma(\theta)}{d\theta} \;=\; -\frac{d}{d\theta}(c^*(\theta)+qk^*(\theta)) \iff$$
$$\frac{dy^*}{d\theta}+q\frac{dk^*}{d\theta} = -1-\frac{d\Gamma(\theta)}{d\theta} \quad \iff \quad -\frac{dy^*}{d\theta} = q\frac{dk^*}{\theta}+1+\frac{d\Gamma(\theta)}{d\theta}$$

By IC and the distortion result of Proposition 1,

$$-\frac{dy^*}{d\theta}/\frac{dk^*}{d\theta} = q\Delta(\theta).$$

Combining these two gives

$$q\frac{dk^*}{d\theta}(\Delta(\theta)-1) \;=\; 1+\frac{d\Gamma(\theta)}{d\theta}$$

□

F Proof of Proposition 4

Proof of Proposition 4 Part 1. Let $s=\frac{\overline{\theta}-\underline{\theta}}{2\lambda(c(\underline{\theta}))^2}$, we rewrite (36) as:

$$\frac{\log\left(1+s\right)\left(1+\sqrt{1+s}\right)}{\sqrt{s}} = 2\sqrt{2\lambda\left(\overline{\theta}-\underline{\theta}\right)}.$$

After lengthy algebras, we show that the left hand side is strictly increasing in s and is equal to 0 at $s=0$ and to ∞ at $s=\infty$. So there exists a unique solution s^* to this equation.

Now, from the IC constraint, we have

$$\frac{1}{k(\theta)}k'(\theta) = \frac{1}{c(\theta)}\left(1 - \frac{1}{4\lambda c(\theta)}\right).$$

Therefore, in order for k to be increasing in θ, we require $4\lambda c(\theta) \geq 1$. Because c is increasing in θ, this is equivalent to $4\lambda c(\underline{\theta}) = 4\lambda q k(\underline{\theta}) \geq 1$. From the definition of s^*, $s^* \leq 8\lambda\left(\overline{\theta} - \underline{\theta}\right)$. Equivalently

$$\frac{\log\left(1 + 8\lambda\left(\overline{\theta} - \underline{\theta}\right)\right)\left(1 + \sqrt{1 + 8\lambda\left(\overline{\theta} - \underline{\theta}\right)}\right)}{\sqrt{8\lambda\left(\overline{\theta} - \underline{\theta}\right)}} > 2\sqrt{2\lambda\left(\overline{\theta} - \underline{\theta}\right)}.$$

After lengthy algebras, we show that

$$\frac{\log(1+t)\left(1 + \sqrt{1+t}\right)}{t}$$

is strictly decreasing in t, and it is equal to 2 at $t = 0$ and to 0 at $t = \infty$. Therefore, the above inequality is equivalent to

$$8\lambda(\overline{\theta} - \underline{\theta}) < t^*,$$

where t^* is uniquely determined by

$$\frac{\left(1 + \sqrt{1+t}\right)\log\left(1+t\right)}{t} = 1. \tag{71}$$

After some algebra manipulation, it is easy to see that $t^* = \delta - 1$, where δ is defined in the Proposition. The numerical value of t^* is 11.3402.

□

Proof of Proposition 4 Part 2. We look for continuous allocations $k(.), y(.)$ and the multipliers $\mu(.), \gamma(.), \eta(.)$ that satisfy the F.O.Cs following the conjecture stated in the proposition.

From (53), because $f(.), y(.), k(.), \xi(.)$ are all continuous in θ, $\mu(\theta)$ is continuous in θ.

From (54), because $f(.), y(.), k(.), \xi(.), \mu(.)$ are all continuous in θ, $\eta(\theta)$ is continuous in θ.

F.O.C. in $y(\theta)$, for $\theta \in [\underline{\theta}, \theta^*]$:

$$\frac{1}{2\left(\theta+\underline{y}\right)}\left(\mu(\theta)+\bar{f}\right)-\lambda\bar{f}=-\frac{1}{2\left(\theta+\underline{y}\right)^2}\xi(\theta) \tag{72}$$

Therefore

$$\frac{1}{2(\theta+\underline{y})}\bar{f}-\lambda\bar{f}=\frac{d}{d\theta}\left(\frac{1}{2(\theta+y)}\xi(\theta)\right)$$

So

$$\frac{1}{2(\theta+\underline{y})}\xi(\theta)=\frac{\bar{f}}{2}\left(\log\left(\theta+\underline{y}\right)-\log\left(\underline{\theta}+\underline{y}\right)\right)-\lambda\bar{f}\left(\theta-\underline{\theta}\right)$$

or

$$\xi(\theta)=\bar{f}\left(\theta+\underline{y}\right)\left(\log\left(\theta+\underline{y}\right)-\log\left(\underline{\theta}+\underline{y}\right)\right)-2\lambda\bar{f}\left(\theta-\underline{\theta}\right)\left(\theta+\underline{y}\right) \tag{73}$$

and

$$\mu(\theta)=-\xi'(\theta)=-\bar{f}\left(\log\left(\theta+\underline{y}\right)-\log\left(\underline{\theta}+\underline{y}\right)\right)-\bar{f}+2\lambda\bar{f}\left(\theta+\underline{y}\right)+2\lambda\bar{f}\left(\theta-\underline{\theta}\right)$$

F.O.C in k:

$$\frac{1}{2\underline{k}}\left(\mu(\theta)+\bar{f}\right)+\eta(\theta)-\bar{f}\lambda q=0. \tag{74}$$

At $\theta=\theta^*$, $\eta(\theta^*)=0$, so

$$\mu(\theta^*)+\bar{f}-2\bar{f}\lambda q\underline{k}=0.$$

Therefore, from the earlier expression for $\mu(\theta)$:

$$-\log\left(c^*\right)+\log\left(c^*-\left(\theta^*-\underline{\theta}\right)\right)+2\lambda c^*+2\lambda\left(\theta^*-\underline{\theta}\right)-2\lambda q\underline{k}=0, \tag{75}$$

where $c^*=\underline{y}+\theta^*$.

