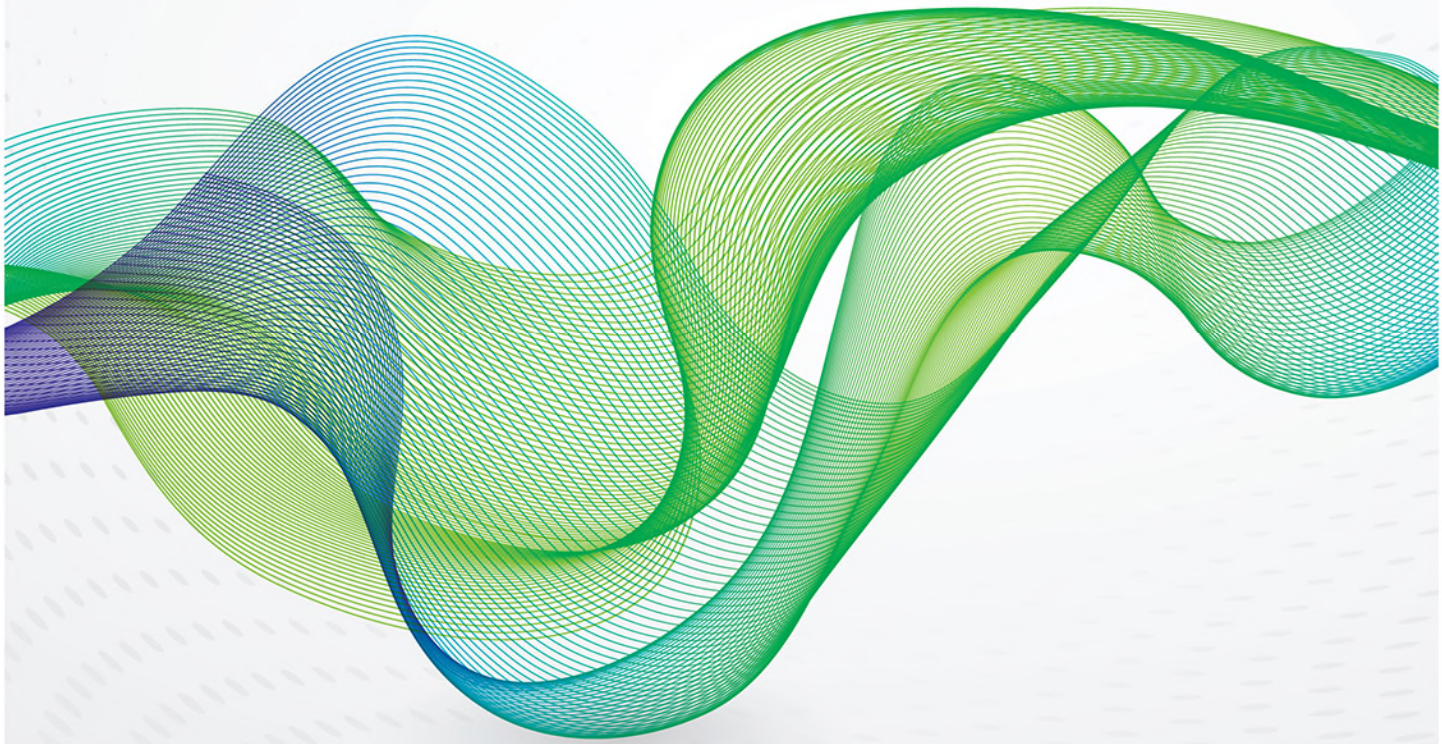
THE OXFORD INSTITUTE FOR ENERGY STUDIES

July 2016

Flexibility-Enabling Contracts in Electricity Markets



OIES PAPER: EL 21

Luis Boscán
Rahmatallah Poudineh

ISBN 978-1-78467-063-4

Contents

Figures

Tables

Flexibility-Enabling Contracts in Electricity Markets

Luis Boscán[*]
PhD fellow, Copenhagen Business School

Rahmatallah Poudineh
Senior research fellow, Oxford Institute for Energy Studies

Abstract

As the share of intermittent renewable energy increases in the generation mix, power systems are exposed to greater levels of uncertainty and risk, which requires planners, policy and business decision makers to incentivise flexibility, that is: their adaptability to unforeseen variations in generation and demand. The greater need for flexibility, along with the fact that its provision is costly, highlights the importance of efficient procurement. As a commodity, flexibility has multiple attributes such as capacity, ramp rate, duration and lead time among which there are complementarities. Additionally, along with traditional sources, which already enable flexibility, a number of business models, such as thermostat-based demand response, aggregators and small storage providers, are emerging in electricity markets and expected to constitute important sources of flexibility in future decentralised power systems. However, due to presence of high transaction costs, relative to the size of resource, the emerging small resources cannot directly participate in an organised electricity market and/or compete. This paper asks the fundamental question of *how should the provision of flexibility, as a multi-dimensional commodity, be incentivised in this context?* We model the procurement of flexibility services from emerging small resources through bilateral contracts in a multidimensional adverse selection setting. We take a normative perspective and show how efficient contracts for flexibility services can be designed given its peculiarity as an economic commodity. Through a simulation analysis we elucidate the applicability of the proposed model and demonstrate the way it can be utilised in, for example, a thermostat based demand response programme.

Keywords: Flexibility, bilateral contracts, electricity markets, mechanism design

JEL classification: L94, D82, D86

[*] Luis *Boscán is currently Postdoc at the Energy Economics and Regulation Group, Technical University of Denmark (DTU)*

1. Introduction

Integrating renewables into electricity systems poses a number of operational challenges (on this topic see, e.g., Morales et al. (2014)). Transmission system operators, which ensure system balancing at all times, must face higher short-term uncertainty as renewable production is stochastic. In the absence of widely available electricity storage, there is a growing need for power systems to adapt to the fluctuation imposed by renewables, that is: to increase their flexibility. While a universal definition for *flexibility* doesn't exist in the technical literature on power systems, it is the encompassing word used to describe the ability of power systems to respond to demand and supply variations over various time horizons.[2] Moreover, while the term is relatively new, the concept is not, as the challenges related to variable demand and generation outage have existed since the dawn of the power industry.

Given the increased deployment of renewables, the technical and policy literatures on power systems have shown greater interest in the topic of flexibility. For example, Lannoye et al. (2012) and Ulbig and Adersson (2015) have focused on the important question of measuring power system flexibility, while associations like EURELECTRIC (2015) have focused on regulatory recommendations to increase the role of demand side flexibility in European electricity markets.

However, very few studies have recognised the need to incentivise the provision of power system flexibility or that it has multiple attributes or dimensions. One notable exception is the report by Ela et al. (2014) who focus on the incentives to enable flexibility in short-term power system operation and claim that "*different types* of resources excel at *different forms of flexibility*, and they also have *different cost impacts* when providing flexibility".

In line with Ela et al. (2014), a key claim of this paper is that - unlike commodities traded in existing electricity markets, such as energy or capacity - power system flexibility is a commodity with multiple attributes. These include, for example, capacity, ramp rate, duration and lead time for demand-side resources. Therefore, buyers have different preferences over the elements that compose flexibility, but sellers are also constrained by the technology they possess, creating thus heterogeneity in the commodity space of flexibility. From the buyer perspective, the value of different attributes of flexibility depends on the specific purpose and conditions (e.g., sometimes for the user, ramp rate is more important than other features of flexibility and some other times duration of response). On the supply side of the market, sellers of flexibility also have different degrees of efficiency across flexibility components.

The inherent multi-attribute nature of flexibility, its heterogeneity and the imperfect complementarity among the consumption and production of its composing elements create a set of unique characteristics that have not been analysed thus far. Its economic implications are much more than a theoretical curiosity and have implications of practical relevance for energy policy makers in general and system operators in particular, because they must incentivise the efficient provision of flexibility if an increasing reliance on renewables is to be achieved.

Furthermore, flexibility is topically relevant as recent technological innovations have sparked new business models that are attracting new actors to trade with different forms of flexibility. Among the new players are distribution system operators, who are expected to have an increasing role as buyers of flexibility to manage congestion. Other market players include retailers, aggregators and balancing-responsible parties who trade for portfolio optimisation purposes (Boscán and Poudineh, 2016).

[2] The reader is referred to Boscán (2016b) for a literature review on power system flexibility and its product design perspective.

The emergence of new technologies along with the greater use of ICT in the power infrastructure have enabled the provision of flexibility from many small resource providers (such as households). This is, in fact, a feature of the future decentralised power systems in which the role of small players will become more valuable to the system especially at an aggregated level. Against this background, the central question of this paper is of an *economic* nature and can be summarised in a sentence, namely: *how should the provision of flexibility be incentivised from these small resources?* Specifically, how does a utility company, system operator or an aggregator compensate owners of flexibility-enabling assets in a power system to supply flexibility? The incentive can be provided through bilateral contracts or auctions (see Boscán, 2016a) specifically designed to account for the economic properties of flexibility. However, due to presence of high transaction cost relative to the size of resource, the small resources providers cannot participate directly in an organised market and compete against each other. Therefore, this situation creates a trading environment in which "efficient" bilateral contracts are the natural method of procurement.

It is important to note that some flexibility services can also be provided in a decentralised manner. An example is demand response in a real time pricing environment when transaction costs are low. However, such a price signal often does not reach end-users and therefore in the absence of a sound price mechanism, and to address market imperfections, bilateral contracts are an alternative option to coordinate the actions of flexibility providers. Additionally, central procurement of certain flexibility services is necessary to maintain security of electricity supply. The standard bilateral contracts between system operators and the owners of operational reserves are examples of this case.

This paper fills an existing gap and contributes to the literature in three different ways. In section two, we discuss the concept of power system flexibility by explaining its relevance, the different sources from which it can be obtained, how it is traded in existing markets, and its main economic properties.

The second and the main contribution of this paper is in power system economics, as section three proposes and solves a static bilateral contracting model with bi-dimensional adverse selection, which appeals to the topic of flexibility but, more generally, to the procurement problem.[3] A buyer of flexibility - the principal - procures the two composing elements of flexibility from a seller - the agent - who has private information about the unit cost of supplying each component. While the model is presented in a bilateral setting, it can be extended to instances where competition among suppliers is feasible, but this paper focuses on trading environments where competition is not realistically viable. For example, contracts between an aggregator and a household or contracts between a DSO and a household for obtaining flexibility through the installation of smart energy management systems.[4] In such cases, transaction costs associated to market access or scale of the offer from a specific agent prevent competition to exist.

The model is presented in a sufficiently general form, which accounts for a wide range of specific functional forms, and the solution assumes non-separability[5] in both the principal's gross utility and the agent's cost function.[6] In this way, instances where separability holds are special cases. To gain

---

[3] Although flexibility is *multi*-dimensional in nature, our model is *bi*-dimensional. Such a simplification was introduced for the sake of tractability.

[4] In separate work, Boscán and Poudineh (in preparation) analyse bilateral contracts in a decentralised competitive market and multi-attribute auction models for procuring flexibility services.

[5] Non-separability means that two (or more) elements of commodity (here flexibility) must be produced or consumed together. In other words, a provider of flexibility either produces all attributes (e.g., capacity, ramp rate, duration) at each time or none of them. Similarly, a user of flexibility either consumes all of the elements together or none of them as they cannot be separated. Later in the paper we show that this non-separability feature has a great impact on the specifications of efficient contracts for flexibility services.

[6] Throughout the paper we refer to the principal's "gross utility" but a mathematically equivalent interpretation is that of a non-separable, multiple input production function where the inputs are the attributes that compose flexibility and the output is used in an internal production process. For example, a DSO that procures "capacity" and "duration" utilises both elements to produce flexibility that alleviates congestion in the network, although this output does not necessarily have a market value but a value that is relevant for the DSO's overall profitability. In the DSO's case, an alternative to procuring the composing elements of

tractability, we solve a relaxed version of the fully constrained optimisation program which gives rise to the most economically relevant situations. Within this program, we analyse five different cases, which stem from the covariance of types – which determines if an agent's efficiency in *one* dimension of flexibility can be used to predict or not its efficiency in the *second* dimension of flexibility – namely: perfect correlation, positive correlation, weak correlation and negative correlation with asymmetry towards each one of the two existing middle types.

The third contribution appears in section four of the paper, where specific functional forms to characterize the optimal contracts in simulated bilateral environments are presented. Specifically, we analyse existing thermostat-based response programs. The last part of the paper concludes.

Related economic literature

On the theoretical side, there is a considerable body of economic related research. First, it has been acknowledged by Che (1993), Parkes and Kalagnanam (2005), and Asker and Cantillon (2010) that procurement is rarely concerned exclusively with one attribute and its price. Buyers of products and services in different industries usually take quality, materials, managerial performance among other considerations into account when offering a contract.

Unlike Che (1993), Asker and Cantillon (2010) and, more recently, Li et al. (2015) depart from the uni-dimensional mechanism design paradigm and analyse procurement in environments with multi-dimensional private information. While our work coincides with theirs in the multi-dimensional procurement approach, it differs with Asker and Cantillon's (2010) in the competitiveness of the environment and with Li et al.'s (2015) in the number of agent types considered.

The second area of related literature is concerned with the multi-dimensional screening approach surveyed by Rochet and Stole (2003). Within this area of research, the paper by Armstrong and Rochet (1999) has been highly influential in our work as it presents a complete and tractable analysis of optimal contracts with four types of agents. Our model, however, generalizes their approach in a significant way as we assume non-separability, whereas they assume additively separable utility and cost functions. This is an important distinction as it introduces a relevant economic insight into the analysis, which we have termed as the "non-separability effect", closely related to the topic of non-separable externalities discussed by Davis and Whinston (1962) and Marchand and Russell (1973). The model by Dana (1993) is also related, as it coincides with ours in the number of agents considered.

From the principal's perspective, the composing elements of flexibility are never perfectly substitutable and create an externality in the sense that the marginal utility of *one* of the components always depends on the *other* component (in a bi-dimensional setting). Symmetrically, from the agent's perspective, the marginal cost to produce one of the composing elements of flexibility depends on the output level selected for the second component. Therefore, whenever the principal is more interested in a specific element (for example, duration of response), he will have to compensate the agent with a price that not only depends on the marginal cost of that element but also on the level of output of the second element (for example, ramp rate or capacity), regardless of his valuation for it. Most other models of procurement under multi-dimensional screening, e.g., Asker and Cantillon (2010), Li et al. (2015), Laffont and Martimort (2002), have avoided such complications.

---

flexibility and using it in the internal production process would be expanding the network, which may not be the most efficient alternative.

2. Power system flexibility

A distinctive feature of power systems is that they require instantaneous equilibrium between supply and demand. Traditionally, utilities have operated with fairly predictable and mature technologies. To deal with the challenges of *uncertainty* and *variability*, which aren't new, a stock of balancing services and reserves have been available to system operators to ensure that the system remains in balance second by second.

However, as decarbonisation climbs up in the policy agenda and renewable generation becomes more relevant in power systems throughout the world, increased uncertainty and variability represent greater challenges for system operation. With substantial shares of renewables, the system operator's problem is to predict fluctuations in the *net load,* which is the difference between total demand (load) and variable generation, that is: demand that must be met by *other sources* if all renewable generation is utilised.[7] This magnitude is harder to predict accurately – i.e. contains greater uncertainty– as it depends on two random variables, namely demand and renewable generation.

Variability of the net load has technical and economic impact on the overall generation base of power systems. In ideal circumstances, demand and renewable generation would be positively correlated: demand is high when renewables are available or, conversely, demand is low when renewables become scarce. If this is the case, generators face *shorter peaks*, implying fewer operating hours and lower economic compensation for existing power plants. Which, of course, raises the related question of resource adequacy: How can system reliability be ensured? How can investments in baseload power plants be incentivised if, as a consequence of the greater reliance on renewables, these receive lower compensation? This is, however, a fundamentally distinct question from that of renewable integration: *How to tackle the operational challenges implied by the greater variability imposed by renewables?* The short answer to this question is "flexibility".

When demand and renewable supply are negatively – and therefore unfavourably – correlated, the remaining generation base experiences *steeper ramp ups* and *deeper turn downs* (Katz and Cochran, 2015).[8] If renewable supply decreases together with increases in demand, system operators must dispatch generation that is able to ramp up quickly. On the contrary, if renewable supply is high when demand is low, the generation base faces deeper turn downs as they must give way for renewables to satisfy demand. In other words, operators require resources – *flexibility-enabling assets* – that modify demand or output in order to follow net load fluctuations. But the question remains: what should the owners of these assets modify in order to help the operator meet net load variations? Supply and demand must indeed be modified, but for how long, for how much and at what cost? More precisely, what are the *exact* requirements of the operator? Is it capacity? Is it duration? Is it the ramp rate? Is it the lead time? Or is it a combination of these elements?

From a *technical* perspective, Ulbig and Andersson (2012), extending the work of Marakov et al. (2009), elucidate these questions. Focusing on "individual power system units" (a synonym of flexibility-enabling assets), they propose the following *flexibility trinity* to measure flexibility:

a) Power capability P for up/down regulation (measured in MW),
b) Energy storage capability E (measured in MWh), and
c) Power ramping capability R (measured in MW/min),

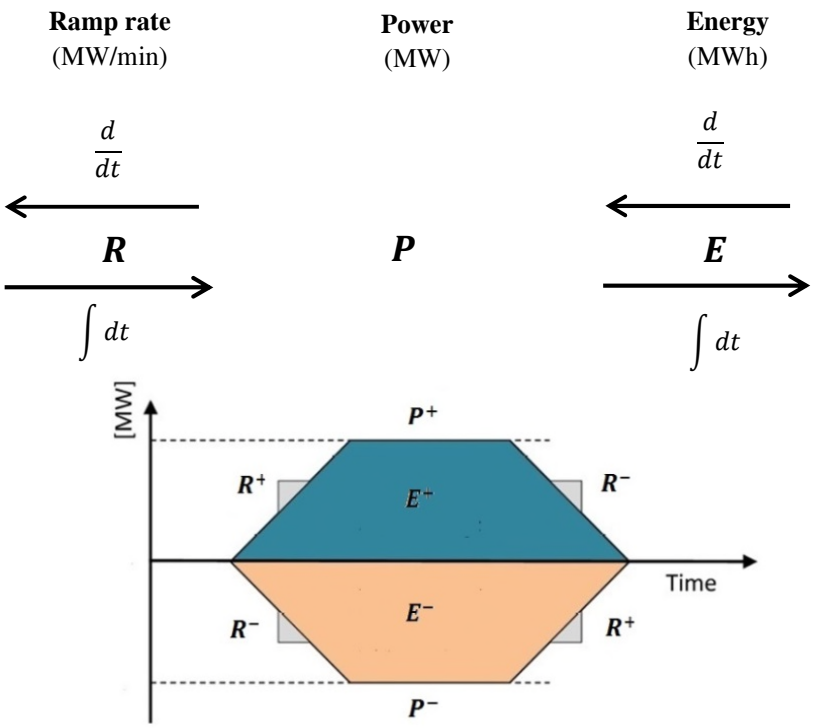
[7] The term "sources" is employed here in its widest possible sense: it could refer to generation, conventional or not, but it could also involve any change in demand that helps to keep the system in balance.
[8] Morales et al. (2013) note that this is typically the case in places like Northern Europe and Texas: renewable supply and demand are negatively correlated.

The three magnitudes are related via integration and differentiation over the time domain, as figure 1 shows. A fourth, related metric is ramping duration, D, which is defined as the ratio of power to the ramp rate, $D = \frac{P}{R}$.

While Ulbig and Andersson (2012)'s trinity could be considered incomplete by some to measure flexibility,[9] their approach highlights its fundamental *economic* characteristics:

1. *Flexibility has multiple attributes*: unlike other commodities traded in existing electricity markets, such as energy or capacity, it is not possible to measure flexibility with a single metric. This feature is economically relevant because ranking the flexibility coming from enabling assets or the agents' cost of supplying flexibility is not straightforward unless precisions regarding the multiple attributes that compose flexibility are made. Statements such as "Flexibility-enabling asset A is more flexible than flexibility-enabling asset B" are not valid, unless a clarification of what dimension of flexibility the statement refers to. Instead of a single good, it is convenient to think of flexibility as a bundle of goods.[10]

Figure 1: Flexibility in Power Systems Operation: Ramp rate (R), Power (P), Energy (E)



Source: Ulbig and Andersson (2012)

2. *Flexibility is a heterogeneous commodity*: in contrast to homogenous commodities, which have the *same* characteristics across its attributes, flexibility naturally has *different*

[9] For example, EURELECTRIC (2014) considers location, while Zhao et al. (2015) include uncertainty and cost as composing elements of flexibility.

[10] It could be argued that *any* commodity has multiple attributes as well. All pencils, for example, have height and colour. But in existing markets (e.g. the day-ahead market), electricity is measured in MWh, whereas flexibility has several dimensions.

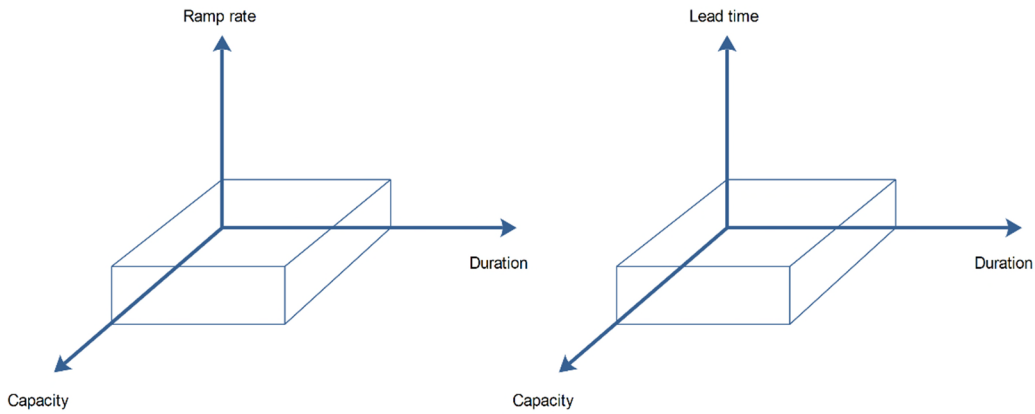
characteristics across them. A flexibility-enabling asset or agent can be efficient in one dimension but inefficient in another, creating heterogeneity. Consider, for example, a comparison between pumped-storage and nuclear power: the first has a very short lead time (can be started up from "cold" quickly) and has a steep ramp rate, whereas the second has a long lead time (cannot be started from "cold") and a much flatter ramp rate (EURELECTRIC, 2011).

3. *The elements that compose flexibility are imperfect complements*: in addition to being technically interrelated as explained before, the components of flexibility are imperfect complements because it is not possible to create any value (e.g. to the system operator) without having positive quantities of at least two flexibility components. Therefore, it is convenient to assume that the final user of flexibility has convex indifference curves (or isoquants) over the bundle of flexibility components. [11] Interestingly, the imperfect complementarity nature of flexibility can be traced back to the work of Marakov et al. (2009) who propose a "*concurrent* consideration of … capacity, ramping and … duration".

The techno-economic vision of flexibility taken in this paper is summarised in figure 2: a multi-dimensional (not necessarily three-dimensional as in figure 2), heterogeneous commodity whose components are imperfect complements. Supply-side resources (e.g. generation) are assumed to be composed – at least – by capacity, duration and ramp rate. Likewise, demand-side resources are assumed to be composed by – at least – capacity, duration and lead time (the time elapsed between agreement and delivery). From a different perspective, if each axis in the figure 2 represents the cost or disutility (when the flexibility provider is for example a household) of providing that dimension of flexibility, then an efficient flexibility contract can be interpreted as a mechanism that minimises the size of the cube shown in the figure.

Moreover, it is worth noting that the economic properties of flexibility refer both to the flexibility-enabling assets and to the agents who own the assets. However, when it comes to incentivising the provision of flexibility, there are significant distinctions between both. When considering supply-side resources like generation, owners of power plants have technical constraints to supply flexibility but will typically behave in an economically rational way, i.e. as profit maximisers. Owners of flexibility-enabling assets participating in, say, a demand response program, can be assumed to be relatively uniform regarding the assets they possess (e.g air conditioners), but their cost of provision (disutilities) should not be straightforwardly assumed to be uniform. The latter are consumers and may react to behavioural elements beyond utility maximisation.

Figure 2: Dimensions of flexibility for a) a supply side resource (left) b) a demand side resource (right)

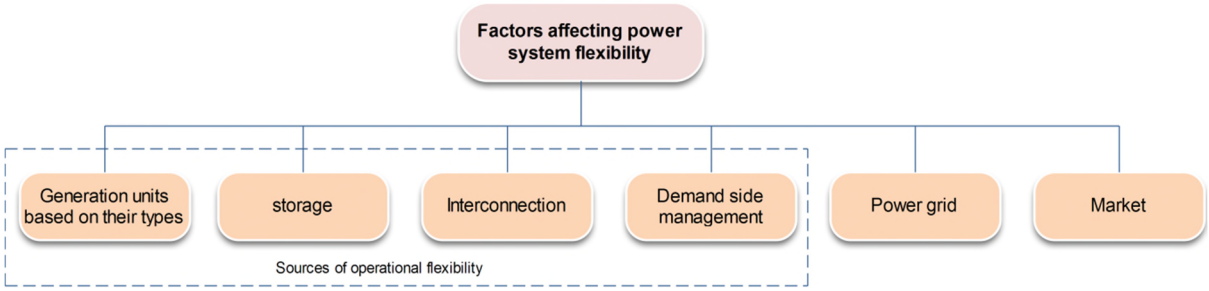


[11] Read footnote 5.

Source: Authors

There are a number of resources that enable operational flexibility in power systems, including both physical (i.e. flexibility-enabling assets) and other elements, including institutional design. The former includes generation facilities – which depending on its attributes – are positioned to supply flexibility,[12] storage (pumped hydro, thermal storage and batteries), interconnections with neighbouring networks and demand-side management. The latter comprehends elements like the power grid, which does not actually provide additional flexibility but severely limits the provision of flexibility if sufficient transmission capacity is unavailable. Market and contract design are another relevant element that can enable flexibility: in the absence of adequate trading mechanisms, market participants are unable to trade flexibility even if sufficient physical resources exist. Figure 3 summarises these elements.

Figure 3: Factors affecting flexibility



Source: Authors

It is worth noting that despite the fact that flexibility has always been needed and indeed used in the operation of power systems, it has never been traded as a distinct commodity. Usually, conventional generators have been able to provide the flexibility that the system requires and have been compensated on the basis of its energy and/or capacity component. However, as the requirement for flexibility increases in power systems, there is a need for designing a product that can be bought and sold, along with other components of electricity services such energy and capacity (on this topic, see Boscán (2016b)).

Mainly, final users of flexibility are transmission system operators (TSOs), Independent System Operators (ISOs) – who are responsible for balancing the high-voltage grid – distribution system operators (DSOs) – who are responsible for the reliable operation of distribution networks and quality of service – together with other market players who may use flexibility to meet their energy and balancing obligations. Access to flexible resources to manage network constraints leads to an effective integration of distributed energy resources, allows network companies to optimise on their network reinforcement capital investments, and improve the reliability and quality of service. It also enables other market players to optimise their energy portfolio in order to meet their energy market and balancing obligations at minimum cost by, for example, arbitrating between generation and demand response (Boscán and Poudineh, 2016).

Overall, electricity markets can be centralised or decentralised: they can be based on a tightly controlled pool, centralised exchange or bilateral contracts. Trade can include physical and/or financial obligations through forward and spot contracts and the market may allow for financial hedging. The main official market can be mandatory or optional and allow or disallow secondary markets.

[12] In a comparative study of the European generation fleet in terms of its technical flexibility, EURELECTRIC (2011) considered that warm and cold start-up times, the rapidity of load change (i.e. the ramp rate), the minimum load level and the shutdown time were the relevant attributes to describe the *technical* flexibility of power plants, also referred to in this paper, as *resource* flexibility.