Integrating (74) from $\underline{\theta}$ to θ^*, we obtain

$$\frac{1}{2\underline{k}}\left(\xi(\underline{\theta})-\xi(\theta^*)+\bar{f}\left(\theta^*-\underline{\theta}\right)\right)+\gamma(\underline{\theta})-\gamma(\theta^*)-\bar{f}\lambda q\left(\theta^*-\underline{\theta}\right)=0.$$

Given that $\gamma(\theta^*)=\gamma(\underline{\theta})=0$ and $\xi(\underline{\theta})=0$ and from the earlier expression for g, this is equivalent to:

$$-c^*\left(\log c^*-\log\left(c^*-\left(\theta^*-\underline{\theta}\right)\right)\right)+2\lambda\left(\theta^*-\underline{\theta}\right)c^*+\left(\theta^*-\underline{\theta}\right)-2\lambda q\underline{k}\left(\theta^*-\underline{\theta}\right)=0 \tag{76}$$

Lastly,

$$\log\left((c^*)^2 + \frac{\overline{\theta}-\theta^*}{2\lambda}\right)$$
$$= \log(q\underline{k}) + 4\lambda\sqrt{(c^*)^2 + \frac{\overline{\theta}-\theta^*}{2\lambda}} - 4\lambda c^* + \log c^*. \tag{77}$$

We show that there is a solution $\left(\underline{y}, \underline{k}, \theta^*\right)$ to (75)-(77).

From the first two equations, (75) and (76), we have

$$c^*\frac{-\log c^* + \log\left(c^* - (\theta^* - \underline{\theta})\right)}{\theta^* - \underline{\theta}} + 2\lambda c^* + 1 = -\log\left(c^*\right) + \log\left(c^* - (\theta^* - \underline{\theta})\right) + 2\lambda c^* + 2\lambda\left(\theta^* - \underline{\theta}\right).$$

Let $\zeta = \frac{\theta^* - \underline{\theta}}{c^*}$, this expression simplifies to:

$$\frac{\log(1-\zeta)}{\zeta} + 1 = \log(1-\zeta) + 2\lambda c^* \zeta$$

So c^* is a function of ζ

$$c^* = \hat{c}(\zeta) \equiv \frac{\frac{\log(1-\zeta)}{\zeta} - \log(1-\zeta) + 1}{2\lambda\zeta}$$

Using Taylor expansion

$$\hat{c}(\zeta) = \frac{1}{2\lambda}\sum_{n=0}^{\infty}\frac{\zeta^n}{(n+1)(n+2)}, \tag{78}$$

which is strictly increasing in ζ.

Notice that $\lim_{\zeta\to 0}\hat{c}(\zeta) = \frac{1}{4\lambda}$ and $\lim_{\zeta\to 1}\hat{c}(\zeta) = \frac{1}{2\lambda}$. Because $8\lambda(\overline{\theta} - \underline{\theta}) > t^* > 4$,

$$1 < 2\lambda(\overline{\theta} - \underline{\theta}).$$

So

$$\lim_{\zeta\to 1}\zeta\hat{c}(\zeta) = \frac{1}{2\lambda} < (\overline{\theta} - \underline{\theta}),$$

which implies at $\zeta = 1$, $\theta^* = \underline{\theta} + \lim_{\zeta\to 1}\zeta\hat{c}(\zeta) < \overline{\theta}$.

Solving for $q\underline{k}$ from (75) and (77), we obtain another equation:

$$\left((c^*)^2 + \frac{\overline{\theta} - \underline{\theta}}{2\lambda} - \frac{c^*\zeta}{2\lambda}\right) \exp\left(-4\lambda\sqrt{(c^*)^2 + \frac{\overline{\theta} - \underline{\theta}}{2\lambda} - \frac{c^*\zeta}{2\lambda}}\right) \frac{\exp(4\lambda c^*)}{c^*} = \frac{\log(1-\zeta)}{2\lambda} + c^* + c^*\zeta. \tag{79}$$

From the earlier expression for c^*, (78), this corresponds to one equation, and one unknown in ζ. Let $\Delta(\zeta)$ denote the difference between the RHS and LHS of this equation.

We will show that $\Delta(0) < 0$ and $\Delta(1^-) > 0$.

It is easy to show that $\lim_{\zeta\uparrow 1} \Delta(\zeta) = +\infty > 0$, since $\lim_{\zeta\uparrow 1} \hat{c}(\zeta) = \frac{1}{2\lambda} < \infty$ and $\lim_{\zeta\uparrow 1} \log(1-\zeta) = -\infty$.