Depending on the specific situation, flexibility can be procured in a competitive setting, where a single buyer incentivises suppliers to compete. Alternatively, when competition isn't feasible – as assumed in this paper – a buyer of flexibility can offer contracts to sellers who do not compete against each other. For example, a contract in which a DSO or an aggregator offers a household to provide demand side flexibility services.

Bilateral procurement contracts without competition among sellers exist across the supply chain of flexibility services for various reasons. For example, direct participation of the flexibility resource provider in an organised market may not be feasible always because of insufficient capacity size. This is important as there are diverse classes of consumers, ranging from households to large industrial units who can provide flexibility. Due to high transaction costs, small resources may face barriers to directly access the market as opposed to large industrial or commercial consumers. Aggregation provides an opportunity for small generation and demand resources to offer their flexibility in the market (Eurelectric, 2014). In this case, the aggregator can be a retailer or a third party who acts as intermediary between providers and buyers of flexibility. Additionally, in many European countries intermittent renewables are being treated as conventional generation in the sense they have the same obligations for their imbalance position and entitlement to participate in balancing market (for example, in Denmark, Finland, Estonia, Netherlands, Spain and Sweden renewable resources have full balancing responsibility). This encourages not only improved forecasting but also entry of competitive aggregators whose role is to minimise balancing risks and offer ancillary services from renewable resources. Furthermore, this provides incentives for renewables to be firmed up by, for example, entering into separate contracts with owners of flexible resources such as residential demand response and storage facilities.

The Nest Learning Thermostat is a good real-world example of how bilateral contracts for flexibility services can be utilised in the integration of renewable resources. Nest, as the manufacturer of smart thermostat technology, partners with utility companies to provide a residential demand response program. Under the so-called "Rush hour scheme", the contracted consumer's consumption (air conditioner temperature for instance) is adjusted automatically by a utility company to manage fluctuation of demand and supply. The consumers are offered a menu of contracts with different lead times (from on-demand to 24-hour notice in advance, for example), duration of adjustments in consumption (30 minutes to 4 hours for instance) and payments. The contract design problem arises because different consumers experience different disutilities for the various dimensions of flexibility they provide, and such disutilities are a privately held piece of information held by the resource provider. For example, one household may incur a high disutility for the short lead time and another household for long duration of load control. Logically, the former household prefers a contract with higher lead time but can sacrifice on load control duration, whereas the latter values more a contract with shorter load control duration. Therefore, the contracts *should* be (and currently are not necessarily being) designed in a way that each participating agent truthfully self-selects its own contract, given the presence of multidimensional information asymmetry between the buyer and sellers. The bilateral contract model of this paper focuses on this category of contractual settings in electricity markets.

3. The model

Let q and t be any two composing elements of flexibility that a buyer – the principal – procures, in exchange for a transfer T, from a seller – the agent – who has asymmetric information parameters θ_q, θ_t. Note that in contrast to the claim that flexibility is a multi-attribute commodity stated in section 2, and for the sake of tractability, the model reduces the *multi*-dimensionality of flexibility to a *bi*-dimensional setting.

In the context of flexibility, q and t can be the lead time, duration, capacity or the ramp rate of a generator, for example. More generally, beyond the context of flexibility, q and t can be any two activities or characteristics of a product delegated on an agent by a principal, who is uncertain about

the unit cost of production θ_q, θ_t. These can also be thought of as relevant parameters of any production process with multiple attributes. For example, materials, design, quality, product features (Li et al., 2015; Asker and Cantillon, 2010).

Net payoffs to the principal and agent are, respectively:

$$W = v(q,t) - T$$

and

$$U = T - c(q,t,\theta_q,\theta_t)$$

where $v(q,t)$ and $c(q,t,\theta_q,\theta_t)$ are the utility and cost function of the principal (buyer) and the agent (seller) respectively, which results from consuming and producing flexibility.[13] Besides linearity of T in W and U, which ensures the risk neutrality of the principal and the agent, few additional assumptions are made:

Assumption 1: the gross utility $v(q,t)$ of the buyer and the seller's cost $c(q,t,\theta_q,\theta_t)$ are twice differentiable functions.

Assumption 2 :$v(q,t)$ satisfies:

2a: Monotonicity, i.e. $\frac{\partial v}{\partial q}, \frac{\partial v}{\partial t} > 0$ and

2b: Concavity, i.e. $\frac{\partial^2 v}{\partial q^2}, \frac{\partial^2 v}{\partial t^2} < 0$

Assumption 3: $c(q,t,\theta_q,\theta_t)$ satisfies:

3a: Monotonicity in all its parameters, i.e., $\frac{\partial c}{\partial q}, \frac{\partial c}{\partial t}, \frac{\partial c}{\partial \theta_q}, \frac{\partial c}{\partial \theta_t} > 0,$

3b: Concavity in θ_q, θ_t, i.e. $\frac{\partial^2 c}{\partial {\theta_q}^2}, \frac{\partial^2 c}{\partial {\theta_t}^2} < 0,$

However, no specific curvature assumption of $c(q,t,\theta_q,\theta_t)$with respect to q and t is made. That is, $c(q,t,\theta_q,\theta_t)$ can be convex (if $\frac{\partial^2 c}{\partial q^2}, \frac{\partial^2 c}{\partial t^2} > 0$), concave (if $\frac{\partial^2 c}{\partial q^2}, \frac{\partial^2 c}{\partial t^2} < 0$) or both (i.e., linear).

Assumption 4: the Spence-Mirrlees (constant sign) conditions hold for U with respect to q and t:

$$\frac{\partial}{\partial \theta_q}\left(\frac{\partial U/\partial q}{\partial U/\partial T}\right), \frac{\partial}{\partial \theta_t}\left(\frac{\partial U/\partial t}{\partial U/\partial T}\right) < 0$$

These assumptions are sufficiently general to account for a wide range of specific functional forms. For example, $v(q,t)$ and $c(q,t,\theta_q,\theta_t)$ can be non-separable, weakly separable or additively separable. The solution to the model, though, assumes non-separability such that instances where separability holds are special cases.

As is standard in static bilateral contracting, nature determines the agent type which, in this case, is a pair of parameters(θ_q,θ_t) in the agent's cost function with two possible realizations. Each parameter can be either "High" (H) or "Low"(L), i.e., $\theta_q \in \{\theta_q^H, \theta_q^L\}$ and $\theta_t \in \{\theta_t^H, \theta_t^L\}$. Nature reveals the type to the agent but not to the principal, resulting thus in an adverse selection problem.

However, the distribution of types is common knowledge to both players. Namely, there are four types of agent:

[13] A mathematically equivalent interpretation of $v(q,t)$ is that of a production function with inputs q and t which combined produce an output (flexibility) employed in an internal production process.

1. Agent type LL characterised by (θ_q^L, θ_t^L) with probability p_{LL}
2. Agent type LH characterised by (θ_q^L, θ_t^H) with probability p_{LH}
3. Agent type HL characterised by (θ_q^H, θ_t^L) with probability p_{HL}
4. Agent type HH characterised by (θ_q^H, θ_t^H) with probability p_{HH}.

We refer to the LL type as the "efficient" because it is capable of offering the lowest cost product in both dimensions of flexibility. The HH type is called the "inefficient" because it has the highest cost in both activities. Likewise, LH and HL types are the "middle types", because they are efficient in one of the activities but inefficient in the other one. We refer to the LH, HL and HH together as the "less efficient" types.

3.1. Dependence of events and covariance of types

The following table summarizes the joint probability distribution of parameters θ_q and θ_t:

Table 1: Joint and marginal probability distribution of types

		θ_q		
		θ_q^L	θ_q^H	
θ_t	θ_t^L	p_{LL}	p_{HL}	$p_t^L = p_{LL} + p_{HL}$
	θ_t^H	p_{LH}	p_{HH}	$p_t^H = p_{LH} + p_{HH}$
		$p_q^L = p_{LL} + p_{LH}$	$p_q^H = p_{HL} + p_{HH}$	

Here, p_q^L, p_q^H and p_t^L, p_t^H are the marginal (unconditional) probabilities that events $\theta_q$and θ_t happen to be, respectively, "High" or "Low". Clearly, it must always be the case that $p_q^L + p_q^H = p_t^L + p_t^H = 1$.

A crucial question is how to formalise a general relationship among the probability distribution parameters to reflect the possibility of dependence or independence of events. For example, if the result of a random experiment in which the two asymmetric information parameters are drawn is the event $\theta_q = \theta_q^H$, what should the principal expect to happen: the event $\theta_t = \theta_t^H$ or $\theta_t = \theta_t^L$? Or are these two events independent of one another?

To analyze this, we use the fact that the probabilistic covariance of two events[14] $\theta_q = \theta_q^i$, $\theta_t = \theta_t^j$, $i = L, H$; $j = L, H$ is defined as:

$$cov(\theta_q = \theta_q^i, \theta_t = \theta_t^j) = p(\theta_q = \theta_q^i \cap \theta_t = \theta_t^j) - p(\theta_q = \theta_q^i)p(\theta_t = \theta_t^j)$$

which leads to the following concise expression for the *covariance of types*:[15]

$$cov = p_{LL}p_{HH} - p_{HL}p_{LH}$$

It is without loss of generality to define the following three possible cases:

[14] The formula for covariance presented in the text is based on a probabilistic view of θ_q and θ_t as events, not as random variables. There is a probability associated to the occurrence of each event. Formally: if A, B are events then $cov(A, B) = P(A \cap B) - P(A)P(B)$.

[15] The discrete set of agents considered and the expression for covariance of types coincides with the setup of Armstrong and Rochet (1999) and Dana (1993).

i. *Positive correlation* if and only if $p_{LL}p_{HH} > p_{HL}p_{LH}$
ii. *Negative correlation* if and only if $p_{LL}p_{HH} < p_{HL}p_{LH}$
iii. *Absence of correlation* if and only if $p_{LL}p_{HH} = p_{HL}p_{LH}$

3.2. Relaxing the fully constrained program

In this bi-dimensional setting, $S(q_{ij}, t_{ij}) = W_{ij} + U_{ij}$ is the total surplus derived from trading goods q_{ij} and t_{ij}. W_{ij} and U_{ij} stand for, respectively, the net payoffs to the principal and agent associated to the ij-type.

Without informational asymmetry, the principal is able to observe cost and obtains first-best levels q_{ij}^{fb} and t_{ij}^{fb} where $q_{ij}^{fb}, t_{ij}^{fb} = \arg\max_{q_{ij},t_{ij}} S(q_{ij}, t_{ij})$, such that:

$$\frac{\partial v(q_{ij}^{fb}, t_{ij})}{\partial q_{ij}} = \frac{\partial c(q_{ij}^{fb}, t_{ij}, {\theta_q}^i, {\theta_t}^j)}{\partial q_{ij}} \text{ for all } q_{ij}$$

$$\frac{\partial v(q_{ij}, t_{ij}^{fb})}{\partial t_{ij}} = \frac{\partial c(q_{ij}, t_{ij}^{fb}, {\theta_q}^i, {\theta_t}^j)}{\partial t_{ij}} \text{ for all } t_{ij}$$

for all $i = L, H$; $j = L, H$

With adverse selection, the principal's problem is to offer a menu of contracts $\{q_{ij}, t_{ij}, T_{ij}\}$ that maximizes expected surplus,

$$E(\pi) = \sum_{ij} p_{ij} S(q_{ij}, t_{ij}) - \sum_{ij} p_{ij} U_{ij} \quad (1)$$

subject to individual rationality (IR):

$$T_{ij} - c(q_{ij}, t_{ij}, \theta_q^i, \theta_t^j) = U_{ij} \geq 0 \text{ for all } i = L, H \,;\, j = L, H$$

and incentive compatibility (IC) constraints:

$$U_{ij} \geq U_{i'j'} + c\left(q_{i'j'}, t_{i'j'}, \theta_q^{i'}, \theta_t^{j'}\right) - c(q_{i'j'}, t_{i'j'}, \theta_q^i, \theta_t^j) \text{ for all pairs } ij \text{ and } i'j'$$

At this point, it is convenient to introduce the following notational convention:

$$\Delta_{ij}^{i'j'}(q_{i'j'}, t_{i'j'}) = c(q_{i'j'}, t_{i'j'}, \theta_q^{i'}, \theta_t^{j'}) - c(q_{i'j'}, t_{i'j'}, \theta_q^i, \theta_t^j)$$

where $\Delta_{ij}^{i'j'}(q_{i'j'}, t_{i'j'})$ stands for the total cost difference between the $i'j'$ and the ij type, evaluated at output levels $(q_{i'j'}, t_{i'j'})$. Note that $\Delta_{ij}^{i'j'}(q_{i'j'}, t_{i'j'})$ can be positive, negative or zero.

Using this convention, it is also possible to write:

$$\frac{\partial \Delta_{ij}^{i'j'}(q_{i'j'}, t_{i'j'})}{\partial z_{i'j'}} = \frac{\partial c(q_{i'j'}, t_{i'j'}, \theta_q^{i'}, \theta_t^{j'})}{\partial z_{i'j'}} - \frac{\partial c(q_{i'j'}, t_{i'j'}, \theta_q^i, \theta_t^j)}{\partial z_{i'j'}} \text{ for } z \in \{q, t\}$$

to denote the *difference in marginal cost* between the $i'j'$ and the ij type when producing either $q_{i'j'}$ or $t_{i'j'}$.

With this notation, IC constraints can be simplified to:

$$U_{ij} \geq U_{i'j'} + \Delta_{ij}^{i'j'}(q_{i'j'}, t_{i'j'}) \text{ for all pairs } ij \text{ and } i'j'$$

The fully constrained program consists of four IR constraints and twelve IC constraints from which the following monotonicity conditions for production levels follow (for details, see sections A.1 and A2.1 in the appendix):

$$q_{LL} \geq q_{HL}\ ,\ q_{LH} \geq q_{HH}\ , t_{LL} \geq t_{LH}\ , t_{HL} \geq t_{HH} \quad (2)$$

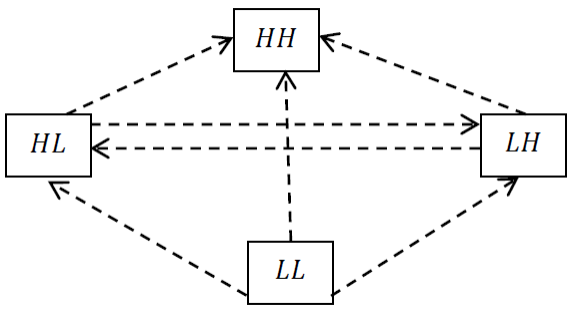
Adding local incentive constraints two by two reveals that it is optimal for the principal to incentivize the efficient types to produce more than the inefficient ones in both dimensions. However, this procedure does not clarify how the outputs of the middle types are ordered relative to each other. It can be shown that assuming a given order in the output levels of one of the middle types in one dimension, does not lead to a definite conclusion about the order in the corresponding output levels of the other dimension. For example, assuming that $q_{LH} \geq q_{HL}$ holds, t_{HL} can be equal, greater or lower than t_{LH} and still satisfy the corresponding cost inequality that results from adding IC constraints. Conversely, assuming a given order for t_{HL} and t_{LH} does not lead to a clear-cut conclusion about the order of q_{LH} and q_{HL}.

To gain tractability, it is convenient to relax the fully constrained program. To this end, economic reasoning helps identifying a number of constraints that can be ignored to construct a relaxed program, which is relevant to the extent that its solution satisfies the general, fully constrained program. First note that it is in the interest of the more efficient types to mimic the relatively less efficient agents because by doing so, the former choose to produce sub-optimal output levels while obtaining positive informational rents. In contrast, the less efficient types lack the incentives to mimic the relatively more efficient ones because this would imply incurring in an unnecessarily high cost. Therefore, it is reasonable to focus on the *upward* IC constraints only. But in the presence of a set of incompletely ordered agents, who mimics whom?

The efficient (LL) can mimic any of the three remaining agents but it is not immediately obvious which would give the LL-type the highest possible informational rent. Consequently, the efficient's optimal choice could involve the possibility of mimicking one, two or three of the less efficient types. Likewise, the middle types can choose to mimic the inefficient (HH) but they could also have incentives to mimic each other (see sections A2.2 and A2.3 in the appendix, for more details).

In consequence, there are seven relevant IC constraints that must be considered in *any* relaxed program that attempts to solve the fully constrained program:

Figure 4: Relevant IC constraints in *any* relaxed program



$$U_{LL} \geq \max\{U_{LH} + \Delta_{LL}^{LH}(q_{LH}, t_{LH}), U_{HL} + \Delta_{LL}^{HL}(q_{HL}, t_{HL}), U_{HH} + \Delta_{LL}^{HH}(q_{HH}, t_{HH})\}$$

$$U_{LH} \geq \max\{U_{HL} + \Delta_{LH}^{HL}(q_{HL}, t_{HL}), U_{HH} + \Delta_{LH}^{HH}(q_{HH}, t_{HH})\}$$

$$U_{HL} \geq \max\{U_{LH} + \Delta_{HL}^{LH}(q_{LH}, t_{LH}), U_{HH} + \Delta_{HL}^{HH}(q_{HH}, t_{HH})\}$$

This situation is reflected in figure 4, where the dashed arrows represent the *potentially* binding IC constraints of a general program relaxation. However, analysing them simultaneously introduces considerable complexity. Instead, this paper focuses on a baseline relaxed program (one of the two possible relaxed programs) which accounts for the most economically relevant cases.[16]

3.3. A baseline relaxed program

The widest range of practically relevant cases can be covered if the IC constraints of the middle type agents to mimic each other are ignored. Specifically, by assuming that:

$$U_{HL} + \Delta_{LH}^{HL}(q_{HL}, t_{HL}) < U_{HH} + \Delta_{LH}^{HH}(q_{HH}, t_{HH})$$

and

$$U_{LH} + \Delta_{HL}^{LH}(q_{LH}, t_{LH}) < U_{HH} + \Delta_{HL}^{HH}(q_{HH}, t_{HH})$$

hold, it is always optimal for the middle types to mimic the inefficient type (HH) while ignoring their counterpart's contract. That is, the baseline relaxed program assumes that the "horizontal" constraints will never bind at the optimum.

The baseline relaxed program thus reduces to maximizing the principal's expected profit (equation (1)) subject to the following five IC constraints:

$$U_{LL} \geq \max\{U_{LH} + \Delta_{LL}^{LH}(q_{LH}, t_{LH}), U_{HL} + \Delta_{LL}^{HL}(q_{HL}, t_{HL}), U_{HH} + \Delta_{LL}^{HH}(q_{HH}, t_{HH})\}$$

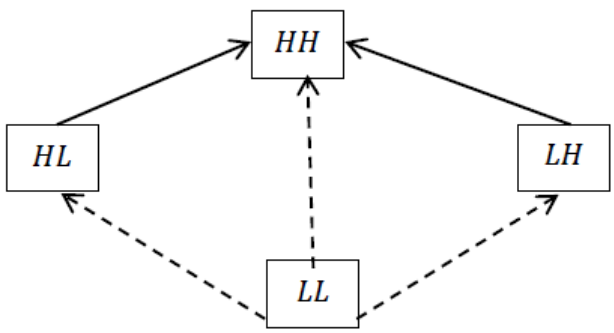
$$U_{LH} \geq U_{HH} + \Delta_{LH}^{HH}(q_{HH}, t_{HH})$$

$$U_{HL} \geq U_{HH} + \Delta_{HL}^{HH}(q_{HH}, t_{HH})$$

and the IR constraint for the inefficient: $U_{HH} \geq 0$

Figure 5 depicts the IC constraints of the baseline relaxed program. Note that the solid lines indicate optimally binding constraints, whereas the dashed lines indicate potentially binding constraints where *at least one* of them binds at the optimum. That is, the middle types' IC constraints with respect to the inefficient always bind while at least one of the inefficient's IC constraints with respect to the less efficient bind.

Figure 5: Relevant IC constraints in the baseline relaxed program



[16] The other relaxed program is the "alternative" relaxed program which accounts for situations in which it may be profitable for the middle types to mimic each other.

At the optimum, the IR constraint for the inefficient binds ($U_{HH} = 0$), but the efficient's and middle types' IC constraints also bind. Then, substituting for the latter into the efficient's IC constraints simplifies the set of constraints to:

$$U_{LL} = \max\{\Delta_{LH}^{HH}(q_{HH}, t_{HH}) + \Delta_{LL}^{LH}(q_{LH}, t_{LH}), \Delta_{HL}^{HH}(q_{HH}, t_{HH}) + \Delta_{LL}^{HL}(q_{HL}, t_{HL}), \Delta_{LL}^{HH}(q_{HH}, t_{HH})\} \quad (3)$$

$$U_{LH} = \Delta_{LH}^{HH}(q_{HH}, t_{HH})$$

$$U_{HL} = \Delta_{HL}^{HH}(q_{HH}, t_{HH})$$

The setup of the baseline relaxed program gives rise to four mutually exclusive cases, which depend on the optimally binding constraints for the efficient type. As will become clear, the emergence of these cases is closely related to the covariance of types and the difference in marginal cost between the least efficient type (HH) and the three other agent types: LL, HL and LH.

3.3.1. Solution to the baseline relaxed program[17]

The output levels produced under adverse selection are always second-best and are written as:

$$q_{ij}^{sb} = q(\cdot),\ t_{ij}^{sb} = t(\cdot) \qquad (4)$$

Here, $q(\cdot), t(\cdot)$ denote that the second-best output levels are, respectively, outputs q and t that deviate from the first-best solution by the expression in the argument. *So, the greater the argument, the greater the deviation from the first best solution is*. In contrast, if the argument is zero, then first-best and second-best output levels coincide[18]. That is, if $q_{ij}^{sb} = q(0) \Rightarrow q_{ij}^{sb} = q_{ij}^{fb}$; $t_{ij}^{sb} = t(0) \Rightarrow t_{ij}^{sb} = t_{ij}^{fb}$.

In the baseline relaxed program, the efficient's output levels always coincide with the first-best:

$$q_{LL}^{sb} = q_{LL}^{fb}\ , t_{LL}^{sb} = t_{LL}^{fb} \qquad (5)$$

but the outputs of the remaining types (HL, LH, HH) are distorted. The size of the distortion, however, will depend on each case.

Case 1: Positive correlation

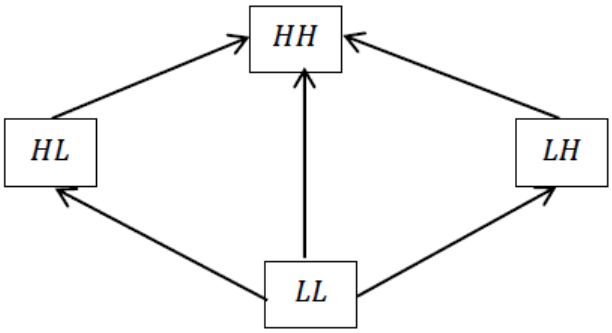
In the second case, equation (3) has no single maximum, implying that the efficient type does not have an incentive to mimic one specific agent but all the three less efficient agents. Note that perfect correlation is a subcase of case 2. The LL type does not ignore the global IC constraint with the inefficient (HH), as it would give him an equally rewarding informational rent as mimicking any of the two adjacent types, LH or HL.

Graphically, this situation is represented in figure 6, which indicates that all the IC constraints of the baseline relaxed program bind at the optimum. In particular, the LL type's IC constraints bind with respect to the three less efficient types.

[17] Technical details are outlined in Appendix, subsection A.3

[18] This happens when marginal cost equals marginal benefit and thus at the optimum point their difference is zero such that the agent produces with no distortion.

Figure 6: Case 2 (Positive correlation) in the baseline relaxed program



Intuitively, this situation arises whenever the types are positively correlated such that the efficient and inefficient types are more likely than the middle types in a given distribution. This case holds whenever the covariance of types is:

$$cov(\theta_q^i, \theta_t^j) > (1 - p_{HH})\left(p_{LH} \frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} + p_{HL} \frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} \right) - p_{LH}\, p_{HL} \text{ for } z \in \{q, t\}$$

In this case, the principal finds it optimal to incentivize the efficient to produce at first-best levels, giving this agent type a rent that equals a convex combination of the informational rents that he would obtain by mimicking the less efficient agents *simultaneously*.

Essentially, the principal distinguishes between the efficient, which is able to offer q and t at the lowest cost, and the less efficient types, who compose a rather indistinguishable group. There is a resemblance of this case with that of the unidimensional adverse selection problem, in which a single parameter determines the difference between the efficient and the inefficient. By implication, the outputs of the inefficient are distorted in a way that leads to a bunching solution, i.e., $q_{LH}^{sb} = q_{HL}^{sb} = q_{HH}^{sb}$; $t_{LH}^{sb} = t_{HL}^{sb} = t_{HH}^{sb}$.

Specifically, the deviations of output relative to the first-best levels are given by:

- For the LH, HL and HH in the q attribute:

$$q_{LH}^{sb} = q_{HL}^{sb} = q_{HH}^{sb} = q\left(\frac{p_{LH}\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}} + p_{HL}\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}} + p_{LL}\frac{\partial \Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}}}{p_{HH} + \frac{\partial \Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}}(p_{HL} + p_{LH}) - p_{HL}\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}} - p_{LH}\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}}} \right) \quad (6a)$$

- For the LH, HL and HH in the t attribute:

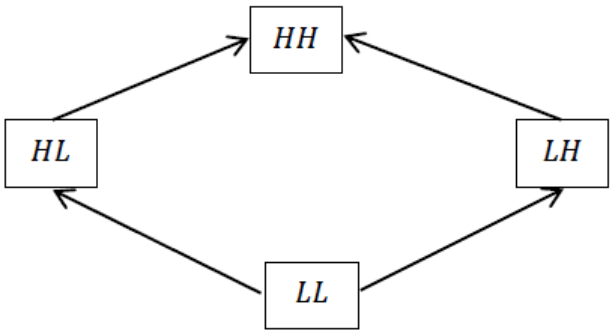
$$t_{LH}^{sb} = t_{HL}^{sb} = t_{HH}^{sb} = t\left(\frac{p_{LH}\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}} + p_{HL}\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}} + p_{LL}\frac{\partial \Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}}}{p_{HH} + \frac{\partial \Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}}(p_{HL} + p_{LH}) - p_{HL}\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}} - p_{LH}\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}}} \right) \quad (6b)$$

Case 2: Weak correlation

In the third case, the global IC constraint between the efficient (LL) and inefficient (HH) becomes irrelevant. This is because the LL type can realize a higher rent by mimicking any of the middle, adjacent types (LH or HL) than by mimicking the inefficient type. In contrast to case two, this case

emerges whenever the middle types become more likely in the distribution, such that the covariance of types is still positive or negative but closer to zero. The situation is depicted in figure 7, where the binding IC constraints are shown.

Figure 7: Case 3 (Weak correlation) in the baseline relaxed program



It is rational for the efficient to always mimic the least efficient type because by doing so he obtains a higher informational rent. To avoid the rent becoming too costly, it is in the principal's interest to make the efficient type indifferent between any of the two middle types. This means that the following equality should hold:

$$\Delta_{LH}^{HH}(q_{HH}^{sb}, t_{HH}^{sb}) + \Delta_{LL}^{LH}(q_{LH}^{sb}, t_{LH}^{sb}) = \Delta_{HL}^{HH}(q_{HH}^{sb}, t_{HH}^{sb}) + \Delta_{LL}^{HL}(q_{HL}^{sb}, t_{HL}^{sb}) \qquad (7)$$

Equation (7) depends on a real number λ, and the principal's problem reduces to finding a $0 < \lambda < 1$ for which the informational rent that the LL type would realize from mimicking any of the two middle types is equal, as in (7).

To this end, the principal distorts output levels which, in turn, depend on the degree of marginal cost symmetry between them. Depending on the assumption about the marginal cost symmetry of the middle types three possibilities emerge.

Output levels, which depend on λ, are:

- For the LH type:

$$q_{LH}^{sb} = q\left(\lambda \frac{p_{LL}}{p_{LH}}\right), t_{LH}^{sb} = t\left(\lambda \frac{p_{LL}}{p_{LH}}\right) \qquad (8a)$$

- For the HL type:

$$q_{HL}^{sb} = q\left((1-\lambda)\frac{p_{LL}}{p_{HL}}\right), t_{HL}^{sb} = t\left((1-\lambda)\frac{p_{LL}}{p_{HL}}\right) \qquad (8b)$$

- For the HH type:

$$q_{HH}^{sb} = q\left(\frac{1}{p_{HH}}\left[(\lambda\, p_{LL} + p_{LH})\frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial q_{HH}} + ((1-\lambda)\, p_{LL} + p_{HL})\frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial q_{HH}}\right]\right)$$

(8c)

$$t_{HH}^{sb} = t\left(\frac{1}{p_{HH}}\left[(\lambda\, p_{LL} + p_{LH})\frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial t_{HH}} + ((1-\lambda)\, p_{LL} + p_{HL})\frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial t_{HH}}\right]\right)$$

The first case is that middle types are equally marginally efficient, i.e., assume that:

$$\frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} = \frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} \text{ for } z \in \{q, t\}$$

holds, meaning that both middle types would realize the same informational rent by mimicking the HH type because they are equally efficient at the margin.

It is, however, too restrictive to assume cost symmetry between the middle types and even in the cases in which there is marginal cost asymmetry we show that there is $0 < \lambda < 1$ that satisfies equation (7). One possibility is that LH *type is more inefficient than the* HL *type:*

$$\frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} > \frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} \text{ for } z \in \{q, t\}$$

The other possibility is that *HL type is more inefficient than the LH type:*

$$\frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} < \frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} \text{ for } z \in \{q, t\}$$

The appendix (section A.3) contains a proof for the existence of a λ that satisfies the indifference condition in (7) in all previous three possibilities.

The sufficient condition for the existence of λ and thus a solution for case three is that the following inequality holds.

$$\Delta_{LL}^{LH}\left(q\left(\frac{p_{LL}}{p_{LH}}\right), t\left(\frac{p_{LL}}{p_{LH}}\right)\right) - \Delta_{LL}^{HL}\big(q(0), t(0)\big) < \Delta_{LH}^{HH}(q_{HH}^{sb}, t_{HH}^{sb}) - \Delta_{HL}^{HH}(q_{HH}^{sb}, t_{HH}^{sb})$$
$$< \Delta_{LL}^{LH}(q(0), t(0)) - \Delta_{LL}^{HL}\left(q\left(\frac{p_{LL}}{p_{HL}}\right), t\left(\frac{p_{LL}}{p_{HL}}\right)\right)$$

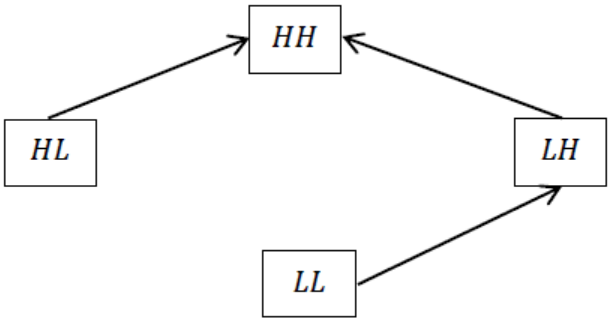
This implies that weak correlation of types case happens when there is no large asymmetry between the middle types.

Case 3: Negative correlation with asymmetry towards the LH type

This case arises whenever the distribution of types is negatively correlated because the middle types are more likely than the LL and HH type together, and contains a higher probability of observing an LH type than any other type. In this situation, the efficient will have an incentive to mimic the LH type but will, essentially, ignore the HL type. The situation is graphically illustrated in figure 8.

It is optimal for the principal to distort the LH type but not the HL type, whose output will coincide with the first-best. The middle types will continue to mimic the inefficient but because of the asymmetry, HH's output will have a higher distortion relative to the LH than to the HL.

Figure 8: Case 4 (Negative correlation with asymmetry towards the LH type)



The output levels in this case are:

- For the LH type:

$$q_{LH}^{sb} = q\left(\frac{p_{LL}}{p_{LH}}\right), t_{LH}^{sb} = t\left(\frac{p_{LL}}{p_{LH}}\right) \quad \text{(9a)}$$

- For the HL type:

$$q_{HL}^{sb} = q(0) = q_{HL}^{fb},\ t_{HL}^{sb} = t(0) = t_{HL}^{fb} \quad \text{(9b)}$$

- For the HH type:

$$q_{HH}^{sb} = q\left(\frac{1}{p_{HH}}\left[(p_{LL} + p_{LH})\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}} + p_{HL}\frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}}\right]\right)$$

(9c)

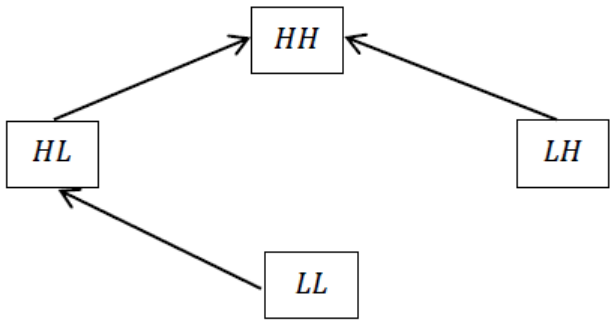
$$t_{HH}^{sb} = t\left(\frac{1}{p_{HH}}\left[(p_{LL} + p_{LH})\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}} + p_{HL}\frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}}\right]\right) \quad \text{(9d)}$$

Case 4: Negative correlation with asymmetry towards the HL type

This case is symmetric to case 3 and happens when the probability of observing the HL type is higher than any other type in the distribution. Consequently, the efficient will have an incentive to mimic the HL while ignoring the LH type. In this situation, shown in figure 9, the principal distorts HL type's output but not LH's, who produces first-best levels.

The middle types will continue to mimic the inefficient but because of the asymmetry, HH's output will have a higher distortion relative to the HL than to the LH.

Figure 9: Case 5: (Negative correlation with asymmetry towards the HL type)



The output levels in this case are:

- For the LH type:

$$q_{LH}^{sb} = q(0) = q_{LH}^{fb}, t_{LH}^{sb} = t(0) = q_{LH}^{fb} \quad (10a)$$

- For the HL type:

$$q_{HL}^{sb} = q\left(\frac{p_{LL}}{p_{HL}}\right), t_{HL}^{sb} = t\left(\frac{p_{LL}}{p_{HL}}\right) \quad (10b)$$

- For the HH type:

$$q_{HH}^{sb} = q\left(\frac{1}{p_{HH}}\left[p_{LH}\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}} + (p_{LL} + p_{HL})\frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}}\right]\right)$$

(10c)

$$t_{HH}^{sb} = q\left(\frac{1}{p_{HH}}\left[p_{LH}\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}} + (p_{LL} + p_{HL})\frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}}\right]\right)$$

4. Simulating bilateral flexibility-enabling contracts

In this section, we emphasise on the applications of the model and delve into its concrete implications for policy and business decision makers. We find inspiration in existing thermostat-based demand response programs in which utilities incentivise their customers to modify their consumption during peak hours or when the system reliability is at stake. Relying on automation, customers allow utilities to automatically reduce their air conditioning (during summer) or electric heating (during winter) consumption in exchange for payments. Some companies pay customers for each season in which the customer enrols or give a rebate on the device. Others give a flat credit on the customer's electricity bill, while others pay per peak hour. If smart metering technology is available, companies will compare actual vs. typical consumption and reward them accordingly.

While an important feature of this approach is that it reduces the customers' transaction cost to act flexibly – a relevant barrier to successfully achieving price responsiveness – a demand response programs cannot be based on a "representative agent" approach in which customers do not differ from one another. We claim that designing "efficient" contracts based on this kind of approach is not possible given that suppliers naturally differ in their cost (or disutility) of provision across the different dimensions of flexibility. This is true even when, for example, consumers have identical flexibility enabling assets (e.g., similar air conditioners).

In contrast, we take a *normative* approach to illustrate how the multi-dimensional adverse selection model discussed in this paper can be employed to design bilateral flexibility-enabling contracts that ensure economic efficiency. The main ingredient required to apply the proposed contract design approach is information regarding the distribution of types and the hidden unit cost parameters in the suppliers' cost functions. A key question that follows is thus how to elicit the relevant information from flexibility suppliers.

For illustration purposes, consider the case of a utility company – the principal – that seeks to procure flexibility through a demand response program from its customer base – the agents – in an area where electric heating or air conditioning is widely used (in practice the programme can include all types of flexible loads such as washing machine, electricity vehicles, among others). The company aims at entering into bilateral contracts with customers who aren't competing against each other and thus designs a menu of contracts $\{q_{ij}, t_{ij}, T_{ij}\}$into which suppliers self-select.. In the menu of contract q_{ij} is the capacity of response (e.g., measured in kW but it is closely correlated with the temperature

of air conditioner or electric heater), t_{ij} is duration that the control of the flexibility-enabling asset is surrounded to the utility company (e.g., measured in hours) and T_{ij} is the payment to the consumer in exchange for proving flexibility (e.g., measured in dollar or any other unit of money). Subscript i and j refer to the type of resource provider in terms of its cost efficiency at each dimensions (capacity and duration) which can be either low cost (L) or high cost (H). This creates four types of consumers LL, HL, LH, HH which their descriptions are presented in Table 2.

Table 2: type of customers and their descriptions

Type of consumer	Description
LL	Experiences a low disutility with regard to change in both the air conditioner temperature and duration of surrounding the control of air conditioner to the utility company.
HL	Experiences a high disutility with respect to change in temperature but a low disutility with respect to duration of control.
LH	Experiences a low disutility with respect to change in temperature but a high disutility with respect to duration of control.
HH	Experiences a high disutility with respect to both change in temperature and duration of control.

In the absence of a Smart Grid infrastructure, which would allow the customer base anonymously revealing their type, an alternative approach would involve conducting experiments among the utility's customer base in which the emphasis is placed on eliciting consumer's preferences (not the technical flexibility of the asset).

Regarding the values of hidden cost parameters (i.e., marginal disutility that customer experience with respect to change in temperature and duration of surrounding the control to the utility company) in the agents' cost functions, these could be estimated through dedicated empirical studies, laboratory experiments or a combination of both. A relatively simple, yet suitable approach would involve conducting an experiment with randomly chosen subjects from the customer base who have identified themselves as LL, HL, LH, HH in the survey.

In what follows, however, we limit ourselves to presenting results that reflect the *simulated* application of the previously described approach. To this end, we take the following steps:

1. *Generate simulation data for survey results to determine the distribution of types in the customer base:* Assuming a sufficiently large sample and an equally likely outcome coming from a uniform distribution for the binary outcome (i.e., High or Low) that represents the disutility which consumers experience with respect to temperature variation and duration of change, we follow a Monte Carlo approach to generate a discrete, joint probability distribution of types (p_{LL}, p_{LH} p_{HL}, and p_{HH} which shows probability of observing each type of agent). Using the Microsoft Excel add-ins developed by Myerson (2005), it is a straightforward process to generate this data in a spreadsheet. In our model, the information about distribution of types is concentrated in one parameter named covariance of types. Covariance of type (cov) show how different dimensions of flexibility are correlated across customer base.

2. *Determine the optimal contracts under the four different cases that emerge from the baseline relaxed program (section 3.3) assuming, respectively* $v(q,t) = A{q_{ij}}^{\alpha}{t_{ij}}^{1-\alpha}$ and $c(q,t,\theta_q,\theta_t) = ({\theta_q}^{i})^{\beta}({\theta_t}^{j})^{1-\beta}(qt)^{\gamma}$ as the principal's gross utility and the agent's cost function. With this specification, net payoffs to the principal and agent are $W = A{q_{ij}}^{\alpha}{t_{ij}}^{1-\alpha} - T$

$U = T - (\theta_q{}^i)^\beta (\theta_t{}^j)^{1-\beta} (q_{ij} t_{ij})^\gamma$. Note that $v(q,t)$ and $c(q,t,\theta_q,\theta_t)$ comply with the assumptions of the model introduced in section 3. First, both are twice differentiable functions. Second, $v(q,t)$ is always monotone and concave, given that $0 < \alpha < 1$. Third, $c(q,t,\theta_q,\theta_t)$ is monotone in all its parameters and concave in θ_q and θ_t, which is guaranteed because $0 < \beta < 1$. Furthermore, the value assigned to γ determines the convexity, concavity or linearity of $c(q,t,\theta_q,\theta_t)$ with respect to q and t. If $\gamma < 1$, the cost function is concave and exhibits *economies of scale*. In contrast, if $\gamma > 1$, the cost function is convex and has *diseconomies of scale*. But if $\gamma = 1$, the cost function is linear. Section A4 in the appendix contains complete technical details of the equations and numerical procedures to compute the simulation results. Section A6 in the appendix contains MATLAB script files used in the computation.

4.1 Results

Before proceeding into setting the value of cost and utility function parameters a caveat worth noting is that in the absence of actual data to feed into the model, the parameters chosen for simulation are arbitrary. Since the simulation is just for illustration of the way our model works therefore, we do not try to give a technical interpretation to the numerical results. However, the results are of sufficient information to establish the point we are trying to make about designing efficient contracts for flexibility services. If actual data becomes available, a more realistic set of results can be produced.

In all simulations, we set the following parameter values: $A = 5$, $\alpha = 0.5$ in $v(q,t)$ and $\beta = 0.5$, $\gamma = 0.6$ in $c(q,t,\theta_q,\theta_t)$. Further, we assume that $\theta_q \in \{\theta_q{}^L = 2, \theta_q{}^H = 4\}$ and that $\theta_t \in \{\theta_t{}^L = 4, \theta_t{}^H = 8\}$, giving rise to four agent types, as summarised in the following table:

Table 3: Cost parameters for the four agent types

	θ_q	θ_t
LL	2	4
LH	2	8
HL	4	4
HH	4	8

That is: the efficient (LL) type experiences the lowest discomfort in both the q and the t dimensions, the inefficient (HH) experiences the highest discomfort in both dimensions, while the middle types (LH and HL) are efficient in only one of the dimensions.

The principal procures one dimension of flexibility, q, which in our case stands for capacity (correlated with change in temperature), for a fixed quantity of second dimension (t) , which stands for duration. The desired capacity is obtained through the change in temperature of air conditioner or electric heating at consumers' premises. In all cases that follow, and for simplicity alone, we assume three different possible values for duration of control, $t_{ij} = 2, 4, 6$ hours. That is: the principal designs an optimal menu of contracts to procure capacity q, given three different options for duration.

The company announces that it will reward provision of capacity, i.e. modifying of household temperature, during 2, 4 or 6 hours, in exchange for a payment T_{ij} and the customer who enters into contact with company must provide the specified amount. Each agent type is expected to (truthfully) self-select into one of the contracts. Under the baseline relaxed program, the incentive compatibility (IC) constraints between the middle types (LH and HL) and the inefficient (HH) always bind at the optimum (see Figure 5), but the correlation of types determines which of the efficient's IC constraints bind at the optimum, as depicted in figures 6 to 9.

The details of the results are presented in cases 1 to 4 below but they can be broadly summarized as follows. The optimal menu of contract for flexibility services have the following properties: (i) the most

efficient type (LL) always receives the first best contract (i.e., undistorted) and the least efficient type' contract (HH) is always the second best (i.e., distorted) (ii) the middle type contracts are the second best except when there is negative correlation with asymmetry in which case one of the middle type receives the first best contract (iii) under most distributions assumed (except when there is strong positive correlation) the least efficient type is shut down in the sense that in practice it provides no flexibility and receives no compensation (in theory the production and compensation for this type are not zero but extremely low). This is likely due to cost parameters assumed in Table 3 in which the marginal cost (disutility) of production (θ_q, θ_t), for the least efficient type is assumed to be twice of the most efficient one. This makes provision of flexibility by the least efficient type very costly to principal.

Case 1: Positive correlation

In a simple language positive correlation means that having information about one dimension of flexibility of a customer we can deduce about its efficiency in other dimension. That is if we randomly select a customer and observe that it is efficient in q dimension it is highly probable that it is also efficient in t dimension (LL). Similarly, if it is inefficient in q dimension it is highly probable that is also inefficient in t dimension (HH). Mathematically, positive correlation happens when the customers are distributed in a way that the product of probabilities of being efficient in both dimensions (LL) and inefficient in both dimensions (HH) is strictly higher than the product of probabilities of being efficient in one dimension and inefficient in the other dimension (HL or LH). This means $p_{LL}p_{HH} > p_{HL}p_{LH}$ and therefore covariance of types is positive ($cov = p_{LL}p_{HH} - p_{HL}p_{LH} > 0$). To illustrate this case, consider the following distributions, with high probability for LL and HH types:

Distribution 1A: where $p_{LL} = 0.5, p_{LH} = 0, p_{HL} = 0, p_{HH} = 0.5, cov = 0.25$

Table 5 contains the menu of contracts for this distribution of types. In table 5, each row contains the menu of contracts for duration of $t_{ij} = 2,4,6$ hours respectively, and each column contains the optimal menu of contracts for each of the four agent types. The fifth column contains the sum of capacity procured, given each duration. For example, when the duration is 2, it is optimal for the principal to procure 24.07 units of q from the efficient type in exchange for a transfer $T_{LL} = 28.95$. The inefficient type (HH), in contrast, produces an output of 4.8E-04 units, which is virtually zero, in exchange for a transfer $T_{HH} = 0.08$.[19] The rows given a duration of 4 and 6 hours can be equivalently interpreted: note that as duration increases, the capacity procured decreases. This is the consequence of the specified $v(q,t)$ in which both inputs are imperfect complements.

In the next two examples for this case the probability of observing middle types (LH and HL) increases as that of the inefficient (HH) and efficient (LL) decrease. The results of for these distributions have been presented in Tables 6 and 7.

In all aforementioned distributions, the three incentive compatibility constraints of the efficient type bind at the optimum, as in figure 6 in Section 3-i.e., for the three different duration considered, the efficient is simultaneously indifferent between his contract, the inefficient's (HH), and the two middle types (HL and LH). The details of calculations of incentive compatibility constraints are presented in appendix A5.

Two points needs to be noted in the case of positive correlation. First, as seen in the theoretical model in Section 3, under positive correlation there is a bunching for the second best contacts offered to the less efficient types, while the efficient type produces at first best levels. This means that although there are four possible types of agents the principal bunch middle type contract with that of inefficient type and thus offer them the same contract. This can be readily confirmed from the simulation results presented in Tables 5,6 and 7. Such a result is true under any alternative distributions that exhibit positive correlation.

[19] The inefficient produces an arbitrarily low amount, which is very close to zero. Actually, it is correct to assume that the inefficient does not produce any amount.

Note, in the results for distributions 1B and 1C, that as the middle types become more likely in the distribution of types, the rent that the principal gives to the efficient becomes costlier. In consequence, the expected profit decreases. Furthermore, there is a relevant qualitative characteristic of the positive correlation case that sets it apart from the remaining cases: first, the principal determines the output level of the inefficient (HH) type and, accordingly, determines the middle types' (LH and HL) output levels, such that the efficient stays indifferent among the three less efficient types' contracts. This explains that even within the same type distribution, the optimal contract differs from one another: given a fixed t_{HH}, the principal determines q_{HH} which, in turn, determines the (q_{LH}, t_{LH}) and (q_{HL}, t_{HL}) output pairs.

Additionally, when there is strong positive correlation, the optimal menu of contract, in practice, leads to shutdown of less efficient types. This can be seen from Table 5 where LH, HL and HH virtually produce nothing and receive no compensation. The sum of q_{ij} for each duration is almost equal to production level of the most efficient type. However, as seen from Tables 6 and 7 when probability of middle type increases the optimal contract involves a non-zero level of production for less efficient types.

Table 5: Menu of contracts under case 1 and type distribution 1A

$\{q_{LL}, t_{LL}, T_{LL}\}$	$\{q_{LH}, t_{LH}, T_{LH}\}$	$\{q_{HL}, t_{HL}, T_{HL}\}$	$\{q_{HH}, t_{HH}, T_{HH}\}$	***Sum of q_{ij}***
{24.07, 2 ,28.95}	{0.00, 2 ,0.08}	{0.00, 2 ,0.08}	{4.8E-04, 2 ,0.08}	24.07
{12.03, 4 , 28.95 }	{0.00, 4 ,0.08}	{0.00, 4 ,0.08}	{2.4E-04, 4 ,0.08}	12.04
{8.02, 6 , 28.95 }	{0.00, 6 ,0.08}	{0.00, 6 ,0.08}	{1.35E-04, 6 ,0.08}	8.03

Distribution 1B: where $p_{LL} = 0.49$, $p_{LH} = 0.01$, $p_{HL} = 0.01$, $p_{HH} = 0.49$, $cov = 0.24$

Table 6: Menu of contracts under case 1 and type distribution 1B

$\{q_{LL}, t_{LL}, T_{LL}\}$	$\{q_{LH}, t_{LH}, T_{LH}\}$	$\{q_{HL}, t_{HL}, T_{HL}\}$	$\{q_{HH}, t_{HH}, T_{HH}\}$	***Sum of q_{ij}***
{24.07, 2, 28.99}	{0.0014, 2 ,0.16}	{0.0014, 2 ,0.16}	{0.0014, 2 ,0.16}	24.07
{12.03, 4 , 29.04 }	{0.0015, 4 ,0.26}	{0.0015, 4 ,0.26}	{0.0015, 4 ,0.26}	12.04
{8.02, 6 , 29.10 }	{0.002, 6 ,0.38}	{0.002, 6 ,0.38}	{0.002, 6 ,0.38}	8.03

Distribution 1C: where $p_{LL} = 0.45$, $p_{LH} = 0.05$, $p_{HL} = 0.05$, $p_{HH} = 0.45$, $cov = 0.20$

Table 7: Menu of contracts under case 1 and type distribution 1C

$\{q_{LL}, t_{LL}, T_{LL}\}$	$\{q_{LH}, t_{LH}, T_{LH}\}$	$\{q_{HL}, t_{HL}, T_{HL}\}$	$\{q_{HH}, t_{HH}, T_{HH}\}$	***Sum of q_{ij}***
{24.07, 2, 29.30}	{0.02, 2 ,0.79}	{0.02, 2 ,0.79}	{0.02, 2 ,0.79}	24.12
{12.03, 4 , 29.83}	{0.04,4,1.84}	{0.04,4,1.84}	{0.04,4,1.84}	12.15
{8.02, 6 , 30.46 }	{0.06, 6 , 3.10 }	{0.06, 6 ,3.10 }	{0.06, 6 , 3.10 }	8.21

Case 2: Weak correlation

When there is weak correlation, the middle types (LH and HL) become more likely in the distribution of types, while the efficient (LL) and inefficient (HH) become less likely. Therefore, types are weakly correlated, i.e. the covariance of types is closer to zero but still positive, and the incentive compatibility (IC) constraint between the efficient (LL) and inefficient (HH) does not bind. Instead, the efficient finds it more profitable to mimic any of the middle types. Therefore, it is in the principal's best interest to make the LL type indifferent between mimicking any of the two middle types (LH or HL). This is ensured by finding a real number $0 < \lambda < 1$ that leads to output levels for which equation (7) in Section 3 holds, which we obtain numerically. Likewise, IC constraints behave as in figure 7 (see Section 3).

The following two type distributions illustrate the relevant features of weak correlation:

Distribution 2A where $p_{LL} = 0.3$, $p_{LH} = 0.25$, $p_{HL} = 0.19$, $p_{HH} = 0.26$, $cov = 0.03$. For $\lambda = 0.57$, the the output levels that satisfy equation (7) have been presented in Table 8.

Distribution 2B where where $p_{LL} = 0.28$, $p_{LH} = 0.22$, $p_{HL} = 0.22$, $p_{HH} = 0.28$, $cov = 0.03$. For $\lambda = 0.50$, the output levels that satisfy equation (7) have presented in Table 9.

As seen from Table 8 and 9, the middle types produce more flexibility compare to the previous case but the least efficient type (HH) almost does not produce. The share of middle type output increases when its probability of being observed increases. Unlike the positive correlation case, in which the principal determines the middle types' output levels in accordance with the inefficient type's (HH) output, under weak correlation it is optimal for the principal to determine output levels that will satisfy equality (7), which ignores the global IC constraint between the HH and LL. If a number λ satisfying this equality exists, then the same number will satisfy it for different levels of the fixed t_{ij}. This explains that the *same* contracts are optimal for different duration and that the efficient's rent and expected profit remain unchanged.

For both 2A and 2B distributions, the efficient type is indifferent between his contract and that of any of the two middle types. However, as predicted in theory presented in Section 3 the efficient's incentive compatibility constraint with respect to the inefficient does not bind. The details of calculations for incentive compatibility check can be found in Appendix A5.

Table 8: Menu of contracts under type distribution 2 A

$\{q_{LL}, t_{LL}, T_{LL}\}$	$\{q_{LH}, t_{LH}, T_{LH}\}$	$\{q_{HL}, t_{HL}, T_{HL}\}$	$\{q_{HH}, t_{HH}, T_{HH}\}$	***Sum of*** q_{ij}
{24.07, 2, 29.42}	{0.12, 2 ,1.72}	{0.12, 2 ,1.72}	{5.47E-05, 2 ,0.02}	24.31
{12.03, 4 , 29.42 }	{0.06,4,1.72}	{0.06,4,1.72}	{2.74E-05,4,0.02}	12.15
{8.02, 6 , 29.42}	{0.05, 6 ,1.72 }	{0.05, 6 ,1.72 }	{1.82E-05, 6 ,0.02}	8.10

Table 9: Menu of contracts under type distribution 2B

$\{q_{LL}, t_{LL}, T_{LL}\}$	$\{q_{LH}, t_{LH}, T_{LH}\}$	$\{q_{HL}, t_{HL}, T_{HL}\}$	$\{q_{HH}, t_{HH}, T_{HH}\}$	***Sum of*** q_{ij}
{24.07, 2, 29.45}	{0.14, 2 ,1.84}	{0.14, 2 ,1.84}	{8.57E-05, 2, 0.03}	24.34
{12.03, 4 , 29.45 }	{0.07,4,1.84}	{0.07,4,1.84}	{4.28E-05,4, 0.03}	12.17
{8.02, 6 , 29.45 }	{0.05, 6 ,1.84 }	{0.05, 6 ,1.84 }	{2.86E-05, 6, 0.03}	8.11

Case 3: Negative correlation with asymmetry towards the LH type

If there is negative correlation, because both of the middle types (LH and HL) are more likely than the efficient (LL) and inefficient (HH) taken together, and the LH type has a greater probability of being observed in the distribution than any other type, then the general model is under case 3. In consequence, at the optimum, the efficient's only binding constraint is the one with respect to the LH type (see figure 8 in Section 3). To avoid the efficient's rent becoming too costly, the principal distorts the LH type's and the HH's output level, while letting the HL type produce at first-best levels.