Now at $\zeta = 0$, $\hat{c}(0) = \frac{1}{4\lambda}$, so

$$\Delta(0) = \left((\hat{c}(0))^2 + \frac{\overline{\theta} - \underline{\theta}}{2\lambda}\right) \exp\left(-4\lambda\sqrt{(\hat{c}(0))^2 + \frac{\overline{\theta} - \underline{\theta}}{2\lambda}}\right) \frac{\exp(4\lambda\hat{c}(0))}{\hat{c}(0)} - \hat{c}(0).$$

Let $\tilde{s} = \frac{\overline{\theta} - \underline{\theta}}{2\lambda(\hat{c}(0))^2}$, after lengthy algebra, we desired inequality $\Delta(0) < 0$ as:

$$\frac{\log(1+\tilde{s})\left(1+\sqrt{1+\tilde{s}}\right)}{\sqrt{\tilde{s}}} < 2\sqrt{2\lambda\left(\overline{\theta} - \underline{\theta}\right)}. \tag{80}$$

Consider the solution to the relaxed problem in Part 1 of this Proposition, $c_R(\theta)$ and $k_R(\theta)$ (k_R might not be increasing). At $\underline{\theta}$,

$$s^* = \frac{\overline{\theta} - \underline{\theta}}{2\lambda\left(c_R(\underline{\theta})\right)^2}$$

and

$$\frac{\log(1+s^*)\left(1+\sqrt{1+s^*}\right)}{\sqrt{s^*}} = 2\sqrt{2\lambda\left(\overline{\theta} - \underline{\theta}\right)}.$$

As shown in the first part of this Proposition, the RHS of (80) is strictly increasing in $\tilde{s}$. Therefore (80) is equivalent to

$$\tilde{s} < s^*$$

or, by definition of $\tilde{s}$ and s^*,

$$\hat{c}(0) = \frac{1}{4\lambda} > c_R(\underline{\theta}).$$

Indeed this is the case since $8\lambda(\overline{\theta}-\underline{\theta}) > t^*$.

So there exists $\zeta^* \in (0,1)$ such that $\Delta(\zeta^*) = 0$. We determine $\theta^*, \underline{y}$ and $\underline{k}$ as $\theta^* = \hat{c}(\zeta^*)\zeta^* + \underline{\theta}$, $\underline{y} = \hat{c}(\zeta^*) - \theta^*$ and $\underline{k}$ is a function of c^* and ζ^* as in either (75) or (77). It is easy to verify that $\left(\theta^*, \underline{y}, \underline{k}\right)$ solves (75)-(77).

We can verify that for all $\theta \in (\underline{\theta}, \theta^*)$:

$$\gamma(\theta) = \frac{1}{2\underline{k}}\left(-\xi(\theta) + \bar{f}\left(\theta - \underline{\theta}\right)\right) - \bar{f}\lambda q\left(\theta - \underline{\theta}\right) > 0.$$

We first show by contradiction that there is no local minimum of γ in $(\underline{\theta}, \theta^*)$. Assume by contradiction that there exists a local minimum $\tilde{\theta} \in (\underline{\theta}, \theta^*)$. Then $\gamma'(\tilde{\theta}) = 0$ and $\gamma''(\tilde{\theta}) \geq 0$.

From the expression for γ:

$$\gamma''(\theta) = -\frac{1}{2\underline{k}}g''(\theta) = -\frac{1}{2\underline{k}}\bar{f}\left(\frac{1}{\theta + \underline{y}} - 4\lambda\right)$$

is increasing in θ. Therefore $\gamma''(\theta) > \gamma''(\tilde{\theta}) \geq 0$ for $\theta \in \left(\tilde{\theta}, \theta^*\right)$. From the definition of θ^*, $\gamma'(\theta^*) = -\eta(\theta^*) = 0$. So $\gamma'(\tilde{\theta}) < 0$ which contradicts the property that $\gamma'(\tilde{\theta}) = 0$. So γ does not have a local minimum in $(\underline{\theta}, \theta^*)$.

At $\underline{\theta}$, $\gamma(\underline{\theta}) = 0$ and by construction, at θ^*, $\gamma(\theta^*) = 0$. So $\gamma \geq 0$ for all $\theta \in [\underline{\theta}, \theta^*]$. In addition, $\gamma(\theta) > 0$ for all $\theta \in (\underline{\theta}, \theta^*)$ (otherwise, γ would have a local minimum in $(\underline{\theta}, \theta^*)$.

Given $\theta^*, \underline{y}, \underline{k}$, we determine the allocation over $\left[\theta^*, \overline{\theta}\right]$ using the derivation in Subsection 5

$$c(\theta) = \sqrt{\left(c^*\right)^2 + \frac{\theta - \theta^*}{2\lambda}}$$

and

$$\begin{aligned}\log(k(\theta)) = \log(\underline{k}) \\ + 4\lambda\sqrt{\left(c^*\right)^2 + \frac{\theta - \theta^*}{2\lambda}} - 4\lambda c^* \\ - \frac{1}{2}\log\left(\left(c^*\right)^2 + \frac{\theta - \theta^*}{2\lambda}\right) + \log\left(c^*\right).\end{aligned} \tag{81}$$

Since $c^* = \hat{c}(\zeta^*) > \hat{c}(0) = \frac{1}{4\lambda}$, $k(.)$ is strictly increasing over $\left(\theta^*, \overline{\theta}\right)$.