The following two distributions illustrate the features of case 3:

Distribution 3A where $p_{LL} = 0.125$, $p_{LH} = 0.5$, $p_{HL} = 0.25$, $p_{HH} = 0.125$, $cov = -0.11$

Table 10: Menu of contracts under type distribution 3A

$\{q_{LL}, t_{LL}, T_{LL}\}$	$\{q_{LH}, t_{LH}, T_{LH}\}$	$\{q_{HL}, t_{HL}, T_{HL}\}$	$\{q_{HH}, t_{HH}, T_{HH}\}$	***Sum of q_{ij}***
{24.07, 2, 29.45}	{0.37, 2 ,1.84}	{0.75, 2 ,1.84}	{3.37E-07, 2, 0.00}	25.19
{12.03, 4 , 29.30 }	{0.19,4,3.34}	{0.38,4,5.11}	{1.68E-07,4, 0.00}	12.59
{8.02, 6 , 29.89 }	{0.12, 6 ,3.34 }	{0.25, 6 ,5.11 }	{1.12E-07, 6, 0.00}	8.40

Distribution 3B where where $p_{LL} = 0.08$, $p_{LH} = 0.74$, $p_{HL} = 0.1$, $p_{HH} = 0.08$, $cov = -0.11$

Table 11: Menu of contracts under type distribution 3B

$\{q_{LL}, t_{LL}, T_{LL}\}$	$\{q_{LH}, t_{LH}, T_{LH}\}$	$\{q_{HL}, t_{HL}, T_{HL}\}$	$\{q_{HH}, t_{HH}, T_{HH}\}$	***Sum of q_{ij}***
{24.07, 2, 30.15}	{0.55, 2, 4.24}	{0.75, 2 , 5.11}	{9.2E-09, 2 ,0.00}	25.37
{12.03, 4 , 30.15 }	{0.28,4, 4.24}	{0.38,4,5.11}	{4.6E-09,4,0.00}	12.68
{8.02, 6 , 30.15}	{0.18, 6 , 4.24 }	{0.25, 6 ,5.11 }	{3E-09, 6 ,0.00}	8.46

Unlike all previous cases in which only the most efficient type receives the first best contract and the rest are offered the second best contract, in the case of negative correlation with asymmetry, one of the middle type receives the first best contract-i.e., the middle type that is least probable to be observed (here HL type). This situation creates a sharp distinction between bilateral contracts under unidimensional and multidimensional information asymmetry in the sense that when there is more than one dimension it is possible to have first best contract for a less efficient type even under information asymmetry. In fact the low probability of HL type averts the need to distort its contract.

In both distribution 3A and 3B, the efficient type is indifferent between his contract and that of the LH type. With the closed form solutions of section A4 in the appendix, it is straightforward to compute output levels for case 3 by letting $\lambda_1 = 1$, $\lambda_2 = \lambda_3 = 0$.

Case 4: Negative correlation with asymmetry towards the HL type

Symmetrically to case 3, if both middle types (LH and HL) are more likely than the efficient (LL) and inefficient (HH) together, and the HL type has a greater probability of being observed in the distribution than any other type, then the general model is under case 4. At the optimum, the efficient's only binding constraint is the one with respect to the HL type (see figure 9 in Section 3), and the principal optimises by distorting both the HL type's and the HH's output level, while letting the LH type produce at first-best levels.

Consider the following two type distributions, which satisfy case 4:

Distribution 4A, where $p_{LL} = 0.125$, $p_{LH} = 0.25$, $p_{HL} = 0.5$, $p_{HH} = 0.125$, $cov = -0.11$

Table 12: Menu of contracts under type distribution 4A

$\{q_{LL}, t_{LL}, T_{LL}\}$	$\{q_{LH}, t_{LH}, T_{LH}\}$	$\{q_{HL}, t_{HL}, T_{HL}\}$	$\{q_{HH}, t_{HH}, T_{HH}\}$	***Sum of q_{ij}***
{24.07, 2, 29.89}	{0.75, 2, 5.11}	{0.37, 2 , 3.34}	{3.3E-07, 2 ,0.00}	25.19
{12.03, 4 , 29.89 }	{0.38,4, 5.11}	{0.19,4,3.34}	{1.6E-07,4,0.00}	12.59
{8.02, 6 , 29.89}	{0.25, 6 , 5.11 }	{0.12, 6 ,3.34 }	{1.12E-07, 6 ,0.00}	8.40

Distribution 4B, where $p_{LL} = 0.04$, $p_{LH} = 0.05$, $p_{HL} = 0.87$, $p_{HH} = 0.04$, $cov = -0.04$

Table 13: Menu of contracts under type distribution 4B

$\{q_{LL}, t_{LL}, T_{LL}\}$	$\{q_{LH}, t_{LH}, T_{LH}\}$	$\{q_{HL}, t_{HL}, T_{HL}\}$	$\{q_{HH}, t_{HH}, T_{HH}\}$	***Sum of*** q_{ij}
{24.07, 2, 30.29}	{0.75, 2, 5.11}	{0.66, 2 , 4.72}	{2.1E-11, 2 ,0.00}	25.48
{12.03, 4 , 30.29 }	{0.38,4, 5.11}	{0.33,4, 4.72}	{1.05E-11,4,0.00}	12.74
{8.02, 6 , 30.29}	{0.25, 6 , 5.11 }	{0.22, 6 , 4.72}	{7E-12, 6 ,0.00}	8.49

The results in this case is symmetric of case 3. The most efficient type and one of the middle types (here LH) receives the first best contract and the rest receive the second best. As under most distributions assumed previously the least efficient type (HH) contract in practice leads to no production and compensation.

In terms of incentive compatibility, the efficient type is indifferent between his contract and that of the HL type under both distributions 4A and 4B. The results for incentive compatibility check have been presented in Appendix A5. Using the formulas derived in section A4 in the appendix, it is straightforward to compute output levels for case 5 by letting $\lambda_2 = 1, \lambda_1 = \lambda_3 = 0$.

5. Conclusions

Given the increased reliance on renewables taking place in many power systems throughout the world, incentivising the provision of flexibility is becoming a priority for system operators, policy and business decision makers alike. The technological improvement in digital communication has resulted in emergence of new players in the electricity market with flexible loads (e.g., households with electric vehicles, electric heater, air conditioner, solar PV with storage) which are valuable sources of flexibility for the system at an aggregated level.

On the other hand, flexibility is inherently *multi-dimensional* as both the cost of producing it and the utility (or output if one considers the equivalent mathematical interpretation of a production function) that is derived from utilizing it depends on more than one factor. Capacity, ramp rate, duration and/or lead time are among the many elements that describe flexibility. Because of this, different flexibility-enabling resources possess differing levels of efficiency, implying that flexibility is not a homogenous commodity. Additionally, an equally relevant economic property of power system flexibility is that its composing elements (i.e., capacity, ramp rate, duration) are best understood as *imperfect complements*: the cost of production and the utility derived from it are always *non-separable*. Not only do buyers and sellers of flexibility have utility and cost functions that depend on more than one factor but these factors enter into these functions multiplicatively.

Therefore, as flexibility differs significantly from commodities traditionally traded in existing electricity markets, such as energy or capacity, correctly accounting for its economic properties is essential to create the incentives to enable it in electricity systems. More specifically, designing a market for flexibility services needs to be compatible with properties of the traded commodity, an important point that has so far remained unaddressed in the existing literature of power system economics. For example, where competition among flexibility providers are possible, a multi-attribute auction is needed to procure flexibility in an efficient manner. The multi-attribute auction is an allocation mechanism in which more than one feature of the commodity is valued (e.g., MW, MW/min and emission performance). Therefore, it allows the principal to incentivise, for example, capacity, flexibility and emission performance simultaneously in a single auction. In situations where competition is not feasible, multi-dimensional bilateral contracts are the alternative method of procurement.

The bilateral contracts are specifically relevant in the case of emerging small resources as, due to high transaction cost relative to the size of resource, these resources cannot directly participate in an organised market and compete. Taking the aforementioned considerations as guiding principles rather than mere theoretical considerations, this paper has taken a contract design perspective to incentivise the bilateral exchange of flexibility from small resources, already happening in some electricity markets in the form of demand response programs and ancillary services. More transactions of the kind are expected to play a greater role in the not so distant future as renewables grow further as an alternative to fossil fuels, and emerging business models that enable power system flexibility – led by technological innovation – consolidate (Boscán and Poudineh, 2016).

The paper introduces an adverse selection model of procurement with multi-dimensional types. The model innovates by presenting solutions to the non-separable case in a baseline relaxed program that accounts for the vast majority of economically relevant cases. The results of this model provide important insights on designing efficient contract for flexibility services which can be utilised in a, for example, demand response programme.

First, the results show that optimal contract for flexibility crucially depends on the way that flexibility providers are distributed-e.g., if we assume flexibility with two dimensions (capacity and duration) which is procured from a group of flexibility providers then distribution of type is a set of probabilities that show what percentage of group is efficient (low cost) in both dimensions (LL) or inefficient (high cost) in both dimensions (HH) or efficient in one dimension and inefficient in the other dimension (LH or HL). In the model presented in the paper, the information about distribution of types is concentrated in one parameter-i.e., covariance of type. The information about distribution of types can be obtained through indirectly observing the consumers' behavior in a smart grid environment or conducting experiments in which the emphasis is placed on eliciting consumer's preferences.

Depending on the covariance of types – whether efficiency in one dimension is independent or not of the other dimension – four mutually exclusive cases arises:

If there is *positive correlation*, meaning most flexibility providers are either low cost (LL) or high costs (HH) in both dimension (middle types are a small percentage of group when there is positive correlation), the efficient type does have incentive to mimic the three less efficient types (HH, H and HL). The optimal contract in this case has two properties: (a) it incentivises the efficient type (LL) to produce at first-best levels (no distortion as if there is no information asymmetry), giving this agent type a rent that equals a convex combination of the informational rents that he would obtain by mimicking the less efficient agents *simultaneously* (b) the other three types are offered the same contract and produce at the second best level (distorted because of information asymmetry). The bunching of the contracts for three less efficient types happens because probability of observing middle types (LH or HL) is low in this case and thus there is no need to offer them a separate contract.

If there is *weak correlation,* meaning that most flexibility providers are middle type agents (LH or HL), the efficient type (LL) can realize a higher rent by mimicking any of the middle, adjacent types (LH or HL) than by mimicking the inefficient type (HH). In this case the optimal contact involves offering the efficient type the first best and the other three types each separately the second best contract. In contrast to the case of positive correlation, there is no bunching of contracts in this case and middle types produce at level which is higher than the least efficient type (HH) although all three are distorted to maintain incentive compatibility.

If there is *negative correlation with asymmetry towards the LH type*, meaning that middle type agents constitute a higher proportion in the group of flexibility providers and probability of observing an LH type is higher than the other middle type, the efficient type will have an incentive to mimic the LH type but will, essentially, ignore the HL type. The optimal contract in this case include the first best for the efficient type (LL) and (HL) types but the rest will receive the second best contracts. An important difference of this case with previous cases is that in additional to efficient type, one of the middle types also receives the first best contract. This is one of the important differences between

unidimensional and multi-dimensional adverse selection procurement in the sense that when the number of dimesons increases, sometimes a first best contract can be given to a not fully efficient agent (here HL) even when there is information asymmetry.

Symmetrically, if there is n*egative correlation with asymmetry towards the* HL *type,* the probability of observing the HL is higher than the other middle type in a distribution in which middle type agents dominate. Consequently, the efficient will have an incentive to mimic the HL while ignoring the LH type. In this situation, the principal gives the first best contract to LL and LH types but distorts HL and HH's types output.

The second important results of the model in this paper is that designing the optimal contract for flexibility services is complicated not only because of multidimensional information asymmetry but also because of the fact that the composing elements of flexibility are non-separable (capacity, ramp rate, duration, cannot be produced or consumed separately). This "non-separable externality" leads to further distortion of inefficient types beyond the fundamental rent efficiency trade off prevailing under information asymmetry. The non-separability distortion is unique to non-conventional commodities such as flexibility and it does not exist when cost and utility function of agent and principal is separable (a condition that has been assumed almost always in the contract theory literature).

Besides the conceptual discussion about flexibility and the theoretical contribution and that constitutes the model itself, by way of a real-life example, the last section of the paper shows that the model is *applicable* by decision makers who wish to incentivise the provision of flexibility or, for that matter, any multi-dimensional commodity with imperfectly substitutable components. Relying on surveys, empirical analyses and suitably designed experiments that elicit the cost of acting flexibly together with the distribution of types in a given area, it is *feasible* to actually design optimal contracts for flexibility services. Existing contracts fail to take any of this information into account and are, therefore, ill-positioned to deliver economic efficiency when for example applied to a demand response programme.

The analysis of the paper has, of course, its limitations. First, we claim that flexibility is multi-dimensional but our model is bi-dimensional only: this is, of course, to gain tractability. The economic intuition, of course, carries over to a multi-dimensional framework. Second, unlike most treatments of the principal-agent model in the literature, we deal with a discrete set of types. However, in line with Vohra (2011) "we know of no modeling reason to prefer a continuous type space to a discrete one". Third, we are aware that the simulation results could cover a wider range of applications and considerations. For example, it would be interesting to obtain real empirical data to apply the contract modelling framework with real information. It would also be interesting to extend the model to consider the situation in which the principal is an intermediary (e.g. an aggregator) who faces *downstream* uncertainty in relation to the distribution of types and its hidden cost parameters but also faces *upstream* uncertainty coming from market fluctuations. Furthermore, analysing the impact of risk attitudes and competitive environments are relevant but beyond the scope of this paper.

References

Armstrong, M., & Rochet, J. C. (1999). Multi-dimensional screening:: A user's guide. *European Economic Review*, *43*(4), 959-979.

Asker, J., and Cantillon, E. (2010). Procurement when price and quality matter. *The Rand journal of economics*, *41*(1), 1-34.

Boscán, L. (2016a). "Product-Mix Exchanges and Incentive Issues". Copenhagen Business School, Department of Economics (mimeo.)

Boscán, L. (2016b). "Power system flexibility: A product design perspective". Copenhagen Business School, Department of Economics (mimeo.)

Boscán, L. and Poudineh, R. (2016) "Business Models for Power System Flexibility: new actors, new roles, new rules" in "The Future of Utilities: Utilities of the Future", Fereidoon P. Sioshansi (ed.), Elsevier.

Boscán, L. and Poudineh, R. "Procuring power system flexibility in competitive environments". In preparation.

Che, Y. (1993). Design competition through multi-dimensional auctions. *The RAND Journal of Economics*. Vol. 24, No. 4, pp. 668-680.

Dana Jr, J. D. (1993). The organization and scope of agents: regulating multiproduct industries. *Journal of Economic Theory*, *59*(2), 288-310.

Davis , O. and Whinston, A. (1962). "Externalities, Welfare and the Theory of Games". *Journal of Political Economy*, Vol. 70, No. 3

De la Fuente, A. (2000). *Mathematical methods and models for economists*. Cambridge University Press.

Ela, E., Milligan, M., Bloom, A., Botterud, A., Townsend, A. and Levin, T. (2014) "Evolution of Wholesale Electricity Market Design with Increasing Levels of Renewable Generation", Technical Report NREL/TP-5D00-61765, September 2014 National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

EURELECTRIC (2014) "Flexibility and Aggregation Requirements for their interaction in the market". A EURELECTRIC paper. http://www.eurelectric.org/media/115877/tf_bal-agr_report_final_je_as-2014-030-0026-01-e.pdf. January 2014.

EURELECTRIC (2015) "Regulatory Recommendations for the Deployment of Flexibility". EG3 Report. Smart Grid Task Force. http://ec.europa.eu/energy/sites/ener/files/documents/EG3%20Final%20-%20January%202015.pdf

Laffont, J. J., & Martimort, D. (2002). *The theory of incentives: the principal-agent model*. Princeton university press.

Lannoye, E., Flynn, D., and M. O'Malley, "Evaluation of Power System Flexibility" in *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 922-931, May 2012.

Li, Z., Ryan, J. K., and Sun, D. (2015). Multi-attribute procurement contracts. *International Journal of Production Economics*, *159*, 137-146.

Katz, J., Milligan, M., & Cochran, J. (2015). Sources of Operational Flexibility, Greening the Grid (No. NREL/FS-6A20-63039). NREL (National Renewable Energy Laboratory (NREL), Golden, CO (United States)).

Makarov, Y. V., Loutan, C., Ma, J., & De Mello, P. (2009). Operational impacts of wind generation on California power systems. *Power Systems, IEEE Transactions on*, *24*(2), 1039-1050.

Marchand, J. and Russell, K. (1973). “Externalities, Separability, and Resource Allocation”. *The American Economic Review*, Vol. 63 , No. 4.

Morales, J. M., Conejo, A. J., Madsen, H., Pinson, P., & Zugno, M. (2013). *Integrating renewables in electricity markets: Operational problems* (Vol. 205). Springer Science & Business Media.

Myerson, R. B. (2005). *Probability models for economic decisions*. Thomson/Brooke/Cole.

Parkes, D. C., and Kalagnanam, J. (2005). Models for iterative multiattribute procurement auctions. *Management Science*, *51*(3), 435-451.

Rochet, J. C., and Stole, L. A. (2003). The economics of multidimensional screening. *Econometric Society Monographs*, *35*, 150-197.

Rocky Mountain institute (2015)” Frequency and duration of positive and negative net load events for six renewable portfolios” online: http://www.rmi.org/RFGraph-frequency_duration_net_loads. Accessed on 15/04/2015.

Sharkey, W. W. (1982). *The theory of natural monopoly* (Vol. 5). Cambridge: Cambridge University Press.

Ulbig, A., & Andersson, G. (2012). On operational flexibility in power systems. In *Power and Energy Society General Meeting, 2012 IEEE* (pp. 1-8). IEEE.

Vohra, R. (2011*). Mechanism Design: A Linear Programming Approach.* Econometric Society Monographs, volume 47. Cambridge University Press.

J. Zhao, T. Zheng and E. Litvinov, "A Unified Framework for Defining and Measuring Flexibility in Power System," in *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 339-347, Jan. 2016.

Appendix

A1. Fully constrained program

This section presents the fully constrained program.

The principal maximises expected profit:

$$\max_{\{q_{ij},t_{ij},U_{ij}\}} E(\pi) = \sum_{ij} p_{ij} S(q_{ij}, t_{ij}) - \sum_{ij} p_{ij} U_{ij}$$

subject to:

<u>Individual Rationality (IR) constraints:</u>

$T_{LL} - c(q_{LL}, t_{LL}, \theta_q^L, \theta_t^L) = U_{LL} \geq 0$ (A1)

$T_{LH} - c(q_{LH}, t_{LH}, \theta_q^L, \theta_t^H) = U_{LH} \geq 0$ (A2)

$T_{HL} - c(q_{HL}, t_{HL}, \theta_q^H, \theta_t^L) = U_{HL} \geq 0$ (A3)

$T_{HL} - c(q_{HL}, t_{HL}, \theta_q^H, \theta_t^L) = U_{HH} \geq 0$ (A4)

Incentive Compatibility (IC) constraints:

For the efficient (LL) type:

$T_{LL} - c(q_{LL}, t_{LL}, \theta_q^L, \theta_t^L) \geq T_{LH} - c(q_{LH}, t_{LH}, \theta_q^L, \theta_t^L)$ which is equivalent to

$U_{LL} \geq U_{LH} + c(q_{LH}, t_{LH}, \theta_q^L, \theta_t^H) - c(q_{LH}, t_{LH}, \theta_q^L, \theta_t^L) = U_{LH} + \Delta_{LL}^{LH}(q_{LH}, t_{LH})$ (A5)

$T_{LL} - c(q_{LL}, t_{LL}, \theta_q^L, \theta_t^L) \geq T_{HL} - c(q_{HL}, t_{HL}, \theta_q^L, \theta_t^L)$ which is equivalent to

$U_{LL} \geq U_{HL} + c(q_{HL}, t_{HL}, \theta_q^H, \theta_t^L) - c(q_{HL}, t_{HL}, \theta_q^L, \theta_t^L) = U_{HL} + \Delta_{LL}^{HL}(q_{HL}, t_{HL})$ (A6)

$T_{LL} - c(q_{LL}, t_{LL}, \theta_q^L, \theta_t^L) \geq T_{HH} - c(q_{HH}, t_{HH}, \theta_q^L, \theta_t^L)$ which is equivalent to

$U_{LL} \geq U_{HH} + c(q_{HH}, t_{HH}, \theta_q^H, \theta_t^H) - c(q_{HH}, t_{HH}, \theta_q^L, \theta_t^L) = U_{HH} + \Delta_{LL}^{HH}(q_{HH}, t_{HH})$ (A7)

For the middle (LH) type:

$T_{LH} - c(q_{LH}, t_{LH}, \theta_q^L, \theta_t^H) \geq T_{LL} - c(q_{LL}, t_{LL}, \theta_q^L, \theta_t^H)$ or

$U_{LH} \geq U_{LL} - c(q_{LL}, t_{LL}, \theta_q^L, \theta_t^H) + c(q_{LL}, t_{LL}, \theta_q^L, \theta_t^L) = U_{LL} + \Delta_{LH}^{LL}(q_{LL}, t_{LL})$ (A8)

$T_{LH} - c(q_{LH}, t_{LH}, \theta_q^L, \theta_t^H) \geq T_{HL} - c(q_{HL}, t_{HL}, \theta_q^L, \theta_t^H)$ or

$U_{LH} \geq U_{HL} + c(q_{HL}, t_{HL}, \theta_q^H, \theta_t^L) - c(q_{HL}, t_{HL}, \theta_q^L, \theta_t^H) = U_{HL} + \Delta_{LH}^{HL}(q_{HL}, t_{HL})$ (A9)

$T_{LH} - c(q_{LH}, t_{LH}, \theta_q^L, \theta_t^H) \geq T_{HH} - c(q_{HH}, t_{HH}, \theta_q^L, \theta_t^H)$ or

$U_{LH} \geq U_{HH} + c(q_{HH}, t_{HH}, \theta_q^H, \theta_t^H) - c(q_{HH}, t_{HH}, \theta_q^L, \theta_t^H) = U_{HH} + \Delta_{LH}^{HH}(q_{HH}, t_{HH})$ (A10)

For the middle (HL) type:

$T_{HL} - c(q_{HL}, t_{HL}, \theta_q^H, \theta_t^L) \geq T_{LL} - c(q_{LL}, t_{LL}, \theta_q^H, \theta_t^L)$ or

$U_{HL} \geq U_{LL} - c(q_{LL}, t_{LL}, \theta_q^H, \theta_t^L) + c(q_{LL}, t_{LL}, \theta_q^L, \theta_t^L) = U_{LL} + \Delta_{HL}^{LL}(q_{LL}, t_{LL})$ (A11)

$T_{HL} - c(q_{HL}, t_{HL}, \theta_q^H, \theta_t^L) \geq T_{LH} - c(q_{LH}, t_{LH}, \theta_q^H, \theta_t^L)$ or

$U_{HL} \geq U_{LH} + c(q_{LH}, t_{LH}, \theta_q^L, \theta_t^H) - c(q_{LH}, t_{LH}, \theta_q^H, \theta_t^L) = U_{LH} + \Delta_{HL}^{LH}(q_{LH}, t_{LH})$ (A12)

$T_{HL} - c(q_{HL}, t_{HL}, \theta_q^H, \theta_t^L) \geq T_{HH} - c(q_{HH}, t_{HH}, \theta_q^H, \theta_t^L)$ or

$U_{HL} \geq U_{HH} + c(q_{HH}, t_{HH}, \theta_q^H, \theta_t^H) - c(q_{HH}, t_{HH}, \theta_q^H, \theta_t^L) = U_{HH} + \Delta_{HL}^{HH}(q_{HH}, t_{HH})$ (A13)

For the inefficient (HH) type:

$T_{HH} - c(q_{HH}, t_{HH}, \theta_q^H, \theta_t^H) \geq T_{LL} - c(q_{LL}, t_{LL}, \theta_q^H, \theta_t^H)$ which is equivalent to

$U_{HH} \geq U_{LL} - c(q_{LL}, t_{LL}, \theta_q^H, \theta_t^H) + c(q_{LL}, t_{LL}, \theta_q^L, \theta_t^L) = U_{LL} + \Delta_{HH}^{LL}(q_{LL}, t_{LL})$ (A14)

$T_{HH} - c(q_{HH}, t_{HH}, \theta_q^H, \theta_t^H) \geq T_{HL} - c(q_{HL}, t_{HL}, \theta_q^H, \theta_t^H)$ or

$U_{HH} \geq U_{HL} - c(q_{HL}, t_{HL}, \theta_q^H, \theta_t^H) + c(q_{HL}, t_{HL}, \theta_q^H, \theta_t^L) = U_{HL} + \Delta_{HH}^{HL}(q_{HL}, t_{HL})$ (A15)

$T_{HH} - c(q_{HH}, t_{HH}, \theta_q^H, \theta_t^H) \geq T_{LH} - c(q_{LH}, t_{LH}, \theta_q^H, \theta_t^H)$ or

$U_{HH} \geq U_{LH} - c(q_{LH}, t_{LH}, \theta_q^H, \theta_t^H) + c(q_{LH}, t_{LH}, \theta_q^L, \theta_t^H) = U_{LH} + \Delta_{HH}^{LH}(q_{LH}, t_{LH})$ (A16)

A2. Analysis of the incentive compatibility constraints

A2.1 Monotonicity of output levels

The monotonicity of output levels follows from the addition of local incentive constraints two by two.

For q:

- Adding (A6) and (A11) establishes $q_{LL} \geq q_{HL}$:

$$c(q_{LL}, t_{LL}, \theta_q^H, \theta_t^L) - c(q_{HL}, t_{HL}, \theta_q^H, \theta_t^L) \geq c(q_{LL}, t_{LL}, \theta_q^L, \theta_t^L) - c(q_{HL}, t_{HL}, \theta_q^L, \theta_t^L)$$

- And adding (A10) with (A16) establishes $q_{LH} \geq q_{HH}$:

$$c(q_{LH}, t_{LH}, \theta_q^H, \theta_t^H) - c(q_{HH}, t_{HH}, \theta_q^H, \theta_t^H) \geq c(q_{LH}, t_{LH}, \theta_q^L, \theta_t^H) - c(q_{HH}, t_{HH}, \theta_q^L, \theta_t^H)$$

Similarly, for t:

- Adding (A5) and (A8) establishes $t_{LL} \geq t_{LH}$:

$$c(q_{LL}, t_{LL}, \theta_q^L, \theta_t^H) - c(q_{LH}, t_{LH}, \theta_q^L, \theta_t^H) \geq c(q_{LL}, t_{LL}, \theta_q^L, \theta_t^L) - c(q_{LH}, t_{LH}, \theta_q^L, \theta_t^L)$$

- Adding (A13) and (A15) establishes $t_{HL} \geq t_{HH}$:

$$c(q_{HL}, t_{HL}, \theta_q^H, \theta_t^H) - c(q_{HH}, t_{HH}, \theta_q^H, \theta_t^H) \geq c(q_{HL}, t_{HL}, \theta_q^H, \theta_t^L) - c(q_{HH}, t_{HH}, \theta_q^H, \theta_t^L)$$

However, adding (A9) with (A12) results in:

$$c(q_{LH}, t_{LH}, \theta_q^H, \theta_t^L) - c(q_{HL}, t_{HL}, \theta_q^H, \theta_t^L) \geq c(q_{LH}, t_{LH}, \theta_q^L, \theta_t^H) - c(q_{HL}, t_{HL}, \theta_q^L, \theta_t^H)$$

which is satisfied if $q_{LH} \geq q_{HL}$ and $t_{HL} \geq t_{LH}$, but also if $q_{LH} \geq q_{HL}$ and $t_{HL} < t_{LH}$.

Re-organizing the inequality,

$$c(q_{HL}, t_{HL}, \theta_q^L, \theta_t^H) - c(q_{LH}, t_{LH}, \theta_q^L, \theta_t^H) \geq c(q_{HL}, t_{HL}, \theta_q^H, \theta_t^L) - c(q_{LH}, t_{LH}, \theta_q^H, \theta_t^L)$$

it is easy to verify that it holds if $t_{HL} \geq t_{LH}$ and $q_{LH} \geq q_{HL}$, but also if $t_{HL} \geq t_{LH}$ and $q_{LH} < q_{HL}$. Therefore, monotonicity of the middle types' output levels cannot be established.

A2.2 Binding constraints for the LL type

Inequalities (A5), (A6) and (A7) are the three IC constraints for the efficient type, and all of them are upward.

By mimicking the LH type, the efficient obtains an informational rent of

$$U_{LL} \geq U_{LH} + \Delta_{LL}^{LH}(q_{LH}, t_{LH})$$

If the efficient mimics the HL type, he would obtain an informational rent of

$$U_{LL} \geq U_{HL} + \Delta_{LL}^{HL}(q_{HL}, t_{HL})$$

And if the LL mimics the HH, the informational rent is

$$U_{LL} \geq U_{HH} + \Delta_{LL}^{HH}(q_{HH}, t_{HH})$$

Because there is no *a priori* ordering of the unit cost difference between the efficient and the inefficient in both dimensions, i.e. $(\theta_t^H - \theta_t^L)$ can be equal, greater or lower than $(\theta_q^H - \theta_q^L)$, and because monotonicity for the middle-types' output levels cannot be established, it is not possible to

determine which of the incentive constraints will bind at the optimum or if only one of them will bind. Therefore:

$$U_{LL} \geq \max\{U_{LH} + \Delta_{LL}^{LH}(q_{LH}, t_{LH}), U_{HL} + \Delta_{LL}^{HL}(q_{HL}, t_{HL}), U_{HH} + \Delta_{LL}^{HH}(q_{HH}, t_{HH})\}$$

A.2.3 Binding constraints for the middle types LH and HL

The LH type agent can mimic the HL type in which case he would realize a rent of:

$$U_{HL} + \Delta_{LH}^{HL}(q_{HL}, t_{HL})$$

Agent LH can also mimic the HH, obtaining a rent of:

$$U_{HH} + \Delta_{LH}^{HH}(q_{HH}, t_{HH})$$

Similarly, the HL can mimic the LH realizing:

$$U_{LH} + \Delta_{HL}^{LH}(q_{LH}, t_{LH})$$

Or it can mimic HH obtaining:

$$U_{HH} + \Delta_{HL}^{HH}(q_{HH}, t_{HH})$$

For the same reasons mentioned previously, that is: an absence of an *a priori* ordering of the unit cost difference between the efficient and the inefficient in both dimensions, i.e. $(\theta_t^H - \theta_t^L)$ and $(\theta_q^H - \theta_q^L)$, and an incomplete ordering of the middle types' output levels, it is not possible to determine which of the incentive constraints will bind at the optimum or if only one of them will bind. Therefore, U_{LH} and U_{HL} are respectively,

$$U_{LH} \geq \max\{U_{HL} + \Delta_{LH}^{HL}(q_{HL}, t_{HL}), U_{HH} + \Delta_{LH}^{HH}(q_{HH}, t_{HH})\}$$

$$U_{HL} \geq \max\{U_{LH} + \Delta_{HL}^{LH}(q_{LH}, t_{LH}), U_{HH} + \Delta_{HL}^{HH}(q_{HH}, t_{HH})\}$$

as in the program relaxation described in section 2.2.

A3. Solution to the baseline relaxed program (technical details)

To solve the problem, it is convenient to re-write equation (3) in the main text as a convex combination of the rents that result from mimicking the (adjacent) middle types and the (non-adjacent) inefficient type:

$$U_{LL} = \lambda_1\left(\Delta_{LH}^{HH}(q_{HH}, t_{HH}) + \Delta_{LL}^{LH}(q_{LH}, t_{LH})\right) + \lambda_2\left(\Delta_{HL}^{HH}(q_{HH}, t_{HH}) + \Delta_{LL}^{HL}(q_{HL}, t_{HL})\right) + \lambda_3(\Delta_{LL}^{HH}(q_{HH}, t_{HH}))$$

where $\lambda_i \geq 0$, $\sum_i \lambda_i = 1$, $i = 1,2,3$ are real numbers. This approach, used by Armstrong and Rochet (1999), provides a simple way to characterize all the economically relevant cases emerging from the baseline relaxed program.

Substituting for the binding IC constraints into the principal's expected payoff (equation (1) on the main text) leads to:

$$\begin{aligned} p_{LL}\big[S(q_{LL}, t_{LL}) - \lambda_1\left(\Delta_{LH}^{HH}(q_{HH}, t_{HH}) + \Delta_{LL}^{LH}(q_{LH}, t_{LH})\right) - \lambda_2\left(\Delta_{HL}^{HH}(q_{HH}, t_{HH}) + \Delta_{LL}^{HL}(q_{HL}, t_{HL})\right) \\ - \lambda_3(\Delta_{LL}^{HH}(q_{HH}, t_{HH}))\big] + p_{LH}[S(q_{LH}, t_{LH}) - \Delta_{LH}^{HH}(q_{HH}, t_{HH})] \\ + p_{HL}[S(q_{HL}, t_{HL}) - \Delta_{HL}^{HH}(q_{HH}, t_{HH})] + p_{HH}[S(q_{HH}, t_{HH})] \end{aligned}$$

The first-order conditions for the LL type are:

- With respect to q_{LL}:

$$p_{LL}\frac{\partial s(q_{LL},t_{LL})}{\partial q_{LL}} = p_{LL}[\frac{\partial v(q_{LL},t_{LL})}{\partial q_{LL}} - \frac{\partial c\left(q_{LL},t_{LL},{\theta_q}^L,{\theta_t}^L\right)}{\partial q_{LL}}] = 0 \quad \text{(A17)}$$

- With respect to t_{LL}:

$$p_{LL}\frac{\partial s(q_{LL},t_{LL})}{\partial t_{LL}} = p_{LL}[\frac{\partial v(q_{LL},t_{LL})}{\partial t_{LL}} - \frac{\partial c\left(q_{LL},t_{LL},{\theta_q}^L,{\theta_t}^L\right)}{\partial t_{LL}}] = 0 \quad \text{(A18)}$$

So, q_{LL}^{sb} and t_{LL}^{sb} solve the equations above, and it follows that $q_{LL}^{sb} = q_{LL}^{fb}$ and $t_{LL}^{sb} = t_{LL}^{fb}$. This proves equations numbered with (4) in the main text.

In contrast, the less efficient types are deviated from their first-best output levels.

Consider the first-order conditions for the LH type:

- With respect to q_{LH}:

$$-\lambda_1\, p_{LL}\frac{\partial \Delta_{LL}^{LH}(q_{LH},t_{LH})}{\partial q_{LH}} + p_{LH}\frac{\partial s(q_{LH},t_{LH})}{\partial q_{LH}} = 0$$

Expanding $\partial s(q_{LH},t_{LH})/\partial q_{LH}$ and re-arranging:

$$\frac{\partial v(q_{LH},t_{LH})}{\partial q_{LH}} = \frac{\partial c\left(q_{LH},t_{LH},\theta_q^L,\theta_t^H\right)}{\partial q_{LH}} + \lambda_1\frac{p_{LL}}{p_{LH}}\frac{\partial \Delta_{LL}^{LH}(q_{LH},t_{LH})}{\partial q_{LH}} \quad \text{(A19)}$$

Further manipulation leads to:

$$\frac{\frac{\partial v(q_{LH},t_{LH})}{\partial q_{LH}} - \frac{\partial c\left(q_{LH},t_{LH},\theta_q^L,\theta_t^H\right)}{\partial q_{LH}}}{\frac{\partial \Delta_{LL}^{LH}(q_{LH},t_{LH})}{\partial q_{LH}}} = \lambda_1\frac{p_{LL}}{p_{LH}} \quad \text{(A20)}$$

- With respect to t_{LH}:

$$-\lambda_1\, p_{LL}\frac{\partial \Delta_{LL}^{LH}(q_{LH},t_{LH})}{\partial t_{LH}} + p_{LH}\frac{\partial S(q_{LH},t_{LH})}{\partial t_{LH}} = 0$$

Expanding $\partial S(q_{LH},t_{LH})/\partial t_{LH}$ and re-arranging:

$$\frac{\partial v(q_{LH},t_{LH})}{\partial t_{LH}} = \frac{\partial c\left(q_{LH},t_{LH},\theta_q^L,\theta_t^H\right)}{\partial t_{LH}} + \lambda_1\frac{p_{LL}}{p_{LH}}\frac{\partial \Delta_{LL}^{LH}(q_{LH},t_{LH})}{\partial t_{LH}} \quad \text{(A21)}$$

Re-arranging to obtain:

$$\frac{\frac{\partial v(q_{LH},t_{LH})}{\partial t_{LH}} - \frac{\partial c\left(q_{LH},t_{LH},\theta_q^L,\theta_t^H\right)}{\partial t_{LH}}}{\frac{\partial \Delta_{LL}^{LH}(q_{LH},t_{LH})}{\partial t_{LH}}} = \lambda_1\frac{p_{LL}}{p_{LH}} \quad \text{(A22)}$$

The second best levels of output q_{LH}^{sb}, t_{LH}^{sb} satisfy equations (A19) and (A21), respectively, and both q_{LH}^{sb} and t_{LH}^{sb} deviate from first-best levels q_{LH}^{fb}, t_{LH}^{fb} by an amount equal to the second term of the corresponding equations. Equations (A20) and (A22) show that $q_{LH}^{sb} = q\left(\lambda_1\frac{p_{LL}}{p_{LH}}\right), t_{LH}^{sb} = t\left(\lambda_1\frac{p_{LL}}{p_{LH}}\right)$. Note also that $\frac{\partial \Delta_{LL}^{LH}(q_{LH},t_{LH})}{\partial q_{LH}}, \frac{\partial \Delta_{LL}^{LH}(q_{LH},t_{LH})}{\partial t_{LH}} \neq 0$ for non-separable cost functions.
The first-order conditions for the HL type are:

- With respect to q_{HL}:

$$-\lambda_2\, p_{LL}\frac{\partial \Delta_{LL}^{HL}(q_{LH},t_{LH})}{\partial q_{HL}} + \; p_{HL}\frac{\partial s(q_{HL},t_{HL})}{\partial q_{HL}} = 0$$

Expanding $\partial s(q_{HL},t_{HL})/\partial q_{HL}$ and re-arranging:

$$\frac{\partial v(q_{HL},t_{HL})}{\partial q_{HL}} = \frac{\partial c(q_{HL},t_{HL},\theta_q^H,\theta_t^L)}{\partial q_{HL}} + \lambda_2\frac{p_{LL}}{p_{HL}}\frac{\partial \Delta_{LL}^{HL}(q_{HL},t_{HL})}{\partial q_{HL}} \quad \text{(A23)}$$

Re-arranging further to obtain:

$$\frac{\frac{\partial v(q_{HL},t_{HL})}{\partial q_{HL}} - \frac{\partial c(q_{HL},q_{HL},\theta_q^H,\theta_t^L)}{\partial q_{HL}}}{\frac{\partial \Delta_{LL}^{HL}(q_{HL},t_{HL})}{\partial q_{HL}}} = \lambda_2\frac{p_{LL}}{p_{HL}} \qquad \text{(A24)}$$

- With respect to t_{HL}:

$$-\lambda_2\, p_{LL}\frac{\partial \Delta_{LL}^{HL}(q_{HL},t_{HL})}{\partial t_{HL}} + \; p_{HL}\frac{\partial s(q_{HL},t_{HL})}{\partial t_{HL}} = 0$$

Expanding $\partial s(q_{HL},t_{HL})/\partial t_{HL}$ and re-arranging:

$$\frac{\partial v(q_{HL},t_{HL})}{\partial t_{HL}} = \frac{\partial c(q_{HL},t_{HL},\theta_q^H,\theta_t^L)}{\partial t_{HL}} + \lambda_2\frac{p_{LL}}{p_{HL}}\frac{\partial \Delta_{LL}^{HL}(q_{HL},t_{HL})}{\partial t_{HL}} \quad \text{(A25)}$$

Re-arranging again shows:

$$\frac{\frac{\partial v(q_{HL},t_{HL})}{\partial t_{HL}} - \frac{\partial c(q_{HL},q_{HL},\theta_q^H,\theta_t^L)}{\partial t_{HL}}}{\frac{\partial \Delta_{LL}^{HL}(q_{HL},t_{HL})}{\partial t_{HL}}} = \lambda_2\frac{p_{LL}}{p_{HL}} \qquad \text{(A26)}$$

The second best levels of output q_{HL}^{sb}, t_{HL}^{sb} are, respectively, the solutions to (A23) and (A25). In both cases, q_{HL}^{sb} and t_{HL}^{sb} deviate from first-best levels q_{HL}^{fb}, t_{HL}^{fb} by an amount equal to the second term of equations the corresponding equations. Equations (A24) and (A26) show that $q_{HL}^{sb} = q\left(\lambda_2\frac{p_{LL}}{p_{HL}}\right), t_{HL}^{sb} = t\left(\lambda_2\frac{p_{LL}}{p_{HL}}\right)$. Note that $\frac{\partial \Delta_{LL}^{HL}(q_{HL},t_{HL})}{\partial q_{HL}}, \frac{\partial \Delta_{LL}^{HL}(q_{HL},t_{HL})}{\partial t_{HL}} \neq 0$ for non-separable cost functions.

Consider the first-order conditions for the HH type:

- With respect to q_{HH}:

$$-\,p_{LL}\left[\lambda_1\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}} + \lambda_2\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}} + \lambda_3\frac{\partial \Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}}\right] - p_{LH}\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}}$$
$$-\,p_{HL}\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}} + p_{HH}\frac{\partial S(q_{HH},t_{HH})}{\partial q_{HH}} = 0$$

Expanding $\partial s(q_{HH},t_{HH})/\partial q_{HH}$ and re-arranging:

$$\frac{\partial v(q_{HH},t_{HH})}{\partial q_{HH}} = \frac{\partial c(q_{HL},t_{HL},\theta_q^H,\theta_t^H)}{\partial q_{HH}} + \frac{(\lambda_1\, p_{LL} + p_{LH})}{p_{HH}}\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}} + \frac{(\lambda_2\, p_{LL} + p_{HL})}{p_{HH}}\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}} + \frac{\lambda_3\, p_{LL}}{p_{HH}}\frac{\partial \Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}}$$

It follows that the deviation from the first-best depends on the difference in marginal cost between the HH type relative to the LH, the HL and the LL:

$$\frac{\partial v(q_{HH},t_{HH})}{\partial q_{HH}} = \frac{\partial c(q_{HL},t_{HL},\theta_q^H,\theta_t^H)}{\partial q_{HH}} + \frac{1}{p_{HH}}\left[(\lambda_1\, p_{LL} + p_{LH})\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}} + (\lambda_2\, p_{LL} + p_{HL})\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}} + \lambda_3\, p_{LL}\frac{\partial \Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}}\right]$$

- With respect to t_{HH}:

$$-p_{LL}\left[\lambda_1\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}} + \lambda_2\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}} + \lambda_3\frac{\partial \Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}}\right] - p_{LH}\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}} - p_{HL}\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}} + p_{HH}\frac{\partial S(q_{HH},t_{HH})}{\partial t_{HH}} = 0$$

Expanding $\partial s(q_{HH},t_{HH})/\partial t_{HH}$ and re-arranging:

$$\frac{\partial v(q_{HH},t_{HH})}{\partial t_{HH}} = \frac{\partial c(q_{HL},t_{HL},\theta_q^H,\theta_t^H)}{\partial t_{HH}} + \frac{(\lambda_1\, p_{LL} + p_{LH})}{p_{HH}}\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}} + \frac{(\lambda_2\, p_{LL} + p_{HL})}{p_{HH}}\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}} + \frac{\lambda_3\, p_{LL}}{p_{HH}}\frac{\partial \Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}}$$

Thus the deviation from the first-best depends on the difference in marginal cost between the HH type relative to the LH, the HL and the LL:

$$\frac{\partial v(q_{HH},t_{HH})}{\partial t_{HH}} - \frac{\partial c(q_{HL},t_{HL},\theta_q^H,\theta_t^H)}{\partial t_{HH}} = \frac{1}{p_{HH}}\left[(\lambda_1\, p_{LL} + p_{LH})\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}} + (\lambda_2\, p_{LL} + p_{HL})\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}} + \lambda_3\, p_{LL}\frac{\partial \Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}}\right]$$

So,

$$q_{HH}^{sb} = q\left(\frac{1}{p_{HH}}\left[(\lambda_1\, p_{LL} + p_{LH})\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}} + (\lambda_2\, p_{LL} + p_{HL})\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}} + \lambda_3\, p_{LL}\frac{\partial \Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}}\right]\right)$$

$$t_{HH}^{sb} = t\left(\frac{1}{p_{HH}}\left[(\lambda_1\, p_{LL} + p_{LH})\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}} + (\lambda_2\, p_{LL} + p_{HL})\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}} + \lambda_3\, p_{LL}\frac{\partial \Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}}\right]\right)$$

To summarize, the *general* solutions to the baseline relaxed program are:

- For the LL type:

$$q_{LL}^{sb} = q_{LL}^{fb}\ ,\ t_{LL}^{sb} = t_{LL}^{fb} \quad \text{(A27)}$$

- For the LH type:

$$q_{LH}^{sb} = q\left(\lambda_1\frac{p_{LL}}{p_{LH}}\right), t_{LH}^{sb} = t\left(\lambda_1\frac{p_{LL}}{p_{LH}}\right) \text{(A28)}$$

- For the HL type:

$$q_{HL}^{sb} = q\left(\lambda_2\frac{p_{LL}}{p_{HL}}\right), t_{HL}^{sb} = t\left(\lambda_2\frac{p_{LL}}{p_{HL}}\right) \text{(A29)}$$

- For the HH type:

$$q_{HH}^{sb} = q\left(\frac{1}{p_{HH}}\left[(\lambda_1\, p_{LL} + \, p_{LH})\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}} + (\lambda_2\, p_{LL} + \, p_{HL})\frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}} + \lambda_3\, p_{LL}\frac{\partial\Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}}\right]\right)$$

$$t_{HH}^{sb} = t\left(\frac{1}{p_{HH}}\left[(\lambda_1\, p_{LL} + \, p_{LH})\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}} + (\lambda_2\, p_{LL} + \, p_{HL})\frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}} + \lambda_3\, p_{LL}\frac{\partial\Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}}\right]\right)$$

(A30)

Case 1: Positive correlation

The second case arises when $\lambda_1, \lambda_2, \lambda_3 > 0$ such that:

$$U_{LL} = \begin{cases} \left(\Delta_{LH}^{HH}(q_{HH}^{sb}, t_{HH}^{sb}) + \Delta_{LL}^{LH}(q_{LH}^{sb}, t_{LH}^{sb})\right) \\ \left(\Delta_{HL}^{HH}(q_{HH}^{sb}, t_{HH}^{sb}) + \Delta_{LL}^{HL}(q_{HL}^{sb}, t_{HL}^{sb})\right) \\ \left(\Delta_{LL}^{HH}(q_{HH}^{sb}, t_{HH}^{sb})\right) \end{cases}$$

That is, in this case it is optimal for the principal to give the efficient type a rent that will make him indifferent between his own contract and any of the three remaining agent's.

- Setting $\Delta_{LH}^{HH}(q_{HH}^{sb}, t_{HH}^{sb}) + \Delta_{LL}^{LH}(q_{LH}^{sb}, t_{LH}^{sb}) = \Delta_{LL}^{HH}(q_{HH}^{sb}, t_{HH}^{sb})$ and expanding:

$$c\left(q_{HH}^{sb}, t_{HH}^{sb}, {\theta_q}^H, {\theta_t}^H\right) - c\left(q_{HH}^{sb}, t_{HH}^{sb}, {\theta_q}^L, {\theta_t}^H\right) + c\left(q_{LH}^{sb}, t_{LH}^{sb}, {\theta_q}^L, {\theta_t}^H\right) - c\left(q_{LH}^{sb}, t_{LH}^{sb}, {\theta_q}^L, {\theta_t}^L\right)$$
$$= c\left(q_{HH}^{sb}, t_{HH}^{sb}, {\theta_q}^H, {\theta_t}^H\right) - c\left(q_{HH}^{sb}, t_{HH}^{sb}, {\theta_q}^L, {\theta_t}^L\right)$$

  Simplifying and re-arranging:

$$c\left(q_{LH}^{sb}, t_{LH}^{sb}, {\theta_q}^L, {\theta_t}^H\right) - c\left(q_{LH}^{sb}, t_{LH}^{sb}, {\theta_q}^L, {\theta_t}^L\right) = c\left(q_{HH}^{sb}, t_{HH}^{sb}, {\theta_q}^L, {\theta_t}^H\right) - c\left(q_{HH}^{sb}, t_{HH}^{sb}, {\theta_q}^L, {\theta_t}^L\right)$$
$$\Delta_{LL}^{LH}(q_{LH}^{sb}, t_{LH}^{sb}) = \Delta_{LL}^{LH}(q_{HH}^{sb}, t_{HH}^{sb}) \Leftrightarrow q_{LH}^{sb} = q_{HH}^{sb}; t_{LH}^{sb} = t_{HH}^{sb}$$

- Likewise, $\Delta_{HL}^{HH}(q_{HH}^{sb}, t_{HH}^{sb}) + \Delta_{LL}^{HL}(q_{HL}^{sb}, t_{HL}^{sb}) = \Delta_{LL}^{HH}(q_{HH}^{sb}, t_{HH}^{sb})$ and expanding:

$$c\left(q_{HH}^{sb}, t_{HH}^{sb}, {\theta_q}^H, {\theta_t}^H\right) - c\left(q_{HH}^{sb}, t_{HH}^{sb}, {\theta_q}^H, {\theta_t}^L\right) + c\left(q_{HL}^{sb}, t_{HL}^{sb}, {\theta_q}^H, {\theta_t}^L\right) - c\left(q_{HL}^{sb}, t_{HL}^{sb}, {\theta_q}^L, {\theta_t}^L\right)$$
$$= c\left(q_{HH}^{sb}, t_{HH}^{sb}, {\theta_q}^H, {\theta_t}^H\right) - c\left(q_{HH}^{sb}, t_{HH}^{sb}, {\theta_q}^L, {\theta_t}^L\right)$$

  Simplifying and re-arranging:

$$c\left(q_{HL}^{sb}, t_{HL}^{sb}, {\theta_q}^H, {\theta_t}^L\right) - c\left(q_{HL}^{sb}, t_{HL}^{sb}, {\theta_q}^L, {\theta_t}^L\right) = c\left(q_{HH}^{sb}, t_{HH}^{sb}, {\theta_q}^H, {\theta_t}^L\right) - c\left(q_{HH}^{sb}, t_{HH}^{sb}, {\theta_q}^L, {\theta_t}^L\right)$$
$$\Delta_{LL}^{HL}(q_{HL}^{sb}, t_{HL}^{sb}) = \Delta_{LL}^{HL}(q_{HH}^{sb}, t_{HH}^{sb}) \Leftrightarrow q_{HL}^{sb} = q_{HH}^{sb}; t_{HL}^{sb} = t_{HH}^{sb}$$

Thus, output levels are equal for the less efficient types: $q_{LH}^{sb} = q_{HL}^{sb} = q_{HH}^{sb}; t_{LH}^{sb} = t_{HL}^{sb} = t_{HH}^{sb}$.

In consequence, from (A28), (A29) and (A30), the deviation from the first best for $q_{LH}^{sb}, q_{HL}^{sb}, q_{HH}^{sb}$ and $t_{LH}^{sb}, t_{HL}^{sb}, t_{HH}^{sb}$ is, respectively:

For element q:

$$\lambda_1\frac{p_{LL}}{p_{LH}} = \lambda_2\frac{p_{LL}}{p_{HL}} = \frac{1}{p_{HH}}\left[(\lambda_1\, p_{LL} + \, p_{LH})\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}} + (\lambda_2\, p_{LL} + \, p_{HL})\frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}} + \lambda_3\, p_{LL}\frac{\partial\Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}}\right]$$

For element t:

$$\lambda_1 \frac{p_{LL}}{p_{LH}} = \lambda_2 \frac{p_{LL}}{p_{HL}} = \frac{1}{p_{HH}} \left[(\lambda_1 \, p_{LL} + \, p_{LH}) \frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial t_{HH}} + (\lambda_2 \, p_{LL} + \, p_{HL}) \frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial t_{HH}} + \lambda_3 \, p_{LL} \frac{\partial \Delta_{LL}^{HH}(q_{HH}, t_{HH})}{\partial t_{HH}} \right]$$

that is, two different systems of three equations in three unknowns. Because of symmetry between the systems, it is more compact to write and solve:

$$\lambda_1 \frac{p_{LL}}{p_{LH}} = \lambda_2 \frac{p_{LL}}{p_{HL}} = \frac{1}{p_{HH}} \left[(\lambda_1 \, p_{LL} + \, p_{LH}) \frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} + (\lambda_2 \, p_{LL} + \, p_{HL}) \frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} + \lambda_3 \, p_{LL} \frac{\partial \Delta_{LL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} \right]$$

where $z \in \{q, t\}$.

Finding the second-best output levels for elements q, t thus amounts to solving the following:

$\lambda_1 + \lambda_2 + \lambda_3 = 1$ (A31)

$$\lambda_1 \frac{p_{LL}}{p_{LH}} = \frac{1}{p_{HH}} \left[(\lambda_1 \, p_{LL} + \, p_{LH}) \frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} + (\lambda_2 \, p_{LL} + \, p_{HL}) \frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} + \lambda_3 \, p_{LL} \frac{\partial \Delta_{LL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} \right] \text{ (A32)}$$

$$\lambda_2 \frac{p_{LL}}{p_{HL}} = \frac{1}{p_{HH}} \left[(\lambda_1 \, p_{LL} + \, p_{LH}) \frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} + (\lambda_2 \, p_{LL} + \, p_{HL}) \frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} + \lambda_3 \, p_{LL} \frac{\partial \Delta_{LL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} \right] \text{ (A33)}$$

Substituting for $\lambda_3 = 1 - \lambda_1 - \lambda_2$ into (A32) and (A33) simplifies the system of equations to two equations in two unknowns:

$$\lambda_1 \frac{p_{LL}}{p_{LH}} = \frac{1}{p_{HH}} \left[(\lambda_1 \, p_{LL} + \, p_{LH}) \frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} + (\lambda_2 \, p_{LL} + \, p_{HL}) \frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} + (1 - \lambda_1 - \lambda_2) \, p_{LL} \frac{\partial \Delta_{LL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} \right] \text{ (A34)}$$

$$\lambda_2 \frac{p_{LL}}{p_{HL}} = \frac{1}{p_{HH}} \left[(\lambda_1 \, p_{LL} + \, p_{LH}) \frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} + (\lambda_2 \, p_{LL} + \, p_{HL}) \frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} + (1 - \lambda_1 - \lambda_2) \, p_{LL} \frac{\partial \Delta_{LL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} \right] \text{ (A35)}$$

Re-organizing the equations:

$$\lambda_1 \left(\frac{p_{LL} \, p_{HH}}{p_{LH}} - \, p_{LL} \frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} + \, p_{LL} \frac{\partial \Delta_{LL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} \right) + \lambda_2 \left(- \, p_{LL} \frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} + \, p_{LL} \frac{\partial \Delta_{LL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} \right) = p_{LH} \frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} + \, p_{HL} \frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} + \, p_{LL} \frac{\partial \Delta_{LL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}}$$

$$\lambda_1 \left(\, p_{LL} \frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} - \, p_{LL} \frac{\partial \Delta_{LL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} \right) + \lambda_2 \left(\, p_{LL} \frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} - \, p_{LL} \frac{\partial \Delta_{LL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} - \frac{p_{LL} \, p_{HH}}{p_{HL}} \right) = - \, p_{LH} \frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} - p_{HL} \frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} - \, p_{LL} \frac{\partial \Delta_{LL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}}$$

After simplification, the expressions for $\lambda_1, \lambda_2, \lambda_3$ are:

$$\lambda_1 = \frac{\left(p_{LH} \frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} + p_{HL} \frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} + p_{LL} \frac{\partial \Delta_{LL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} \right)}{\frac{p_{LL}}{p_{LH}} \left(p_{HH} - p_{HL} \left(\frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} - \frac{\partial \Delta_{LL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} \right) - p_{LH} \left(\frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} - \frac{\partial \Delta_{LL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} \right) \right)} \text{ (A36)}$$

$$\lambda_2 = \frac{\left(p_{LH}\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + p_{HL}\frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + p_{LL}\frac{\partial\Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}\right)}{\frac{p_{LL}}{p_{HL}}\left(p_{HH} - p_{HL}\left(\frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} - \frac{\partial\Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}\right) - p_{LH}\left(\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} - \frac{\partial\Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}\right)\right)} \quad \text{(A37)}$$

$$\lambda_3 = 1 - \frac{\left(p_{LH}\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + p_{HL}\frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + p_{LL}\frac{\partial\Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}\right)}{p_{LL}\left(p_{HH} - p_{HL}\left(\frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} - \frac{\partial\Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}\right) - p_{LH}\left(\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} - \frac{\partial\Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}\right)\right)}(p_{LH} + p_{HL}) \quad \text{(A38)}$$

Finally, it remains to establish that $\lambda_1, \lambda_2, \lambda_3 > 0$. To this end, we prove the following:

<u>Lemma</u>

1. $\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}, \frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}, \frac{\partial\Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} > 0$ for $z \in \{q, t\}$
2. $\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} - \frac{\partial\Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} < 0$ and $\frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} - \frac{\partial\Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} < 0$, for $z \in \{q, t\}$

<u>Proof:</u>

From the agent's payoff function $U = T - c(q, t, \theta_q, \theta_t)$, we know that $\frac{\partial U}{\partial \theta_z} = -\frac{\partial c}{\partial \theta_z}$, $z \in \{q, t\}$ the payoff is decreasing in the agent's type. Alternatively, we write $\frac{\partial c}{\partial \theta_z} > 0$, so the cost increases with the agent's type.