□

G Proof of Proposition 5

To prove Proposition 5, we apply the general analysis in Subsection 6.1 to this special case with separable utility function:

$$U(c,k) = u_1(c) + u_2(k).$$

Then

$$\begin{aligned} x =& u_1'(c) \\ z =& u_2'(k), \end{aligned}$$

and

$$\begin{aligned} c =& H^c(x) = \left(u_1'\right)^{-1}(x) \\ k =& H^k(z) = \left(u_2'\right)^{-1}(z). \end{aligned}$$

Furthermore,

$$\frac{\partial^2 U(c,k)}{\partial c^2} = \frac{1}{(H^c)'(x)}$$

and

$$\frac{\partial^2 U(c,k)}{\partial c \partial k} = 0.$$

From the expression for g, (41),

$$\xi(\theta) = \lambda f(\theta) K(x(\theta), z(\theta)),$$

where, in this case,

$$\begin{aligned} K(x,z) =& -\frac{z - qx}{z}(H^c)'(x) \\ =& -(H^c)'(x) + \frac{qx}{z}(H^c)'(x). \end{aligned}$$

Therefore

$$\frac{\partial K}{\partial x} = -(H^c)''(x) + \frac{q}{z}(H^c)'(x) + \frac{qx}{z}(H^c)''(x),$$

and

$$\frac{\partial K}{\partial z} = -\frac{qx}{z^2}\left(H^c\right)'(x).$$

The differential equations (42) and (43), for $x(\theta)$ and $z(\theta)$ can be re-written as:

$$\begin{aligned} & x\left(1-\lambda\frac{f'}{f}K(x,z)-\lambda\left(\frac{\partial K}{\partial x}x'+\frac{\partial K}{\partial z}z'\right)\right)-\lambda \\ & = \frac{1}{\left(H^c\right)'(x)}\lambda K(x,z). \end{aligned} \tag{82}$$

and

$$x\left(\left(H^c\right)'(x)x'-1\right)+z\left(\left(H^k\right)'(z)z'\right)=0. \tag{83}$$

Armed with these results, we can show the following lemma.

Lemma 9. *Assuming that* $\frac{(H^c)'(x)}{\left(H^k\right)'(qx)}$ *is weakly increasing in* x*. We have* $\frac{dk}{d\theta}(\bar{\theta}) > \frac{dk}{d\theta}(\underline{\theta})$*.*

Proof. As shown in Proposition 1, there is no pooling at the top, i.e. $\frac{dk(\bar{\theta})}{d\theta} > 0$. Therefore, if the desired inequality is immediately satisfied if there is pooling at the bottom, i.e. $\frac{dk(\underline{\theta})}{d\theta} = 0$.

We just have to show the inequality if there is no pooling at the bottom. In this case, at $\theta = \underline{\theta}$ or $\bar{\theta}$, $z = qx$. Therefore, at $\theta = \underline{\theta}$ or $\bar{\theta}$, $K(x,z) = 0$ and

$$\frac{\partial K}{\partial x} = \frac{1}{x}\left(J^c\right)'(x),$$

and

$$\frac{\partial K}{\partial z} = -\frac{1}{x}\left(J^c\right)'(x).$$

Conditions (82) and (83) become

$$\begin{aligned} & x-\lambda\left(H^c\right)'(x)x'\left(1+\frac{\left(H^c\right)'(x)}{\left(H^k\right)'(z)}\right)+\frac{\lambda\left(H^c\right)'(x)}{q\left(H^k\right)'(z)}-\lambda \\ & = 0, \end{aligned}$$

and

$$\left(H^c\right)'(x)x'+q\left(H^k\right)'(z)z'=1.$$

After algebra manipulations:

$$\left(H^c\right)'(x)x' = \frac{x-2\lambda}{\lambda\left(q+\frac{(H^c)'(x)}{q\left(H^k\right)'(qx)}\right)}+1.$$

Therefore

$$\begin{aligned}\frac{dk}{d\theta} &= \left(H^k\right)'(z)z' \\ &= \frac{1}{q}\left(1-\left(H^c\right)'(x)x'\right) \\ &= \frac{1}{q}\frac{2\lambda-x}{\lambda\left(q+\frac{(H^c)'(x)}{q\left(H^k\right)'(qx)}\right)}.\end{aligned}$$

Since $\frac{(H^c)'(x)}{q\left(H^k\right)'(qx)}$ is weakly increasing in x, $\frac{dk}{d\theta}$ is strictly decreasing in x. From the assumption that the optimal contract is separating at the top and at the bottom, $k(\overline{\theta}) > k(\underline{\theta})$, so $x(\overline{\theta}) = \frac{1}{q}z(\overline{\theta}) < x(\underline{\theta}) = \frac{1}{q}z(\underline{\theta})$. Therefore, $\frac{dk}{d\theta}(\bar{\theta}) > \frac{dk}{d\theta}(\underline{\theta})$. □

Proof of Proposition 5. This proposition is a direct application of Lemma 9. □

H Figures for Robustness

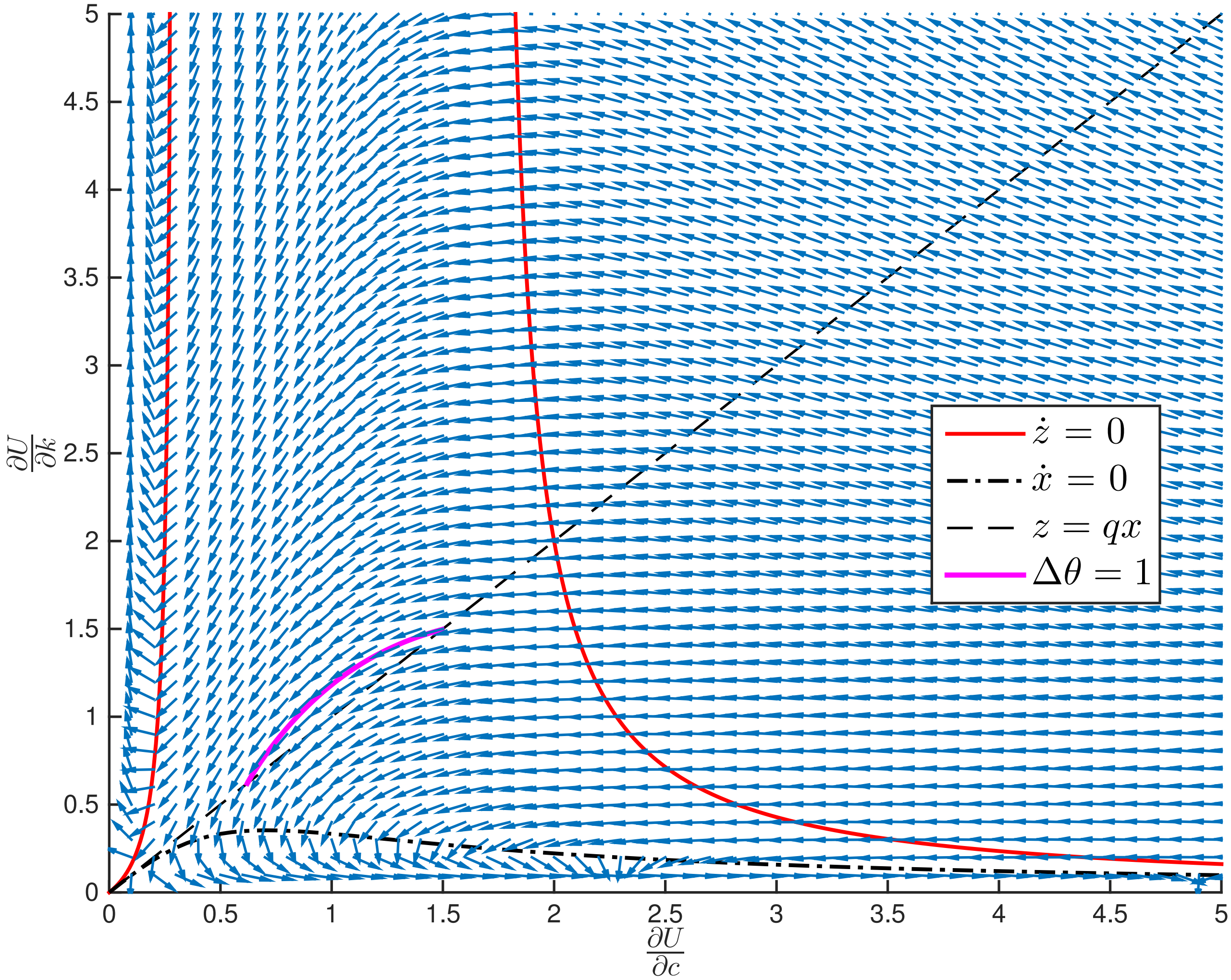