Therefore, if the HH type produces outputs (q_{HH}, t_{HH}), its cost is always strictly greater than if it were produced by any other type. So we can write:

For $ij \in \{LH, HL, LL\}$:

$$c(q_{HH}, t_{HH}, \theta_q^H, \theta_t^H) > c(q_{HH}, t_{HH}, \theta_q^i, \theta_t^j)$$

$$c(q_{HH}, t_{HH}, \theta_q^H, \theta_t^H) - c(q_{HH}, t_{HH}, \theta_q^i, \theta_t^j) > 0$$

$$\frac{\partial c(q_{HH}, t_{HH}, \theta_q^H, \theta_t^H)}{\partial z_{HH}} - \frac{\partial c(q_{HH}, t_{HH}, \theta_q^i, \theta_t^j)}{\partial z_{HH}} > 0$$

$$\frac{\partial\Delta_{ij}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} > 0$$

which proves statement 1.

Also note that, for $ij \in \{LH, HL\}$:

$$c(q_{HH}, t_{HH}, \theta_q^i, \theta_t^j) > c(q_{HH}, t_{HH}, \theta_q^L, \theta_t^L)$$

$$-c(q_{HH}, t_{HH}, \theta_q^i, \theta_t^j) < -c(q_{HH}, t_{HH}, \theta_q^L, \theta_t^L)$$

Adding $c(q_{HH}, t_{HH}, \theta_q^H, \theta_t^H)$ at both sides of the inequality:

$$c\left(q_{HH}, t_{HH}, \theta_q^H, \theta_t^H\right) - c\left(q_{HH}, t_{HH}, \theta_q^i, \theta_t^j\right) < c\left(q_{HH}, t_{HH}, \theta_q^H, \theta_t^H\right) - c\left(q_{HH}, t_{HH}, \theta_q^L, \theta_t^L\right)$$

$$\Delta_{ij}^{HH}(q_{HH}, t_{HH}) < \Delta_{LL}^{HH}(q_{HH}, t_{HH})$$

Deriving with respect to $z \in \{q, t\}$:

$$\frac{\partial \Delta_{ij}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} < \frac{\partial \Delta_{LL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}}$$

$$\frac{\partial \Delta_{ij}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} - \frac{\partial \Delta_{LL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} < 0$$

which proves statement 2.

It follows that λ_1, λ_2 are strictly positive. It remains to prove that $\lambda_3 > 0$:

$$\lambda_3 = 1 - \frac{\left(p_{LH}\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + p_{HL}\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + p_{LL}\frac{\partial \Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}\right)}{p_{LL}\left(p_{HH} - p_{HL}\left(\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} - \frac{\partial \Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}\right) - p_{LH}\left(\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} - \frac{\partial \Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}\right)\right)}(p_{LH} + p_{HL}) > 0$$

$$\frac{\left(p_{LH}\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + p_{HL}\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + p_{LL}\frac{\partial \Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}\right)}{p_{LL}\left(p_{HH} - p_{HL}\left(\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} - \frac{\partial \Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}\right) - p_{LH}\left(\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} - \frac{\partial \Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}\right)\right)}(p_{LH} + p_{HL}) < 1$$

$$\left(p_{LH}\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + p_{HL}\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + p_{LL}\frac{\partial \Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}\right)(p_{LH} + p_{HL}) < p_{LL}\left(p_{HH} - p_{HL}\left(\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} - \frac{\partial \Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}\right) - p_{LH}\left(\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} - \frac{\partial \Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}\right)\right)$$

$$p_{LH}{}^2\frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} + p_{LH}\,p_{HL}\frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} + p_{HL}\,p_{LH}\frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} + p_{HL}{}^2\frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} < p_{LL}\,p_{HH} - p_{LL}\,p_{HL}\frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} - p_{LL}\,p_{LH}\frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}}$$

which simplifies to the following condition:

$$cov\left(\theta_q^i, \theta_t^j\right) > (1 - p_{HH})\left(p_{LH}\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + p_{HL}\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}\right) - p_{LH}\,p_{HL} \text{ for } z \in \{q, t\} \quad \text{(A39)}$$

Substituting for $\lambda_1, \lambda_2, \lambda_3$ in (A28), (A29), (A30) shows:

$$q_{LH}^{sb} = q_{HL}^{sb} = q_{HH}^{sb} = q\left(\frac{p_{LH}\frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial q_{HH}} + p_{HL}\frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial q_{HH}} + p_{LL}\frac{\partial \Delta_{LL}^{HH}(q_{HH}, t_{HH})}{\partial q_{HH}}}{p_{HH} + \frac{\partial \Delta_{LL}^{HH}(q_{HH}, t_{HH})}{\partial q_{HH}}(p_{HL} + p_{LH}) - p_{HL}\frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial q_{HH}} - p_{LH}\frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial q_{HH}}}\right)$$

$$t_{LH}^{sb} = t_{HL}^{sb} = q_{HH}^{sb} = t\left(\frac{p_{LH}\frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial t_{HH}} + p_{HL}\frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial t_{HH}} + p_{LL}\frac{\partial \Delta_{LL}^{HH}(q_{HH}, t_{HH})}{\partial t_{HH}}}{p_{HH} + \frac{\partial \Delta_{LL}^{HH}(q_{HH}, t_{HH})}{\partial t_{HH}}(p_{HL} + p_{LH}) - p_{HL}\frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial t_{HH}} - p_{LH}\frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial t_{HH}}}\right)$$

which are equations (6a-6c) in the main text.

Case 2: Weak correlation

The third case arises when $\lambda_3 = 0$ and $\lambda_1 + \lambda_2 = 1$. For simplicity, let $\lambda_1 = \lambda$ so $\lambda_2 = 1 - \lambda$

$$U_{LL} = \begin{cases} \left(\Delta_{LH}^{HH}(q_{HH}^{sb}, t_{HH}^{sb}) + \Delta_{LL}^{LH}(q_{LH}^{sb}, t_{LH}^{sb})\right) \\ \left(\Delta_{HL}^{HH}(q_{HH}^{sb}, t_{HH}^{sb}) + \Delta_{LL}^{HL}(q_{HL}^{sb}, t_{HL}^{sb})\right) \end{cases}$$

That is, in this case it is optimal for the principal to give the efficient type a rent that will make him indifferent between his own contract and any of the two adjacent, middle types:

$$\Delta_{LH}^{HH}(q_{HH}^{sb}, t_{HH}^{sb}) + \Delta_{LL}^{LH}(q_{LH}^{sb}, t_{LH}^{sb}) = \Delta_{HL}^{HH}(q_{HH}^{sb}, t_{HH}^{sb}) + \Delta_{LL}^{HL}(q_{HL}^{sb}, t_{HL}^{sb}) \quad \text{(A40)}$$

which is the same as equation (7) in the main text.

The "deltas" in (A40) are increasing in their arguments. From the lemma proved above (see the subsection on Case 2), we know that $\Delta_{LH}^{HH}(q_{HH}^{sb}, t_{HH}^{sb})$, $\Delta_{HL}^{HH}(q_{HH}^{sb}, t_{HH}^{sb})$ are increasing in $(q_{HH}^{sb}, t_{HH}^{sb})$.

By a similar argument, we show that $\Delta_{LL}^{LH}(q_{LH}^{sb}, t_{LH}^{sb})$ is increasing in $(q_{LH}^{sb}, t_{LH}^{sb})$ and that $\Delta_{LL}^{HL}(q_{HL}^{sb}, t_{HL}^{sb})$ is increasing in $(q_{HL}^{sb}, t_{HL}^{sb})$:

For $ij \in \{LH, HL\}$:

$$c\left(q_{ij}, t_{ij}, \theta_q^i, \theta_t^j\right) > c\left(q_{ij}, t_{ij}, \theta_q^L, \theta_t^L\right)$$

$$c\left(q_{ij}, t_{ij}, \theta_q^i, \theta_t^j\right) - c\left(q_{ij}, t_{ij}, \theta_q^L, \theta_t^L\right) > 0$$

$$\frac{\partial c\left(q_{ij}, t_{ij}, \theta_q^i, \theta_t^j\right)}{\partial z_{ij}} - \frac{\partial c\left(q_{ij}, t_{ij}, \theta_q^L, \theta_t^L\right)}{\partial z_{ij}} > 0$$

$$\frac{\partial \Delta_{ij}^{HH}\left(q_{ij}, t_{ij}\right)}{\partial z_{ij}} > 0 \text{ for } z \in \{q, t\}$$

However, substituting for $\lambda_1 = \lambda$, $\lambda_2 = 1 - \lambda$, $\lambda_3 = 0$ in (A28), (A29), (A30):

$$q_{LH}^{sb} = q\left(\lambda \frac{p_{LL}}{p_{LH}}\right), t_{LH}^{sb} = t\left(\lambda \frac{p_{LL}}{p_{LH}}\right) \quad \text{(A41)}$$

$$q_{HL}^{sb} = q\left((1-\lambda)\frac{p_{LL}}{p_{HL}}\right), t_{HL}^{sb} = t\left((1-\lambda)\frac{p_{LL}}{p_{HL}}\right) \quad \text{(A42)}$$

$$q_{HH}^{sb} = q\left(\frac{1}{p_{HH}}\left[(\lambda\, p_{LL} + p_{LH})\frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial q_{HH}} + \left((1-\lambda)\, p_{LL} + p_{HL}\right)\frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial q_{HH}}\right]\right) \quad \text{(A43)}$$

$$t_{HH}^{sb} = t\left(\frac{1}{p_{HH}}\left[(\lambda\, p_{LL} + p_{LH})\frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial t_{HH}} + \left((1-\lambda)\, p_{LL} + p_{HL}\right)\frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial t_{HH}}\right]\right) \quad \text{(A44)}$$

Which are equations (8a-8c) in the main text.

reveals that the deviation of $(q_{LH}^{sb}, t_{LH}^{sb})$ relative to the first best $\left(q_{LH}^{fb}, t_{LH}^{fb}\right)$ is increasing in λ, whereas the deviation of $(q_{HL}^{sb}, t_{HL}^{sb})$ relative to $\left(q_{HL}^{fb}, t_{HL}^{fb}\right)$ is decreasing in this parameter.

In contrast, the effect of λ on $(q_{HH}^{sb}, t_{HH}^{sb})$ is ambiguous and depends on the marginal cost difference between the HH and the LH types and the marginal cost difference between the HH and the HL types, $\frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}}, \frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}}$ for $z \in \{q, t\}$.

So, under weak correlation, the principal's problem of defining the optimal output reduces to proving the existence of a λ for which (A40) holds. There are three cases to consider:

i. *Marginal cost symmetry between the middle types:* suppose that

$$\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} = \frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} \text{ for } z \in \{q,t\} \qquad \text{(A45)}$$

holds.

Substituting (A45) into equations (A43) and (A44) leads to the following expressions for q_{HH}^{sb}, t_{HH}^{sb}, which are independent of λ:

$$q_{HH}^{sb} = q\left(\frac{1-p_{HH}}{p_{HH}}\left[\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial q_{HH}}\right]\right)$$

$$t_{HH}^{sb} = t\left(\frac{1-p_{HH}}{p_{HH}}\left[\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial t_{HH}}\right]\right)$$

Therefore, to prove the existence of a λ that satisfies (A40), the arguments of $\Delta_{LL}^{LH}(q_{LH}^{sb}, t_{LH}^{sb})$ and $\Delta_{LL}^{HL}(q_{HL}^{sb}, t_{HL}^{sb})$ are the only relevant terms. The former is decreasing whereas the latter is increasing in λ.

Re-arranging (A40), we obtain:

$$\Delta_{LL}^{LH}(q_{LH}^{sb}, t_{LH}^{sb}) - \Delta_{LL}^{HL}(q_{HL}^{sb}, t_{HL}^{sb}) = \Delta_{HL}^{HH}(q_{HH}^{sb}, t_{HH}^{sb}) - \Delta_{LH}^{HH}(q_{HH}^{sb}, t_{HH}^{sb})$$

From (A45) we know that the difference $\Delta_{HL}^{HH}(q_{HH}^{sb}, t_{HH}^{sb}) - \Delta_{LH}^{HH}(q_{HH}^{sb}, t_{HH}^{sb})$ is a constant k^*, that is:

$$\Delta_{HL}^{HH}(q_{HH}^{sb}, t_{HH}^{sb}) - \Delta_{LH}^{HH}(q_{HH}^{sb}, t_{HH}^{sb}) = k^* \text{ (A46)}$$

Substituting k^* into (A46) and rearranging yields a function $f(\lambda), 0 \leq \lambda \leq 1$:

$$f(\lambda) = \Delta_{LL}^{LH}(q_{LH}^{sb}, t_{LH}^{sb}) - \Delta_{LL}^{HL}(q_{HL}^{sb}, t_{HL}^{sb}) - k^* = 0 \qquad \text{(A47)}$$

The problem now reduces to showing that (A47) has a root and to show this we use a classical result, namely the:

<u>Lemma *(Intermediate Value Theorem, as stated in de La Fuente, 2000)):*</u> Let f be a continuous real-valued function on the closed bounded interval $[a,b]$. Then for each number γ (strictly) between $f(a)$ and $f(b)$, there exists a point $c \in (a,b)$ such that $f(c) = \gamma$.

To be able to apply the result, $f(\lambda)$ must be continuous over λ, which can be easily verified by simple inspection of (A41) and (A42), which are well-defined over the admissible range of λ. Therefore, the sufficient condition for the existence of a root for (A47) over the domain of $f(\lambda)$ is that $f(0)f(1) < 0$.

If $\lambda = 0$:

$$f(0) = \Delta_{LL}^{LH}(q(0), t(0)) - \Delta_{LL}^{HL}\left(q\left(\frac{p_{LL}}{p_{HL}}\right), t\left(\frac{p_{LL}}{p_{HL}}\right)\right) - k^*$$

If $\lambda = 1$:

$$f(1) = \Delta_{LL}^{LH}\left(q\left(\frac{p_{LL}}{p_{LH}}\right), t\left(\frac{p_{LL}}{p_{LH}}\right)\right) - \Delta_{LL}^{HL}(q(0), t(0)) - k^*$$

We now show that $\Delta_{LL}^{LH}(q(0),t(0)) > \Delta_{LL}^{HL}\left(q\left(\frac{p_{LL}}{p_{HL}}\right),t\left(\frac{p_{LL}}{p_{HL}}\right)\right)$ and that $\Delta_{LL}^{LH}\left(q\left(\frac{p_{LL}}{p_{LH}}\right),t\left(\frac{p_{LL}}{p_{LH}}\right)\right) < \Delta_{LL}^{HL}(q(0),t(0))$ hold.

<u>Proof:</u>

Let us assume that $\Delta_{LL}^{LH}(q(0),t(0)) > \Delta_{LL}^{HL}\left(q\left(\frac{p_{LL}}{p_{HL}}\right),t\left(\frac{p_{LL}}{p_{HL}}\right)\right)$ does not hold. In other words, $\Delta_{LL}^{LH}(q(0),t(0)) < \Delta_{LL}^{HL}\left(q\left(\frac{p_{LL}}{p_{HL}}\right),t\left(\frac{p_{LL}}{p_{HL}}\right)\right)$ or $\Delta_{LL}^{LH}(q(0),t(0)) = \Delta_{LL}^{HL}\left(q\left(\frac{p_{LL}}{p_{HL}}\right),t\left(\frac{p_{LL}}{p_{HL}}\right)\right)$ hold. We know from monotonicity of the cost difference that $\Delta_{LL}^{HL}\left(q\left(\frac{p_{LL}}{p_{HL}}\right),t\left(\frac{p_{LL}}{p_{HL}}\right)\right) < \Delta_{LL}^{HL}(q(0),t(0))$ always holds and thus $\Delta_{LL}^{LH}(q(0),t(0)) < \Delta_{LL}^{HL}(q(0),t(0))$. In this case, the LL type will have a strict preference to mimic the HL type, which contradicts our assumption in (A40). The same logic applies when $\Delta_{LL}^{LH}(q(0),t(0)) = \Delta_{LL}^{HL}\left(q\left(\frac{p_{LL}}{p_{HL}}\right),t\left(\frac{p_{LL}}{p_{HL}}\right)\right)$ holds. A similar argument proves the claim that $\Delta_{LL}^{LH}\left(q\left(\frac{p_{LL}}{p_{LH}}\right),t\left(\frac{p_{LL}}{p_{LH}}\right)\right) < \Delta_{LL}^{HL}(q(0),t(0))$ holds.

Therefore, if there is no large difference between $\Delta_{HL}^{HH}(q_{HH}^{sb},t_{HH}^{sb})$ and $\Delta_{LH}^{HH}(q_{HH}^{sb},t_{HH}^{sb})$, that is, if the value of k^* is within the following interval:

$$\Delta_{LL}^{LH}\left(q(\frac{p_{LL}}{p_{LH}}),t(\frac{p_{LL}}{p_{LH}})\right) - \Delta_{LL}^{HL}(q(0),t(0)) < k^* < \Delta_{LL}^{LH}(q(0),t(0)) - \Delta_{LL}^{HL}\left(q\left(\frac{p_{LL}}{p_{HL}}\right),t\left(\frac{p_{LL}}{p_{HL}}\right)\right)$$

then we always have $f(0) > 0$ and $f(1) < 0$ implying that $f(0)f(1) < 0$ and, by the Intermediate Value Theorem, there is certainly a $0 < \lambda^* < 1$ satisfying (A46). QED.

ii. *The LH type is more marginally inefficient than the HL type:* to analyse this case suppose that

$$\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} > \frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} \text{ for } z \in \{q,t\}$$

which implies $\Delta_{LH}^{HH}(q_{HH}^{sb},t_{HH}^{sb}) - \Delta_{HL}^{HH}(q_{HH}^{sb},t_{HH}^{sb}) > k^*$, where k^* is as in (A46), defined in the case (i), so $\Delta_{LH}^{HH}(q_{HH}^{sb},t_{HH}^{sb}) - \Delta_{HL}^{HH}(q_{HH}^{sb},t_{HH}^{sb})$ belongs to the interval (k^*,∞). This means that there is a constant $k \in (k^*,\infty)$ for which $\Delta_{LH}^{HH}(q_{HH}^{sb},t_{HH}^{sb}) - \Delta_{HL}^{HH}(q_{HH}^{sb},t_{HH}^{sb}) = k$ holds.

Then, in a similar way to (A47) we can write the following equation:

$$f(\lambda) = \Delta_{LL}^{LH}(q_{LH}^{sb},t_{LH}^{sb}) - \Delta_{LL}^{HL}(q_{HL}^{sb},t_{HL}^{sb}) - k = 0 \qquad \text{(A48)}$$

(A48) is the same as (A47), except that the constant is now k rather than k^* where $k \in (k^*,\infty)$. Thus, the same approach used to prove that $f(0) > 0$ and $f(1) < 0$ for (A47) can also be used for (A48). We avoid repeating this step and only present the acceptable range of k that guarantees the existence of $0 < \lambda^* < 1$ that satisfies (A48):

$$\Delta_{LL}^{LH}\left(q\left(\frac{p_{LL}}{p_{LH}}\right),t\left(\frac{p_{LL}}{p_{LH}}\right)\right) - \Delta_{LL}^{HL}(q(0),t(0)) < k^* < k < \Delta_{LL}^{LH}(q(0),t(0)) - \Delta_{LL}^{HL}\left(q\left(\frac{p_{LL}}{p_{HL}}\right),t\left(\frac{p_{LL}}{p_{HL}}\right)\right)$$

iii. *The HL type is more marginally inefficient than the LH type:* to analyse this case suppose that

$$\frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} < \frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} \text{ for } z \in \{q, t\}$$

This means $\Delta_{LH}^{HH}(q_{HH}^{sb}, t_{HH}^{sb}) - \Delta_{HL}^{HH}(q_{HH}^{sb}, t_{HH}^{sb}) < k^*$ where k^* is as in (A46) same constant as in case (i). In a similar manner to the previous case there is a $k' \in (-\infty, k^*)$ for which $\Delta_{LH}^{HH}(q_{HH}^{sb}, t_{HH}^{sb}) - \Delta_{HL}^{HH}(q_{HH}^{sb}, t_{HH}^{sb}) = k'$.

This means we need prove that there is a $0 < \lambda^* < 1$ that satifies equation below.

$$f(\lambda) = \Delta_{LL}^{LH}(q_{LH}^{sb}, t_{LH}^{sb}) - \Delta_{LL}^{HL}(q_{HL}^{sb}, t_{HL}^{sb}) - k' = 0$$

Repeating the previous proof procedure we show that the acceptable range of k' that guarantees the existence of $0 < \lambda^* < 1$ as a root to above equation is:

$$\Delta_{LL}^{LH}\left(q(\frac{p_{LL}}{p_{LH}}), t(\frac{p_{LL}}{p_{LH}})\right) - \Delta_{LL}^{HL}\big(q(0), t(0)\big) < k' < k^* < \Delta_{LL}^{LH}(q(0), t(0)) - \Delta_{LL}^{HL}\left(q(\frac{p_{LL}}{p_{HL}}), t(\frac{p_{LL}}{p_{HL}})\right)$$

The above results show that irrespective of how marginal cost differences are (i.e., $\frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} = \frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}}$ or $\frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} > \frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}}$ or $\frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} < \frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}}$), there is always a λ that satisfies (A40). The condition for the existence of λ across three aforementioned subcases can be compactly presented as follows:

$$\Delta_{LL}^{LH}\left(q\left(\frac{p_{LL}}{p_{LH}}\right), t\left(\frac{p_{LL}}{p_{LH}}\right)\right) - \Delta_{LL}^{HL}\big(q(0), t(0)\big) < k' < k^* < k < \Delta_{LL}^{LH}(q(0), t(0)) - \Delta_{LL}^{HL}\left(q\left(\frac{p_{LL}}{p_{HL}}\right), t\left(\frac{p_{LL}}{p_{HL}}\right)\right)$$

Case 3:
Let $\lambda_1 = 1, \lambda_2 = 0, \lambda_3 = 0$ in (A28), (A29), (A30) to obtain equations (9a-9c) in the main text.

Case 4:
Let $\lambda_1 = 0, \lambda_2 = 1, \lambda_3 = 0$ in (A28), (A29), (A30) to obtain equations (10a-10c) in the main text.

A4. Equations for the simulation in Section 4

Consider the following Cobb-Douglas specification

$$v(q, t) = A{q_{ij}}^{\alpha}{t_{ij}}^{1-\alpha}$$

$$c(q, t, \theta_q, \theta_t) = \left({\theta_q}^{i}\right)^{\beta}\left({\theta_t}^{j}\right)^{1-\beta}\left(q_{ij}t_{ij}\right)^{\gamma}$$

A4.1 Output levels

1. *For the inefficient (HH) type:*

From:

$$\frac{\partial v(q_{HH},t_{HH})}{\partial z_{HH}} = \frac{\partial c\left(q_{HL},t_{HL},\theta_q^H,\theta_t^H\right)}{\partial z_{HH}} + \frac{(\lambda_1\, p_{LL} +\, p_{LH})}{p_{HH}}\frac{\partial \Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + \frac{(\lambda_2\, p_{LL} +\, p_{HL})}{p_{HH}}\frac{\partial \Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + \frac{\lambda_3\, p_{LL}}{p_{HH}}\frac{\partial \Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}$$

for $z \in \{q,t\}$

- q_{HH}:

With respect to q_{HH}:

$$\begin{aligned} A\alpha {q_{HH}}^{\alpha-1}{t_{HH}}^{1-\alpha} &= (q_{HH}t_{HH})^{\gamma-1}t_{HH}\,\gamma\left(\theta_q^H\right)^{\beta}\,(\theta_t^H)^{1-\beta} \\ &+ \frac{(\lambda_1\, p_{LL} +\, p_{LH})}{p_{HH}}\left((q_{HH}t_{HH})^{\gamma-1}t_{HH}\,\gamma\left(\left(\theta_q^H\right)^{\beta}\,(\theta_t^H)^{1-\beta} - \left(\theta_q^L\right)^{\beta}\,(\theta_t^H)^{1-\beta}\right)\right) \\ &+ \frac{(\lambda_2\, p_{LL} +\, p_{HL})}{p_{HH}}\left((q_{HH}t_{HH})^{\gamma-1}t_{HH}\,\gamma\left(\left(\theta_q^H\right)^{\beta}\,(\theta_t^H)^{1-\beta} - \left(\theta_q^H\right)^{\beta}\,(\theta_t^L)^{1-\beta}\right)\right) \\ &+ \frac{\lambda_3\, p_{LL}}{p_{HH}}\left((q_{HH}t_{HH})^{\gamma-1}t_{HH}\,\gamma\left(\left(\theta_q^H\right)^{\beta}\,(\theta_t^H)^{1-\beta} - \left(\theta_q^L\right)^{\beta}\,(\theta_t^L)^{1-\beta}\right)\right) \end{aligned}$$

Let:

$$K_1 = \gamma\left(\theta_q^H\right)^{\beta}\,(\theta_t^H)^{1-\beta}$$

$$K_2 = \gamma\left(\left(\theta_q^H\right)^{\beta}\,(\theta_t^H)^{1-\beta} - \left(\theta_q^L\right)^{\beta}\,(\theta_t^H)^{1-\beta}\right)$$

$$K_3 = \gamma\left(\left(\theta_q^H\right)^{\beta}\,(\theta_t^H)^{1-\beta} - \left(\theta_q^H\right)^{\beta}\,(\theta_t^L)^{1-\beta}\right)$$

$$K_4 = \gamma\left(\left(\theta_q^H\right)^{\beta}\,(\theta_t^H)^{1-\beta} - \left(\theta_q^L\right)^{\beta}\,(\theta_t^L)^{1-\beta}\right)$$

The expression is now:

$$\begin{aligned} A\alpha {q_{HH}}^{\alpha-1}{t_{HH}}^{1-\alpha} &= (q_{HH}t_{HH})^{\gamma-1}t_{HH}\,K_1 + \frac{(\lambda_1\, p_{LL} +\, p_{LH})}{p_{HH}}(q_{HH}t_{HH})^{\gamma-1}t_{HH}\,K_2 \\ &+ \frac{(\lambda_2\, p_{LL} +\, p_{HL})}{p_{HH}}(q_{HH}t_{HH})^{\gamma-1}t_{HH}\,K_3 + \frac{\lambda_3\, p_{LL}}{p_{HH}}(q_{HH}t_{HH})^{\gamma-1}t_{HH}\,K_4 \end{aligned}$$

From which a closed form solution for q_{HH} can be derived:

$$q_{HH} = \left[\frac{1}{A\alpha}\left(K_1 + \frac{(\lambda_1\, p_{LL} + p_{LH})}{p_{HH}}K_2 + \frac{(\lambda_2\, p_{LL} + p_{HL})}{p_{HH}}\,K_3 + \frac{\lambda_3\, p_{LL}}{p_{HH}}K_4\right)\right]^{\frac{1}{\alpha-\gamma}} {t_{HH}}^{\frac{\gamma+\alpha-1}{\alpha-\gamma}} \quad \text{(A48)}$$

- t_{HH}:

Likewise, with respect to t_{HH}:

$$A(1-\alpha){q_{HH}}^{\alpha}{t_{HH}}^{-\alpha} = (q_{HH}t_{HH})^{\gamma-1}q_{HH}\,\gamma\left(\theta_q^H\right)^{\beta}\,(\theta_t^H)^{1-\beta} + \frac{(\lambda_1\,p_{LL}+\,p_{LH})}{p_{HH}}\left((q_{HH}t_{HH})^{\gamma-1}q_{HH}\,\gamma\left(\left(\theta_q^H\right)^{\beta}\,(\theta_t^H)^{1-\beta}-\left(\theta_q^L\right)^{\beta}\,(\theta_t^H)^{1-\beta}\right)\right) + \frac{(\lambda_2\,p_{LL}+\,p_{HL})}{p_{HH}}\left((q_{HH}t_{HH})^{\gamma-1}q_{HH}\,\gamma\left(\left(\theta_q^H\right)^{\beta}\,(\theta_t^H)^{1-\beta}-\left(\theta_q^H\right)^{\beta}\,(\theta_t^L)^{1-\beta}\right)\right) + \frac{\lambda_3\,p_{LL}}{p_{HH}}\left((q_{HH}t_{HH})^{\gamma-1}q_{HH}\,\gamma\left(\left(\theta_q^H\right)^{\beta}\,(\theta_t^H)^{1-\beta}-\left(\theta_q^L\right)^{\beta}\,(\theta_t^L)^{1-\beta}\right)\right)$$

Recall that:

$$K_1 = \gamma\left(\theta_q^H\right)^{\beta}\,(\theta_t^H)^{1-\beta}$$

$$K_2 = \gamma\left(\left(\theta_q^H\right)^{\beta}\,(\theta_t^H)^{1-\beta}-\left(\theta_q^L\right)^{\beta}\,(\theta_t^H)^{1-\beta}\right)$$

$$K_3 = \gamma\left(\left(\theta_q^H\right)^{\beta}\,(\theta_t^H)^{1-\beta}-\left(\theta_q^H\right)^{\beta}\,(\theta_t^L)^{1-\beta}\right)$$

$$K_4 = \gamma\left(\left(\theta_q^H\right)^{\beta}\,(\theta_t^H)^{1-\beta}-\left(\theta_q^L\right)^{\beta}\,(\theta_t^L)^{1-\beta}\right)$$

The expression is now:

$$A(1-\alpha){q_{HH}}^{\alpha}{t_{HH}}^{-\alpha} = (q_{HH}t_{HH})^{\gamma-1}q_{HH}\,K_1 + \frac{(\lambda_1\,p_{LL}+p_{LH})}{p_{HH}}(q_{HH}t_{HH})^{\gamma-1}q_{HH}\,K_2 + \frac{(\lambda_2\,p_{LL}+p_{HL})}{p_{HH}}(q_{HH}t_{HH})^{\gamma-1}q_{HH}K_3 + \frac{\lambda_3\,p_{LL}}{p_{HH}}(q_{HH}t_{HH})^{\gamma-1}q_{HH}\,K_4$$

From which a closed form solution for t_{HH} can be derived:

$$t_{HH} = \left[\frac{1}{A(1-\alpha)}\left(K_1 + \frac{(\lambda_1\,p_{LL}+p_{LH})}{p_{HH}}K_2 + \frac{(\lambda_2\,p_{LL}+p_{HL})}{p_{HH}}\,K_3 + \frac{\lambda_3\,p_{LL}}{p_{HH}}K_4\right)\right]^{\frac{1}{1-\gamma-\alpha}} {q_{HH}}^{\frac{\gamma-\alpha}{1-\gamma-\alpha}} \quad \text{(A49)}$$

2. *For the middle (LH) type:*

From:

$$\frac{\partial v(q_{LH},t_{LH})}{\partial z_{LH}} = \frac{\partial c\left(q_{LH},t_{LH},\theta_q^L,\theta_t^H\right)}{\partial z_{LH}} + \lambda_1\frac{p_{LL}}{p_{LH}}\frac{\partial \Delta_{LL}^{LH}(q_{LH},t_{LH})}{\partial z_{LH}}$$

for $z \in \{q,t\}$

- q_{LH}:

With respect to q_{LH}:

$$A\alpha{q_{LH}}^{\alpha-1}{t_{LH}}^{1-\alpha} = (q_{LH}t_{LH})^{\gamma-1}t_{LH}\,\gamma\left(\theta_q^L\right)^{\beta}\,(\theta_t^H)^{1-\beta} + \left(\lambda_1\frac{p_{LL}}{p_{LH}}\right)(q_{LH}t_{LH})^{\gamma-1}t_{LH}\,\gamma\left[\left(\theta_q^L\right)^{\beta}\,(\theta_t^H)^{1-\beta} - \left(\theta_q^L\right)^{\beta}\,(\theta_t^L)^{1-\beta}\right]$$

Let:

$$Q_1 = \gamma\left(\theta_q^L\right)^{\beta}\,(\theta_t^H)^{1-\beta}$$

$$Q_2 = \gamma\left[\left(\theta_q^L\right)^\beta \left(\theta_t^H\right)^{1-\beta} - \left(\theta_q^L\right)^\beta \left(\theta_t^L\right)^{1-\beta}\right]$$

and we have:

$$A\alpha {q_{LH}}^{\alpha-1} {t_{LH}}^{1-\alpha} = (q_{LH}t_{LH})^{\gamma-1} t_{LH}\, Q_1 + \left(\lambda_1 \frac{p_{LL}}{p_{LH}}\right)(q_{LH}t_{LH})^{\gamma-1} t_{LH}\, Q_2$$

A closed form solution for q_{LH} is:

$$q_{LH} = \left[\frac{1}{A\alpha}\left(Q_1 + \left(\lambda_1 \frac{p_{LL}}{p_{LH}}\right) Q_2\right)\right]^{\frac{1}{\alpha-\gamma}} {t_{LH}}^{\frac{\gamma+\alpha-1}{\alpha-\gamma}} \quad \text{(A50)}$$

- t_{LH}:

With respect to t_{LH}:

$$A(1-\alpha){q_{LH}}^{\alpha}{t_{LH}}^{-\alpha} = (q_{LH}t_{LH})^{\gamma-1} q_{LH}\, \gamma\left(\theta_q^L\right)^\beta \left(\theta_t^H\right)^{1-\beta} + \left(\lambda_1 \frac{p_{LL}}{p_{LH}}\right)(q_{LH}t_{LH})^{\gamma-1} q_{LH}\, \gamma\left[\left(\theta_q^L\right)^\beta \left(\theta_t^H\right)^{1-\beta} - \left(\theta_q^L\right)^\beta \left(\theta_t^L\right)^{1-\beta}\right]$$

Recall that:

$$Q_1 = \gamma\left(\theta_q^L\right)^\beta \left(\theta_t^H\right)^{1-\beta}$$

$$Q_2 = \gamma\left[\left(\theta_q^L\right)^\beta \left(\theta_t^H\right)^{1-\beta} - \left(\theta_q^L\right)^\beta \left(\theta_t^L\right)^{1-\beta}\right]$$

to obtain:

$$A(1-\alpha){q_{LH}}^{\alpha}{t_{LH}}^{-\alpha} = (q_{LH}t_{LH})^{\gamma-1} q_{LH}\, Q_1 + \left(\lambda_1 \frac{p_{LL}}{p_{LH}}\right)(q_{LH}t_{LH})^{\gamma-1} q_{LH}\, Q_2$$

A closed form solution for t_{LH} is:

$$t_{LH} = \left[\frac{1}{A(1-\alpha)}\left(Q_1 + \left(\lambda_1 \frac{p_{LL}}{p_{LH}}\right) Q_2\right)\right]^{\frac{1}{1-\gamma-\alpha}} {q_{LH}}^{\frac{\gamma-\alpha}{1-\gamma-\alpha}} \quad \text{(A51)}$$

3. *For the middle (HL) type:*

From:

$$\frac{\partial v(q_{HL}, t_{HL})}{\partial z_{HL}} = \frac{\partial c\left(q_{HL}, t_{HL}, \theta_q^H, \theta_t^L\right)}{\partial z_{HL}} + \lambda_2 \frac{p_{LL}}{p_{HL}} \frac{\partial \Delta_{LL}^{HL}(q_{HL}, t_{HL})}{\partial z_{HL}}$$

for $z \in \{q, t\}$

- q_{HL}:

$$A\alpha {q_{HL}}^{\alpha-1} {t_{HL}}^{1-\alpha} = (q_{HL}t_{HL})^{\gamma-1} t_{HL}\, \gamma\left(\theta_q^H\right)^\beta \left(\theta_t^L\right)^{1-\beta} + \left(\lambda_2 \frac{p_{LL}}{p_{HL}}\right)(q_{HL}t_{HL})^{\gamma-1} t_{HL}\, \gamma\left[\left(\theta_q^H\right)^\beta \left(\theta_t^L\right)^{1-\beta} - \left(\theta_q^L\right)^\beta \left(\theta_t^L\right)^{1-\beta}\right]$$

Let:

$$R_1 = \gamma\left(\theta_q^H\right)^\beta \left(\theta_t^L\right)^{1-\beta}$$

$$R_2 = \gamma\left[\left(\theta_q^H\right)^\beta \left(\theta_t^L\right)^{1-\beta} - \left(\theta_q^L\right)^\beta \left(\theta_t^L\right)^{1-\beta}\right]$$

A closed form solution for q_{HL} is:

$$q_{HL} = \left[\frac{1}{A\alpha}\left(R_1 + \left(\lambda_2 \frac{p_{LL}}{p_{HL}}\right) R_2\right)\right]^{\frac{1}{\alpha-\gamma}} t_{HL}{}^{\frac{\gamma+\alpha-1}{\alpha-\gamma}} \quad \text{(A52)}$$

- t_{HL}:

$$\begin{aligned} A(1-\alpha) q_{HL}{}^{\alpha} t_{LH}{}^{-\alpha} &= (q_{HL} t_{HL})^{\gamma-1} q_{HL}\, \gamma \left(\theta_q^H\right)^\beta \left(\theta_t^L\right)^{1-\beta} \\ &+ \left(\lambda_2 \frac{p_{LL}}{p_{HL}}\right) (q_{HL} t_{HL})^{\gamma-1} q_{HL}\, \gamma \left[\left(\theta_q^H\right)^\beta \left(\theta_t^L\right)^{1-\beta} - \left(\theta_q^L\right)^\beta \left(\theta_t^L\right)^{1-\beta}\right] \end{aligned}$$

Recall that:

$$R_1 = \gamma\left(\theta_q^H\right)^\beta \left(\theta_t^L\right)^{1-\beta}$$

$$R_2 = \gamma\left[\left(\theta_q^H\right)^\beta \left(\theta_t^L\right)^{1-\beta} - \left(\theta_q^L\right)^\beta \left(\theta_t^L\right)^{1-\beta}\right]$$

$$A(1-\alpha) q_{HL}{}^{\alpha} t_{LH}{}^{-\alpha} = (q_{HL} t_{HL})^{\gamma-1} q_{HL}\, R_1 + \left(\lambda_2 \frac{p_{LL}}{p_{HL}}\right) (q_{HL} t_{HL})^{\gamma-1} q_{HL}\, R_2$$

A closed form solution for t_{HL} is:

$$t_{HL} = \left[\frac{1}{A(1-\alpha)}\left(R_1 + \left(\lambda_2 \frac{p_{LL}}{p_{HL}}\right) R_2\right)\right]^{\frac{1}{1-\gamma-\alpha}} q_{HL}{}^{\frac{\gamma-\alpha}{1-\gamma-\alpha}} \quad \text{(A53)}$$

4. *For the efficient (LL) type:*

From :

$$p_{LL}\left[\frac{\partial v(q_{LL}, t_{LL})}{\partial z_{LL}} - \frac{\partial c\left(q_{LL}, t_{LL}, \theta_q{}^L, \theta_t{}^L\right)}{\partial z_{LL}}\right] = 0$$

for $z \in \{q, t\}$

- q_{LL}:

$$A\alpha q_{LL}{}^{\alpha-1} t_{LL}{}^{1-\alpha} = (q_{LL} t_{LL})^{\gamma-1} t_{LL}\, \gamma \left(\theta_q^L\right)^\beta \left(\theta_t^L\right)^{1-\beta}$$

Let:

$$S_1 = \gamma\left(\theta_q^H\right)^\beta \left(\theta_t^L\right)^{1-\beta}$$

to obtain

$$q_{LL} = \left[\frac{1}{A\alpha} S_1\right]^{\frac{1}{\alpha-\gamma}} t_{LL}{}^{\frac{\gamma+\alpha-1}{\alpha-\gamma}} \quad \text{(A54)}$$

- t_{LL}:

$$A(1-\alpha) q_{LL}{}^{\alpha} t_{LL}{}^{-\alpha} = (q_{LL} t_{LL})^{\gamma-1} q_{LL}\, \gamma \left(\theta_q^L\right)^\beta \left(\theta_t^L\right)^{1-\beta}$$

Recall that:

$$S_1 = \gamma\left(\theta_q^H\right)^{\beta}\left(\theta_t^L\right)^{1-\beta}$$

to obtain:

$$t_{LL} = \left[\frac{1}{A(1-\alpha)} S_1\right]^{\frac{1}{1-\gamma-\alpha}} q_{LL}^{\frac{\gamma-\alpha}{1-\gamma-\alpha}} \quad \text{(A55)}$$

A4.2 Cases

Case 1: Positive correlation

<u>Focus on the λs:</u>

The solution to the system of equations is:

$$\lambda_1 = \frac{p_{LH}\left(p_{LH}\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + p_{HL}\frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + p_{HH}p_{LL}\frac{\partial\Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}\right)}{p_{LL}\left(p_{HH} - p_{HL}\frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} - p_{LH}\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + \frac{\partial\Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}(p_{HH}p_{HL}) + \frac{\partial\Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}(p_{HH}p_{LH})\right)}$$

More compactly:

$$\lambda_1 = \frac{p_{LH}\left(p_{LH}\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + p_{HL}\frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + p_{HH}p_{LL}\frac{\partial\Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}\right)}{p_{LL}\left(p_{HH}\left(1 + \frac{\partial\Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}(p_{HL} + p_{LH})\right) - p_{LH}\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} - p_{HL}\frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}\right)}$$

$$\lambda_2 = \frac{p_{HL}\left(p_{LH}\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + p_{HL}\frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + p_{HH}p_{LL}\frac{\partial\Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}\right)}{p_{LL}\left(p_{HH} - p_{HL}\frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} - p_{LH}\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + \frac{\partial\Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}(p_{HH}p_{HL}) + \frac{\partial\Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}(p_{HH}p_{LH})\right)}$$

More compactly:

$$\lambda_2 = \frac{p_{HL}\left(p_{LH}\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + p_{HL}\frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + p_{HH}p_{LL}\frac{\partial\Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}\right)}{p_{LL}\left(p_{HH}\left(1 + \frac{\partial\Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}(p_{HL} + p_{LH})\right) - p_{LH}\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} - p_{HL}\frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}\right)}$$

$$\lambda_3 = \frac{\left(p_{HH}\,p_{LL} - (p_{LH}p_{LL})\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} - (p_{HL}\,p_{LH})\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} - (p_{LH}{}^2)\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} - (p_{HL}\,p_{LL})\frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} - (p_{HL}\,p_{LH})\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} - (p_{HL}{}^2)\frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}\right)}{p_{LL}\left(p_{HH} - p_{HL}\frac{\partial\Delta_{HL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} - p_{LH}\frac{\partial\Delta_{LH}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}} + \frac{\partial\Delta_{LL}^{HH}(q_{HH},t_{HH})}{\partial z_{HH}}(p_{HH}{}^2 p_{HL}\,p_{LH})\right)}$$

More compactly:

$$\lambda_3 = \frac{\left(p_{HH}\, p_{LL} - \frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}}\, p_{LH}(\, p_{LL} + \, p_{LH} + \, p_{HL}) - \frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}}\, p_{HL}(\, p_{LL} + \, p_{LH} + \, p_{HL}) \right)}{p_{LL}\left(p_{HH}\left(1 + \frac{\partial \Delta_{LL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}}(\, p_{HL} + \, p_{LH})\right) - \, p_{LH}\frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}} - \, p_{HL}\frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}}\right)}$$

Now we derive the Cobb-Douglas instances of $\lambda_1, \lambda_2, \lambda_3$.

1. Consider the **denominator** in the expressions for$\lambda_1, \lambda_2, \lambda_3$:

The full expressions are:

- With respect to q_{HH}:

$$p_{LL}\Bigg(p_{HH}\Bigg(1 + \Big((q_{HH}t_{HH})^{\gamma-1} t_{HH}\, \gamma \left((\theta_q^H)^{\beta}\, (\theta_t^H)^{1-\beta} - (\theta_q^L)^{\beta}\, (\theta_t^L)^{1-\beta} \right)\Big)(\, p_{HL} + \, p_{LH})\Bigg)$$
$$- \, p_{LH}\left((q_{HH}t_{HH})^{\gamma-1} t_{HH}\, \gamma \left((\theta_q^H)^{\beta}\, (\theta_t^H)^{1-\beta} - (\theta_q^L)^{\beta}\, (\theta_t^H)^{1-\beta} \right)\right)$$
$$- \, p_{HL}\left((q_{HH}t_{HH})^{\gamma-1} t_{HH}\, \gamma \left((\theta_q^H)^{\beta}\, (\theta_t^H)^{1-\beta} - (\theta_q^H)^{\beta}\, (\theta_t^L)^{1-\beta} \right)\right)\Bigg)$$

Let:

$$x = (q_{HH}t_{HH})^{\gamma-1} t_{HH}$$

Recall that:

$$K_2 = \gamma\left((\theta_q^H)^{\beta}\, (\theta_t^H)^{1-\beta} - (\theta_q^L)^{\beta}\, (\theta_t^H)^{1-\beta} \right)$$

$$K_3 = \gamma\left((\theta_q^H)^{\beta}\, (\theta_t^H)^{1-\beta} - (\theta_q^H)^{\beta}\, (\theta_t^L)^{1-\beta} \right)$$

$$K_4 = \gamma\left((\theta_q^H)^{\beta}\, (\theta_t^H)^{1-\beta} - (\theta_q^L)^{\beta}\, (\theta_t^L)^{1-\beta} \right)$$

The denominator for $\lambda_1, \lambda_2, \lambda_3$ with respect to q_{HH}:

$$p_{LL}\left(p_{HH}\big(1 + (xK_4)(\, p_{HL} + \, p_{LH})\big) - \, p_{LH}(x\, K_2) - \, p_{HL}(xK_3)\right)$$

Or more compactly:

$$p_{LL}(\, p_{HH}(1 + K_4(\, p_{HL} + \, p_{LH})x) - (\, p_{LH}K_2 + \, p_{HL}K_3)x)$$

- With respect to t_{HH}:

$$p_{LL}\Bigg(p_{HH}\Bigg(1+\Big((q_{HH}t_{HH})^{\gamma-1}q_{HH}\,\gamma\left(\left(\theta_q^H\right)^{\beta}\,(\theta_t^H)^{1-\beta}-\left(\theta_q^L\right)^{\beta}\,(\theta_t^L)^{1-\beta}\right)\Big)(p_{HL}+p_{LH})\Bigg)$$
$$-\,p_{LH}\Big((q_{HH}t_{HH})^{\gamma-1}q_{HH}\,\gamma\left(\left(\theta_q^H\right)^{\beta}\,(\theta_t^H)^{1-\beta}-\left(\theta_q^L\right)^{\beta}\,(\theta_t^H)^{1-\beta}\right)\Big)$$
$$-\,p_{HL}\Bigg((q_{HH}t_{HH})^{\gamma-1}q_{HH}\,\gamma\left(\left(\theta_q^H\right)^{\beta}\,(\theta_t^H)^{1-\beta}-\left(\theta_q^H\right)^{\beta}\,(\theta_t^L)^{1-\beta}\right)\Bigg)\Bigg)$$

Let:

$$y=(q_{HH}t_{HH})^{\gamma-1}q_{HH}$$

Recall that:

$$K_2=\gamma\left(\left(\theta_q^H\right)^{\beta}\,(\theta_t^H)^{1-\beta}-\left(\theta_q^L\right)^{\beta}\,(\theta_t^H)^{1-\beta}\right)$$

$$K_3=\gamma\left(\left(\theta_q^H\right)^{\beta}\,(\theta_t^H)^{1-\beta}-\left(\theta_q^H\right)^{\beta}\,(\theta_t^L)^{1-\beta}\right)$$

$$K_4=\gamma\left(\left(\theta_q^H\right)^{\beta}\,(\theta_t^H)^{1-\beta}-\left(\theta_q^L\right)^{\beta}\,(\theta_t^L)^{1-\beta}\right)$$

The denominator for $\lambda_1,\lambda_2,\lambda_3$ with respect to t_{HH}:

$$p_{LL}\left(p_{HH}\big(1+(yK_4)(p_{HL}+p_{LH})\big)-p_{LH}(y\,K_2)-p_{HL}(yK_3)\right)$$

Or more compactly:

$$p_{LL}(p_{HH}(1+K_4(p_{HL}+p_{LH})y)-(p_{LH}K_2+p_{HL}K_3)y)$$

2. Consider the **numerator** in the expressions for λ_1,λ_2:

With respect to q_{HH}:

- The numerator of λ_1 is:

$$p_{LH}xT_1$$

- The numerator of λ_2 is:

$$p_{HL}xT_1$$

where $x=(q_{HH}t_{HH})^{\gamma-1}t_{HH}$

With respect to t_{HH}:

- The numerator of λ_1 is:

$$p_{LH}(q_{HH}t_{HH})^{\gamma-1}q_{HH}T_1$$

- The numerator of λ_2 is:

$$p_{HL}(q_{HH}t_{HH})^{\gamma-1}q_{HH}T_1$$

where $y=(q_{HH}t_{HH})^{\gamma-1}q_{HH}$

$$T_1 = \gamma \left[p_{LH} \left(\left(\theta_q^H\right)^\beta (\theta_t^H)^{1-\beta} - \left(\theta_q^L\right)^\beta (\theta_t^H)^{1-\beta} \right) + p_{HL} \left(\left(\theta_q^H\right)^\beta (\theta_t^H)^{1-\beta} - \left(\theta_q^H\right)^\beta (\theta_t^L)^{1-\beta} \right) + p_{HH} p_{LL} \left(\left(\theta_q^H\right)^\beta (\theta_t^H)^{1-\beta} - \left(\theta_q^L\right)^\beta (\theta_t^L)^{1-\beta} \right) \right]$$

Recall that the full expressions are:

With respect to q_{HH}:

- The numerator of λ_1 is:

$$p_{LH} (q_{HH} t_{HH})^{\gamma-1} t_{HH} \gamma \left[p_{LH} \left(\left(\theta_q^H\right)^\beta (\theta_t^H)^{1-\beta} - \left(\theta_q^L\right)^\beta (\theta_t^H)^{1-\beta} \right) + p_{HL} \left(\left(\theta_q^H\right)^\beta (\theta_t^H)^{1-\beta} - \left(\theta_q^H\right)^\beta (\theta_t^L)^{1-\beta} \right) + p_{HH} p_{LL} \left(\left(\theta_q^H\right)^\beta (\theta_t^H)^{1-\beta} - \left(\theta_q^L\right)^\beta (\theta_t^L)^{1-\beta} \right) \right]$$

- The numerator of λ_2 is:

$$p_{HL} (q_{HH} t_{HH})^{\gamma-1} t_{HH} \gamma \left[p_{LH} \left(\left(\theta_q^H\right)^\beta (\theta_t^H)^{1-\beta} - \left(\theta_q^L\right)^\beta (\theta_t^H)^{1-\beta} \right) + p_{HL} \left(\left(\theta_q^H\right)^\beta (\theta_t^H)^{1-\beta} - \left(\theta_q^H\right)^\beta (\theta_t^L)^{1-\beta} \right) + p_{HH} p_{LL} \left(\left(\theta_q^H\right)^\beta (\theta_t^H)^{1-\beta} - \left(\theta_q^L\right)^\beta (\theta_t^L)^{1-\beta} \right) \right]$$

With respect to t_{HH}:

- The numerator of λ_1 is:

$$p_{LH} (q_{HH} t_{HH})^{\gamma-1} q_{HH} \gamma \left[p_{LH} \left(\left(\theta_q^H\right)^\beta (\theta_t^H)^{1-\beta} - \left(\theta_q^L\right)^\beta (\theta_t^H)^{1-\beta} \right) + p_{HL} \left(\left(\theta_q^H\right)^\beta (\theta_t^H)^{1-\beta} - \left(\theta_q^H\right)^\beta (\theta_t^L)^{1-\beta} \right) + p_{HH} p_{LL} \left(\left(\theta_q^H\right)^\beta (\theta_t^H)^{1-\beta} - \left(\theta_q^L\right)^\beta (\theta_t^L)^{1-\beta} \right) \right]$$

- The numerator of λ_2 is:

$$p_{HL} (q_{HH} t_{HH})^{\gamma-1} q_{HH} \gamma \left[p_{LH} \left(\left(\theta_q^H\right)^\beta (\theta_t^H)^{1-\beta} - \left(\theta_q^L\right)^\beta (\theta_t^H)^{1-\beta} \right) + p_{HL} \left(\left(\theta_q^H\right)^\beta (\theta_t^H)^{1-\beta} - \left(\theta_q^H\right)^\beta (\theta_t^L)^{1-\beta} \right) + p_{HH} p_{LL} \left(\left(\theta_q^H\right)^\beta (\theta_t^H)^{1-\beta} - \left(\theta_q^L\right)^\beta (\theta_t^L)^{1-\beta} \right) \right]$$

3. Consider the **numerator** in the expression for λ_3:

$$\left(p_{HH}\, p_{LL} - \frac{\partial \Delta_{LH}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}}\, p_{LH}(p_{LL} + p_{LH} + p_{HL}) - \frac{\partial \Delta_{HL}^{HH}(q_{HH}, t_{HH})}{\partial z_{HH}}\, p_{HL}(1 - p_{HH}) \right)$$

The full expression is:

- With respect to q_{HH}:

$$\left(p_{HH}\, p_{LL} - \left((q_{HH} t_{HH})^{\gamma-1} t_{HH} \gamma \left(\left(\theta_q^H\right)^\beta (\theta_t^H)^{1-\beta} - \left(\theta_q^L\right)^\beta (\theta_t^H)^{1-\beta} \right) \right) p_{LH}(1 - p_{HH}) - \left((q_{HH} t_{HH})^{\gamma-1} t_{HH} \gamma \left(\left(\theta_q^H\right)^\beta (\theta_t^H)^{1-\beta} - \left(\theta_q^H\right)^\beta (\theta_t^L)^{1-\beta} \right) \right) p_{HL}(1 - p_{HH}) \right)$$

Recall that:

$x = (q_{HH} t_{HH})^{\gamma-1} t_{HH}$

$K_2 = \gamma \left(\left(\theta_q^H\right)^\beta (\theta_t^H)^{1-\beta} - \left(\theta_q^L\right)^\beta (\theta_t^H)^{1-\beta} \right)$

$K_3 = \gamma\left(\left(\theta_q^H\right)^\beta (\theta_t^H)^{1-\beta} - \left(\theta_q^H\right)^\beta (\theta_t^L)^{1-\beta}\right)$

And substitute to obtain:

$$\left(p_{HH}\, p_{LL} - (xK_2)\, p_{LH}(\, p_{LL} +\, p_{LH} +\, p_{HL}) - (xK3)\, p_{HL}(1 -\, p_{HH})\right)$$