Figure 5: Phase Diagram for Example 1

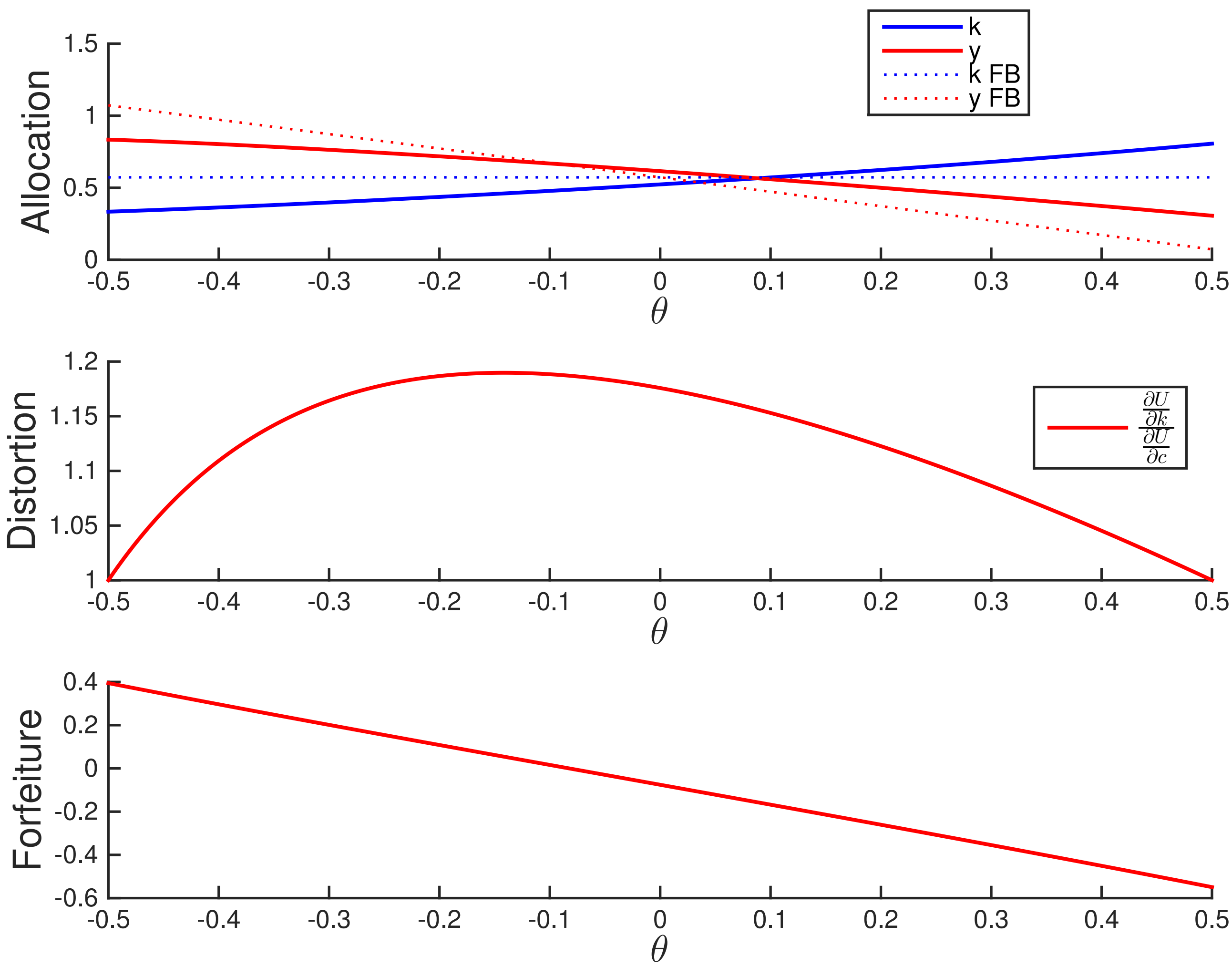

Figure 6: Allocations, Distortion, and Forfeiture for Example 1

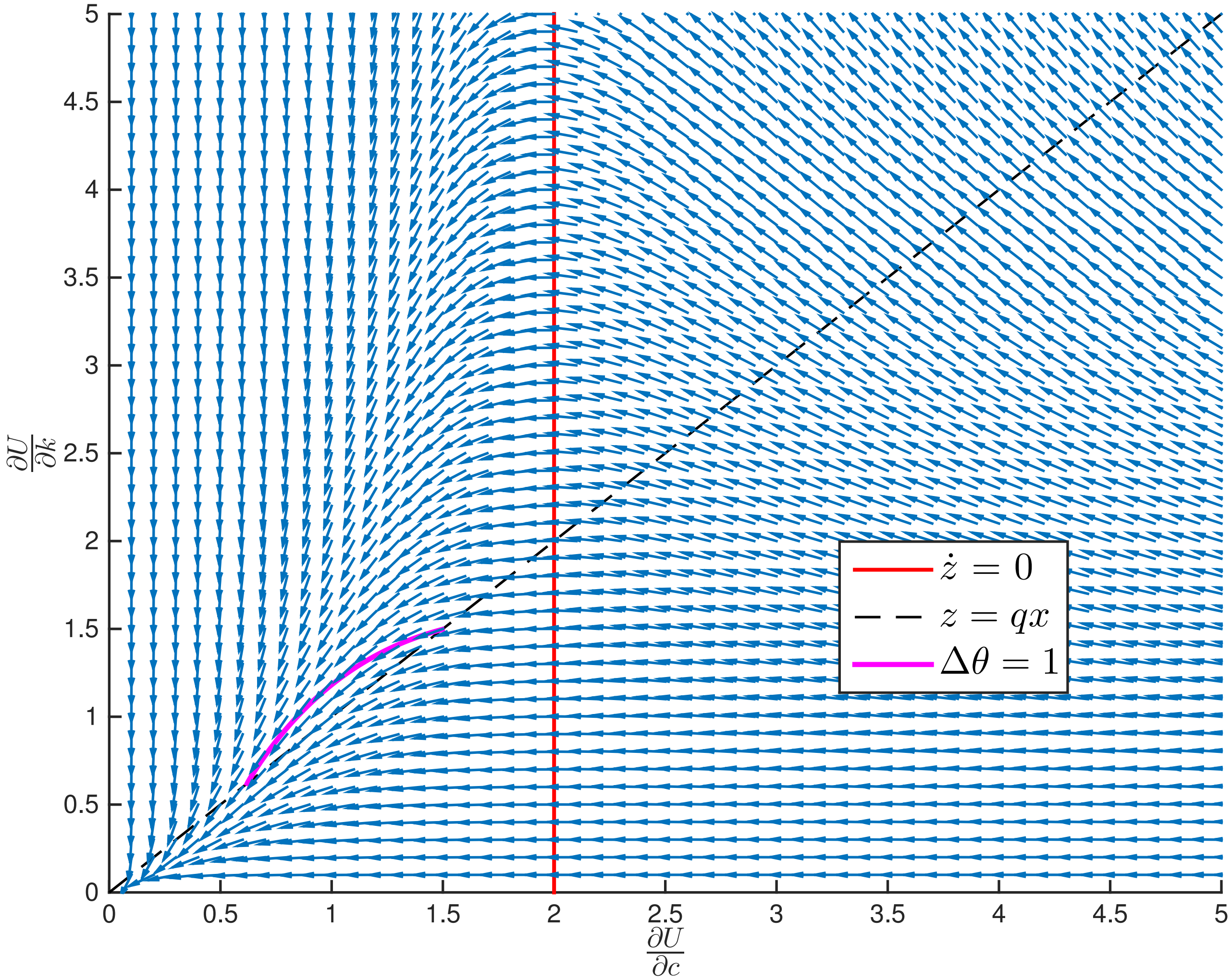

Figure 7: Phase Diagram for Example 2

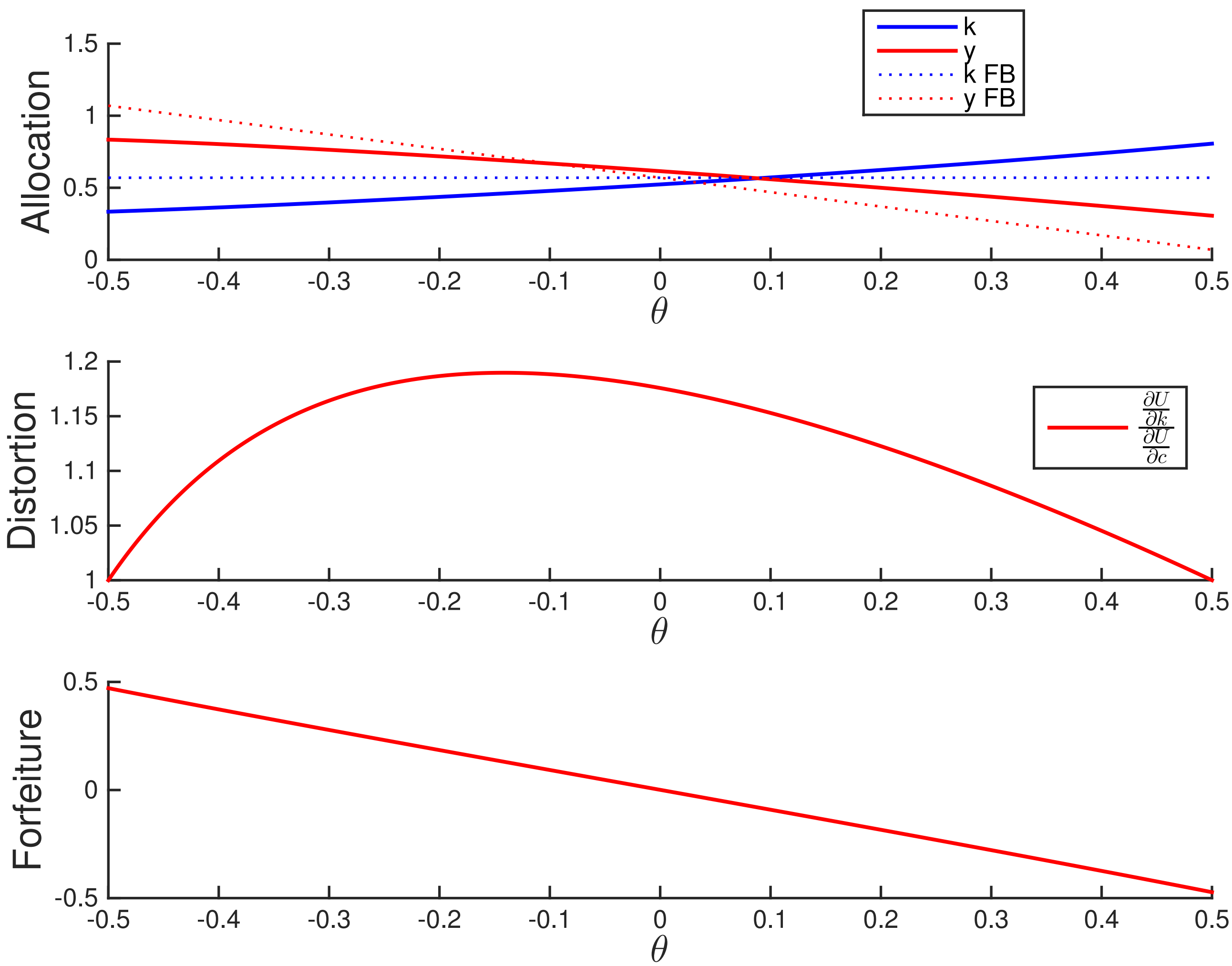

Figure 8: Allocations, Distortion, and Forfeiture for Example 2

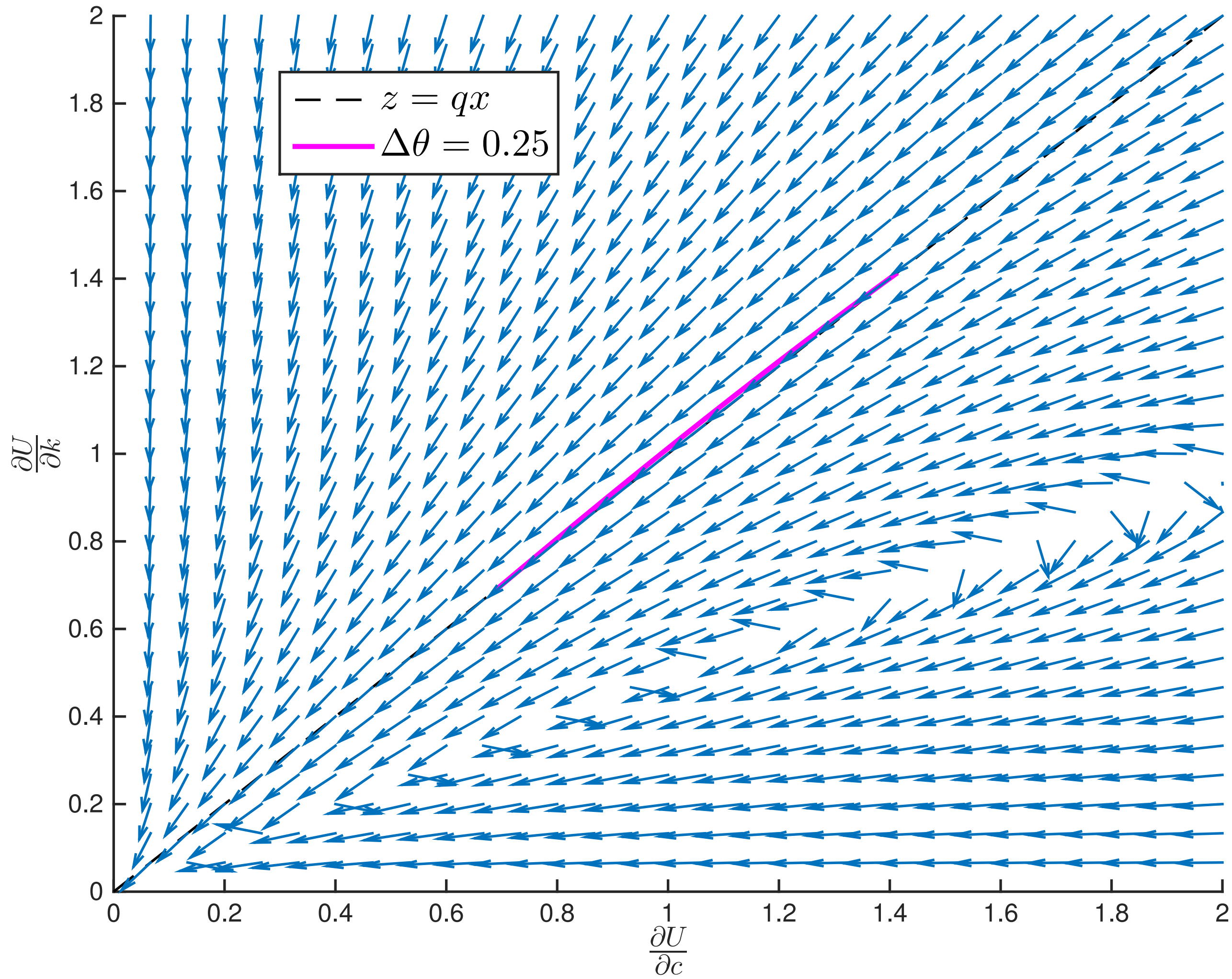

Figure 9: Phase Diagram for Example 3

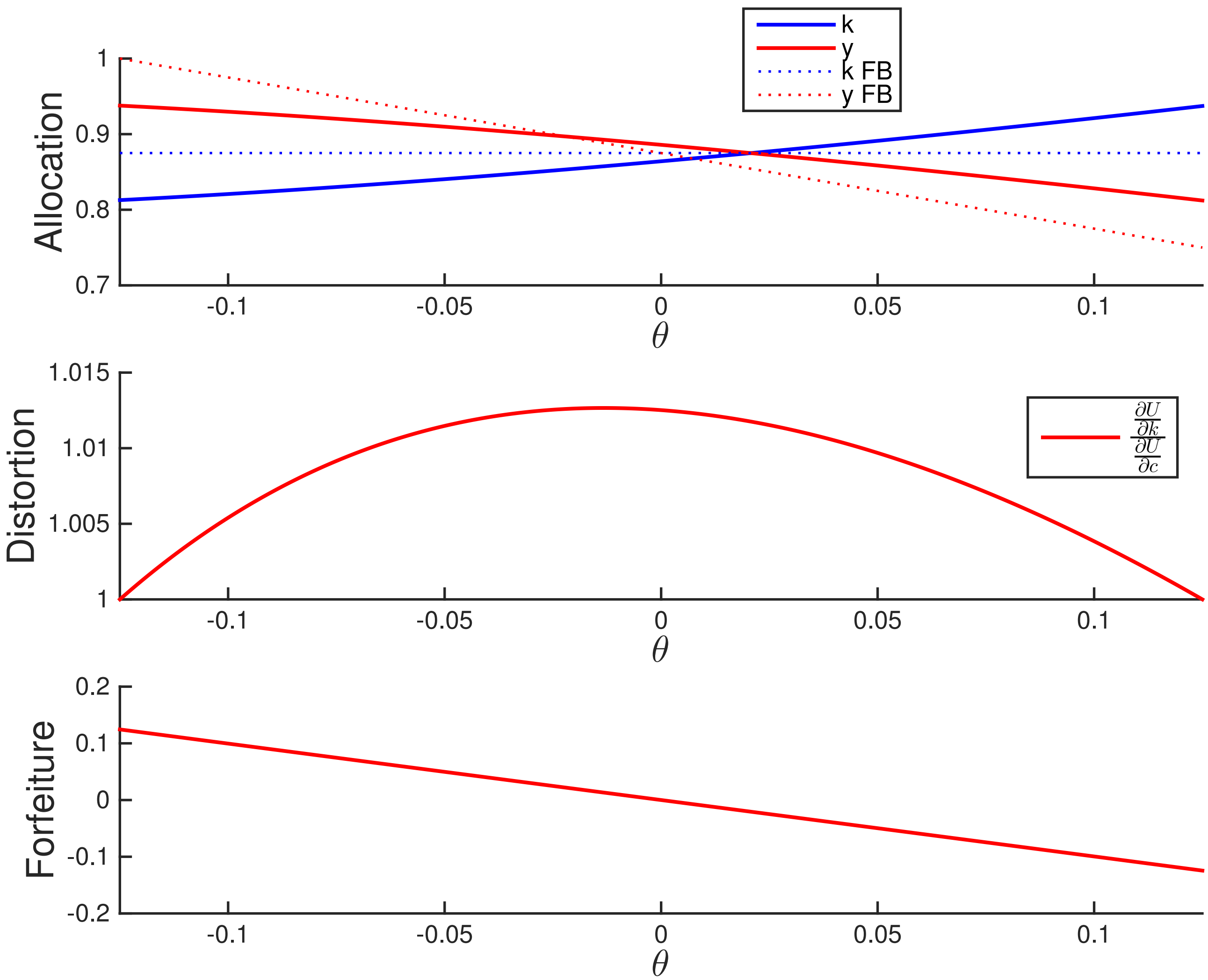

Figure 10: Allocations, Distortion, and Forfeiture for Example 3