- With respect to t_{HH}:

$$\left(p_{HH}\, p_{LL} - \left((q_{HH}t_{HH})^{\gamma-1}q_{HH}\,\gamma\left(\left(\theta_q^H\right)^\beta (\theta_t^H)^{1-\beta} - \left(\theta_q^L\right)^\beta (\theta_t^H)^{1-\beta}\right)\right) p_{LH}(1 -\, p_{HH}) - \left((q_{HH}t_{HH})^{\gamma-1}q_{HH}\,\gamma\left(\left(\theta_q^H\right)^\beta (\theta_t^H)^{1-\beta} - \left(\theta_q^H\right)^\beta (\theta_t^L)^{1-\beta}\right)\right) p_{HL}(1 -\, p_{HH})\right)$$

Recall that:

$y = (q_{HH}t_{HH})^{\gamma-1}q_{HH}$

$K_2 = \gamma\left(\left(\theta_q^H\right)^\beta (\theta_t^H)^{1-\beta} - \left(\theta_q^L\right)^\beta (\theta_t^H)^{1-\beta}\right)$

$K_3 = \gamma\left(\left(\theta_q^H\right)^\beta (\theta_t^H)^{1-\beta} - \left(\theta_q^H\right)^\beta (\theta_t^L)^{1-\beta}\right)$

$\left(p_{HH}\, p_{LL} - (yK_2)\, p_{LH}(\, p_{LL} +\, p_{LH} +\, p_{HL}) - (yK_3)\, p_{HL}(1 -\, p_{HH})\right)$

Finally, the λs are:

- With respect to q_{HH}:

$$\lambda_1 = \frac{p_{LH}xT_1}{p_{LL}(\, p_{HH}(1 + K_4(\, p_{HL} +\, p_{LH})x) - (\, p_{LH}K_2 +\, p_{HL}K_3)x)}$$

$$\lambda_2 = \frac{p_{HL}xT_1}{p_{LL}(\, p_{HH}(1 + K_4(\, p_{HL} +\, p_{LH})x) - (\, p_{LH}K_2 +\, p_{HL}K_3)x)}$$

$$\lambda_3 = \frac{p_{HH}\, p_{LL} - (xK_2)\, p_{LH}(1 -\, p_{HH}) - (xK_3)\, p_{HL}(1 -\, p_{HH})}{p_{LL}(\, p_{HH}(1 + K_4(\, p_{HL} +\, p_{LH})x) - (\, p_{LH}K_2 +\, p_{HL}K_3)x)}$$

- With respect to t_{HH}:

$$\lambda_1 = \frac{p_{LH}yT_1}{p_{LL}(\, p_{HH}(1 + K_4(\, p_{HL} +\, p_{LH})y) - (\, p_{LH}K_2 +\, p_{HL}K_3)y)}$$

$$\lambda_2 = \frac{p_{HL}yT_1}{p_{LL}(\, p_{HH}(1 + K_4(\, p_{HL} +\, p_{LH})y) - (\, p_{LH}K_2 +\, p_{HL}K_3)y)}$$

$$\lambda_3 = \frac{p_{HH}\, p_{LL} - (yK_2)\, p_{LH}(1 -\, p_{HH}) - (yK3)\, p_{HL}(1 -\, p_{HH})}{p_{LL}(\, p_{HH}(1 + K_4(\, p_{HL} +\, p_{LH})y) - (\, p_{LH}K_2 +\, p_{HL}K_3)y)}$$

<u>Focus on the output levels:</u>

The HH type:

- q_{HH}:

To obtain q_{HH} numerically, plug the λs and substitute for x into:

$$A\alpha {q_{HH}}^{\alpha-1}{t_{HH}}^{1-\alpha} = x\,K_1 + \frac{\left(\frac{p_{LH}xT_1}{p_{LL}(p_{HH}(1+K_4(p_{HL}+p_{LH})x)-(p_{LH}K_2+p_{HL}K_3)x)}\,p_{LL}+p_{LH}\right)}{p_{HH}}x\,K_2 + \frac{\left(\frac{p_{HL}xT_1}{p_{LL}(p_{HH}(1+K_4(p_{HL}+p_{LH})x)-(p_{LH}K_2+p_{HL}K_3)x)}\,p_{LL}+p_{HL}\right)}{p_{HH}}x\,K_3 + \frac{\frac{p_{HH}\,p_{LL}-(xK_2)\,p_{LH}(1-p_{HH})-(xK_3)\,p_{HL}(1-p_{HH})}{p_{LL}(p_{HH}(1+K_4(p_{HL}+p_{LH})x)-(p_{LH}K_2+p_{HL}K_3)x)}\,p_{LL}}{p_{HH}}x\,K_4$$

- t_{HH}:

To obtain t_{HH} numerically, plug the λs and substitute for y into:

$$A(1-\alpha){q_{HH}}^{\alpha}{t_{HH}}^{-\alpha} = y\,K_1 + \frac{\left(\frac{p_{LH}yT_1}{p_{LL}(p_{HH}(1+K_4(p_{HL}+p_{LH})y)-(p_{LH}K_2+p_{HL}K_3)y)}\,p_{LL}+p_{LH}\right)}{p_{HH}}y\,K_2 + \frac{\left(\frac{p_{HL}yT_1}{p_{LL}(p_{HH}(1+K_4(p_{HL}+p_{LH})y)-(p_{LH}K_2+p_{HL}K_3)y)}\,p_{LL}+p_{HL}\right)}{p_{HH}}yK_3 + \frac{\frac{p_{HH}\,p_{LL}-(yK_2)\,p_{LH}(1-p_{HH})-(yK3)\,p_{HL}(1-p_{HH})}{p_{LL}(p_{HH}(1+K_4(p_{HL}+p_{LH})y)-(p_{LH}K_2+p_{HL}K_3)y)}\,p_{LL}}{p_{HH}}y\,K_4$$

The LH type:

- q_{LH}:

Use the closed form solution derived above and plug λ_1 (with respect to q_{HH}) to obtain:

$$q_{LH} = \left[\frac{1}{A\alpha}\left(Q_1 + \left(\frac{p_{LH}xT_1}{p_{LL}(p_{HH}(1+K_4(p_{HL}+p_{LH})x)-(p_{LH}K_2+p_{HL}K_3)x)}\frac{p_{LL}}{p_{LH}}\right)Q_2\right)\right]^{\frac{1}{\alpha-\gamma}} {t_{LH}}^{\frac{\gamma+\alpha-1}{\alpha-\gamma}}$$

- t_{LH}:

Use the closed form solution derived above and plug λ_1 (with respect to t_{HH}) to obtain:

$$t_{LH} = \left[\frac{1}{A(1-\alpha)}\left(Q_1 + \left(\frac{p_{LH}yT_1}{p_{LL}(p_{HH}(1+K_4(p_{HL}+p_{LH})y)-(p_{LH}K_2+p_{HL}K_3)y)}\frac{p_{LL}}{p_{LH}}\right)Q_2\right)\right]^{\frac{1}{1-\gamma-\alpha}} {q_{LH}}^{\frac{\gamma-\alpha}{1-\gamma-\alpha}}$$

The HL type:

- q_{HL}:

Use the closed form solution derived above and plug λ_2 (with respect to q_{HH}) to obtain:

$$q_{HL} = \left[\frac{1}{A\alpha}\left(R_1 + \left(\frac{p_{HL}xT_1}{p_{LL}(p_{HH}(1+K_4(p_{HL}+p_{LH})x)-(p_{LH}K_2+p_{HL}K_3)x)}\frac{p_{LL}}{p_{HL}}\right)R_2\right)\right]^{\frac{1}{\alpha-\gamma}} {t_{HL}}^{\frac{\gamma+\alpha-1}{\alpha-\gamma}}$$

- t_{HL}:

Use the closed form solution derived above and plug λ_2 (with respect to q_{HH}) to obtain:

$$t_{HL} = \left[\frac{1}{A(1-\alpha)}\left(R_1 + \left(\frac{p_{HL} y T_1}{p_{LL}(\,p_{HH}(1+K_4(\,p_{HL}+\,p_{LH})y) - (\,p_{LH}K_2 + \,p_{HL}K_3)y)}\frac{p_{LL}}{p_{HL}}\right)R_2\right)\right]^{\frac{1}{1-\gamma-\alpha}} q_{HL}{}^{\frac{\gamma-\alpha}{1-\gamma-\alpha}}$$

Case 2: Weak correlation

In case 3, simply let $\lambda_1 = \lambda$, $\lambda_2 = 1-\lambda$ and $\lambda_3 = 0$ and plug these values into the closed form solutions derived above.

For computational purposes, it is convenient to use the following expressions:

For the LH type:

- q_{LH}:

$$q_{LH} = \left[\frac{1}{A\alpha}\left(Q_1 + \left(\lambda\frac{p_{LL}}{p_{LH}}\right)Q_2\right)\right]^{\frac{1}{\alpha-\gamma}} t_{LH}{}^{\frac{\gamma+\alpha-1}{\alpha-\gamma}}$$

- t_{LH}:

$$t_{LH} = \left[\frac{1}{A(1-\alpha)}\left(Q_1 + \left(\lambda\frac{p_{LL}}{p_{LH}}\right)Q_2\right)\right]^{\frac{1}{1-\gamma-\alpha}} q_{LH}{}^{\frac{\gamma-\alpha}{1-\gamma-\alpha}}$$

For the HL type:

- q_{HL}:

$$q_{HL} = \left[\frac{1}{A\alpha}\left(R_1 + \left((1-\lambda)\frac{p_{LL}}{p_{HL}}\right)\right)R_2\right]^{\frac{1}{\alpha-\gamma}} t_{HL}{}^{\frac{\gamma+\alpha-1}{\alpha-\gamma}}$$

- t_{HL}:

$$t_{HL} = \left[\frac{1}{A(1-\alpha)}\left(R_1 + \left((1-\lambda)\frac{p_{LL}}{p_{HL}}\right)R_2\right)\right]^{\frac{1}{1-\gamma-\alpha}} q_{HL}{}^{\frac{\gamma-\alpha}{1-\gamma-\alpha}}$$

For the HH type:

- q_{HH}:

$$q_{HH} = \left[\frac{1}{A\alpha}\left(K_1 + \frac{(\lambda\,p_{LL} + \,p_{LH})}{p_{HH}}K_2 + \frac{\left((1-\lambda)\,p_{LL} + \,p_{HL}\right)}{p_{HH}}\,K_3\right)\right]^{\frac{1}{\alpha-\gamma}} t_{HH}{}^{\frac{\gamma+\alpha-1}{\alpha-\gamma}}$$

- t_{HH}:

$$t_{HH} = \left[\frac{1}{A(1-\alpha)}\left(K_1 + \frac{(\lambda\,p_{LL}+\,p_{LH})}{p_{HH}}K_2 + \frac{\left((1-\lambda)\,p_{LL}\,p_{LL}+\,p_{HL}\right)}{p_{HH}}\,K_3\right)\right]^{\frac{1}{1-\gamma-\alpha}} q_{HH}{}^{\frac{\gamma-\alpha}{1-\gamma-\alpha}}$$

Case 3: Negative correlation with asymmetry towards the LH type

It holds when $\lambda_1 = 1$ and $\lambda_2 = \lambda_3 = 0$:

For the LH type:

- q_{LH}:

$$q_{LH} = \left[\frac{1}{A\alpha}\left(Q_1 + \left(\frac{p_{LL}}{p_{LH}}\right) Q_2\right)\right]^{\frac{1}{\alpha-\gamma}} t_{LH}{}^{\frac{\gamma+\alpha-1}{\alpha-\gamma}}$$

- t_{LH}:

$$t_{LH} = \left[\frac{1}{A(1-\alpha)}\left(Q_1 + \left(\frac{p_{LL}}{p_{LH}}\right) Q_2\right)\right]^{\frac{1}{1-\gamma-\alpha}} q_{LH}{}^{\frac{\gamma-\alpha}{1-\gamma-\alpha}}$$

For the HL type:

- q_{HL}:

$$q_{HL} = \left[\frac{1}{A\alpha} R_1\right]^{\frac{1}{\alpha-\gamma}} t_{HL}{}^{\frac{\gamma+\alpha-1}{\alpha-\gamma}}$$

- t_{HL}:

$$t_{HL} = \left[\frac{1}{A(1-\alpha)} (R_1)\right]^{\frac{1}{1-\gamma-\alpha}} q_{HL}{}^{\frac{\gamma-\alpha}{1-\gamma-\alpha}}$$

For the HH type:

- q_{HH}:

$$q_{HH} = \left[\frac{1}{A\alpha}\left(K_1 + \frac{(p_{LL} + p_{LH})}{p_{HH}} K_2 + \frac{p_{HL}}{p_{HH}} K_3\right)\right]^{\frac{1}{\alpha-\gamma}} t_{HH}{}^{\frac{\gamma+\alpha-1}{\alpha-\gamma}}$$

- t_{HH}:

$$t_{HH} = \left[\frac{1}{A(1-\alpha)}\left(K_1 + \frac{(p_{LL} + p_{LH})}{p_{HH}} K_2 + \frac{p_{HL}}{p_{HH}} K_3\right)\right]^{\frac{1}{1-\gamma-\alpha}} q_{HH}{}^{\frac{\gamma-\alpha}{1-\gamma-\alpha}}$$

Case 4: Negative correlation with asymmetry towards the HL type

It holds when $\lambda_2 = 1$ and $\lambda_1 = \lambda_3 = 0$:

For the LH type:

- q_{LH}:

$$q_{LH} = \left[\frac{1}{A\alpha} Q_1\right]^{\frac{1}{\alpha-\gamma}} t_{LH}{}^{\frac{\gamma+\alpha-1}{\alpha-\gamma}}$$

- t_{LH}:

$$t_{LH} = \left[\frac{1}{A(1-\alpha)} Q_1\right]^{\frac{1}{1-\gamma-\alpha}} q_{LH}{}^{\frac{\gamma-\alpha}{1-\gamma-\alpha}}$$

For the HL type:

- q_{HL}:

$$q_{HL} = \left[\frac{1}{A\alpha}\left(R_1 + \left(\frac{p_{LL}}{p_{HL}}\right)R_2\right)\right]^{\frac{1}{\alpha-\gamma}} t_{HL}^{\frac{\gamma+\alpha-1}{\alpha-\gamma}}$$

- t_{HL}:

$$t_{HL} = \left[\frac{1}{A(1-\alpha)}\left(R_1 + \left(\frac{p_{LL}}{p_{HL}}\right)R_2\right)\right]^{\frac{1}{1-\gamma-\alpha}} q_{HL}^{\frac{\gamma-\alpha}{1-\gamma-\alpha}}$$

For the HH type:

- q_{HH}:

$$q_{HH} = \left[\frac{1}{A\alpha}\left(K_1 + \frac{p_{LH}}{p_{HH}}K_2 + \frac{(p_{LL} + p_{HL})}{p_{HH}}K_3\right)\right]^{\frac{1}{\alpha-\gamma}} t_{HH}^{\frac{\gamma+\alpha-1}{\alpha-\gamma}}$$

- t_{HH}:

$$t_{HH} = \left[\frac{1}{A(1-\alpha)}\left(K_1 + \frac{p_{LH}}{p_{HH}}K_2 + \frac{(p_{LL} + p_{HL})}{p_{HH}}K_3\right)\right]^{\frac{1}{1-\gamma-\alpha}} q_{HH}^{\frac{\gamma-\alpha}{1-\gamma-\alpha}}$$

A5. Incentive compatibility constraints for simulation results

Case 1 (positive correlation):

Distribution 1A: where $p_{LL} = 0.5$, $p_{LH} = 0$, $p_{HL} = 0$, $p_{HH} = 0.5$, $cov = 0.25$

- The efficient is *simultaneously* indifferent between his contract, the inefficient's (HH), and the two middle types' (HL and LH):

  $U_{LL} = U_{HH} + \Delta_{LL}^{HH}(q_{HH}, t_{HH})$, because $U_{LL} = 0.04$, $U_{HH} = 0$, $\Delta_{LL}^{HH}(q_{HH}, t_{HH}) = 0.04$

  $U_{LL} = U_{LH} + \Delta_{LL}^{LH}(q_{LH}, t_{LH})$, because $U_{LL} = 0.04$, $U_{LH} = 0.02$, $\Delta_{LL}^{LH}(q_{LH}, t_{LH}) = 0.02$

  $U_{LL} = U_{HL} + \Delta_{LL}^{HL}(q_{LH}, t_{LH})$, because $U_{LL} = 0.04$, $U_{HL} = 0.02$, $\Delta_{LL}^{HL}(q_{HL}, t_{HL}) = 0.02$

  Expected profit is $E(\pi) = 2.90$

Distribution 1B: where $p_{LL} = 0.49$, $p_{LH} = 0.01$, $p_{HL} = 0.01$, $p_{HH} = 0.49$, $cov = 0.24$

When the principal fixes $t_{ij} = 2$:

$U_{LL} = U_{HH} + \Delta_{LL}^{HH}(q_{HH}, t_{HH})$, because $U_{LL} = 0.08$, $U_{HH} = 0$, $\Delta_{LL}^{HH}(q_{HH}, t_{HH}) = 0.08$

$U_{LL} = U_{LH} + \Delta_{LL}^{LH}(q_{LH}, t_{LH})$, because $U_{LL} = 0.08$, $U_{LH} = 0.05$, $\Delta_{LL}^{LH}(q_{LH}, t_{LH}) = 0.03$

$U_{LL} = U_{HL} + \Delta_{LL}^{HL}(q_{LH}, t_{LH})$, because $U_{LL} = 0.08$, $U_{HL} = 0.05$, $\Delta_{LL}^{HL}(q_{HL}, t_{HL}) = 0.03$

Expected profit is $E(\pi) = 2.84$

When the principal fixes $t_{ij} = 4$:

$U_{LL} = U_{HH} + \Delta_{LL}^{HH}(q_{HH}, t_{HH})$, because $U_{LL} = 0.13$, $U_{HH} = 0$, $\Delta_{LL}^{HH}(q_{HH}, t_{HH}) = 0.13$

$U_{LL} = U_{LH} + \Delta_{LL}^{LH}(q_{LH}, t_{LH})$, because $U_{LL} = 0.13$, $U_{LH} = 0.08$, $\Delta_{LL}^{LH}(q_{LH}, t_{LH}) = 0.05$

$U_{LL} = U_{HL} + \Delta_{LL}^{HL}(q_{LH}, t_{LH})$, because $U_{LL} = 0.13$, $U_{HL} = 0.08$, $\Delta_{LL}^{HL}(q_{HL}, t_{HL}) = 0.05$

Expected profit is $E(\pi) = 2.83$

When the principal fixes $t_{ij} = 6$:

$U_{LL} = U_{HH} + \Delta_{LL}^{HH}(q_{HH}, t_{HH})$, because $U_{LL} = 0.19$, $U_{HH} = 0$, $\Delta_{LL}^{HH}(q_{HH}, t_{HH}) = 0.19$

$U_{LL} = U_{LH} + \Delta_{LL}^{LH}(q_{LH}, t_{LH})$, because $U_{LL} = 0.19$, $U_{LH} = 0.11$, $\Delta_{LL}^{LH}(q_{LH}, t_{LH}) = 0.08$

$U_{LL} = U_{HL} + \Delta_{LL}^{HL}(q_{LH}, t_{LH})$, because $U_{LL} = 0.19$, $U_{HL} = 0.11$, $\Delta_{LL}^{HL}(q_{HL}, t_{HL}) = 0.08$

Expected profit is $E(\pi) = 2.82$

Distribution 1C: where $p_{LL} = 0.45$, $p_{LH} = 0.05$, $p_{HL} = 0.05$, $p_{HH} = 0.45$, $cov = 0.20$

When the principal fixes $t_{ij} = 2$:

$U_{LL} = U_{HH} + \Delta_{LL}^{HH}(q_{HH}, t_{HH})$, because $U_{LL} = 0.40$, $U_{HH} = 0$, $\Delta_{LL}^{HH}(q_{HH}, t_{HH}) = 0.40$

$U_{LL} = U_{LH} + \Delta_{LL}^{LH}(q_{LH}, t_{LH})$, because $U_{LL} = 0.40$, $U_{LH} = 0.23$, $\Delta_{LL}^{LH}(q_{LH}, t_{LH}) = 0.17$

$U_{LL} = U_{HL} + \Delta_{LL}^{HL}(q_{LH}, t_{LH})$, because $U_{LL} = 0.40$, $U_{HL} = 0.23$ $\Delta_{LL}^{HL}(q_{HL}, t_{HL}) = 0.17$

Expected profit is $E(\pi) = 2.62$

When the principal fixes $t_{ij} = 4$:

$U_{LL} = U_{HH} + \Delta_{LL}^{HH}(q_{HH}, t_{HH})$, because $U_{LL} = 0.92$, $U_{HH} = 0$, $\Delta_{LL}^{HH}(q_{HH}, t_{HH}) = 0.92$

$U_{LL} = U_{LH} + \Delta_{LL}^{LH}(q_{LH}, t_{LH})$, because $U_{LL} = 0.92$, $U_{LH} = 0.54$, $\Delta_{LL}^{LH}(q_{LH}, t_{LH}) = 0.38$

$U_{LL} = U_{HL} + \Delta_{LL}^{HL}(q_{LH}, t_{LH})$, because $U_{LL} = 0.92$, $U_{HL} = 0.54$ $\Delta_{LL}^{HL}(q_{HL}, t_{HL}) = 0.38$

Expected profit is $E(\pi) = 2.47$

When the principal fixes $t_{ij} = 6$:

$U_{LL} = U_{HH} + \Delta_{LL}^{HH}(q_{HH}, t_{HH})$, because $U_{LL} = 1.55$, $U_{HH} = 0$, $\Delta_{LL}^{HH}(q_{HH}, t_{HH}) = 1.55$

$U_{LL} = U_{LH} + \Delta_{LL}^{LH}(q_{LH}, t_{LH})$, because $U_{LL} = 1.55$, $U_{LH} = 0.91$, $\Delta_{LL}^{LH}(q_{LH}, t_{LH}) = 0.64$

$U_{LL} = U_{HL} + \Delta_{LL}^{HL}(q_{LH}, t_{LH})$, because $U_{LL} = 1.55$, $U_{HL} = 0.91$ $\Delta_{LL}^{HL}(q_{HL}, t_{HL}) = 0.64$

Expected profit is $E(\pi) = 2.23$

Case 2: Weak correlation

Distribution 2A where where $p_{LL} = 0.3$, $p_{LH} = 0.25$, $p_{HL} = 0.19$, $p_{HH} = 0.26$, $cov = 0.03$. For $\lambda = 0.57$, the following output levels hold and satisfy equation (7):

The efficient is indifferent between his contract and that of any of the two middle types:

$U_{LL} = U_{LH} + \Delta_{LL}^{LH}(q_{LH}, t_{LH})$, because $U_{LL} = 0.51$, $U_{LH} = 0.01$, $\Delta_{LL}^{LH}(q_{LH}, t_{LH}) = 0.50$

$U_{LL} = U_{HL} + \Delta_{LL}^{HL}(q_{HL}, t_{HL})$, because $U_{LL} = 0.51$, $U_{HL} = 0.01$, $\Delta_{LL}^{HL}(q_{HL}, t_{HL}) = 0.50$

However, the efficient's incentive compatibility constraint with respect to the inefficient does not bind:

$U_{LL} = U_{HH} + \Delta_{LL}^{HH}(q_{HH}, t_{HH})$, because $U_{LL} = 0.55$, $U_{HH} = 0$, $\Delta_{LL}^{HH}(q_{HH}, t_{HH}) = 0.01$

Expected profit is $E(\pi) = 1.96$

Distribution 2B where where $p_{LL} = 0.28$, $p_{LH} = 0.22$, $p_{HL} = 0.22$, $p_{HH} = 0.28$, $cov = 0.03$. For $\lambda = 0.50$, the following output levels hold and satisfy equation (7).

The efficient is indifferent between his contract and that of any of the two middle types:

$U_{LL} = U_{LH} + \Delta_{LL}^{LH}(q_{LH}, t_{LH})$, because $U_{LL} = 0.55$, $U_{LH} = 0.01$, $\Delta_{LL}^{LH}(q_{LH}, t_{LH}) = 0.54$

$U_{LL} = U_{HL} + \Delta_{LL}^{HL}(q_{HL}, t_{HL})$, because $U_{LL} = 0.55$, $U_{HL} = 0.01$, $\Delta_{LL}^{HL}(q_{HL}, t_{HL}) = 0.54$

However, the incentive compatibility constraint with respect to the inefficient does not bind:

$U_{LL} = U_{HH} + \Delta_{LL}^{HH}(q_{HH}, t_{HH})$, because $U_{LL} = 0.55$, $U_{HH} = 0$, $\Delta_{LL}^{HH}(q_{HH}, t_{HH}) = 0.02$

Expected profit is $E(\pi) = 1.84$

Case 3: Negative correlation with asymmetry towards the LH type

Distribution 3A where $p_{LL} = 0.125$, $p_{LH} = 0.5$, $p_{HL} = 0.25$, $p_{HH} = 0.125$, $cov = -0.11$

The efficient is indifferent between his contract and that of the LH type:

$U_{LL} = U_{LH} + \Delta_{LL}^{LH}(q_{LH}, t_{LH})$, because $U_{LL} = 0.98$, $U_{LH} = 0$, $\Delta_{LL}^{LH}(q_{LH}, t_{LH}) = 0.98$

Expected profit is $E(\pi) = 1.34$

Distribution 3B where where $p_{LL} = 0.08$, $p_{LH} = 0.74$, $p_{HL} = 0.1$, $p_{HH} = 0.08$, $cov = -0.11$

The efficient is indifferent between his contract and that of the LH type:

$U_{LL} = U_{LH} + \Delta_{LL}^{LH}(q_{LH}, t_{LH})$, because $U_{LL} = 1.24$, $U_{LH} = 0$, $\Delta_{LL}^{LH}(q_{LH}, t_{LH}) = 1.24$

Expected profit is $E(\pi) = 1.2$

Case 4: Negative correlation with asymmetry towards the HL type

Distribution 4A, where $p_{LL} = 0.125$, $p_{LH} = 0.25$, $p_{HL} = 0.5$, $p_{HH} = 0.125$, $cov = -0.11$

The efficient is indifferent between his contract and that of the HL type:

$U_{LL} = U_{HL} + \Delta_{LL}^{HL}(q_{HL}, t_{HL})$, because $U_{LL} = 0.98$, $U_{LH} = 0$, $\Delta_{LL}^{HL}(q_{HL}, t_{HL}) = 0.98$

Distribution 4B, where $p_{LL} = 0.04$, $p_{LH} = 0.05$, $p_{HL} = 0.87$, $p_{HH} = 0.04$, $cov = -0.04$

The efficient is indifferent between his contract and that of the LH type:

$U_{LL} = U_{HL} + \Delta_{LL}^{HL}(q_{HL}, t_{HL})$, because $U_{LL} = 1.38$, $U_{HL} = 0$, $\Delta_{LL}^{HL}(q_{HL}, t_{HL}) = 1.38$

A6. MATLAB script files to compute simulation results

For Case 1 (Positive Correlation):

```
%**************************************************%
%      Parameters of the cost function           %
%**************************************************%
A = 5;
alpha = 0.5;
beta = 0.5;
gama = 0.6;
%**************************************************%
%      Distribution of types                     %
%**************************************************
pll = 0.45;
plh = 0.05;
phl = 0.05;
phh = 0.45;
%**************************************************%
%     Asymmetric information parameters          %
%**************************************************%
theta_qL = 2;
theta_qH = 4;
theta_tL= 4;
theta_tH= 8;
%**************************************************%
%       Deltas                                   %
%**************************************************
delta_HH_LH = ((theta_qH)^beta)*((theta_tH)^(1-beta))...
   -((theta_qL)^beta)*((theta_tH)^(1-beta));
delta_HH_HL = ((theta_qH)^beta)*((theta_tH)^(1-beta))...
   -((theta_qH)^beta)*((theta_tL)^(1-beta));
delta_HH_LL = ((theta_qH)^beta)*((theta_tH)^(1-beta))...
   -((theta_qL)^beta)*((theta_tL)^(1-beta));
delta_LH_LL = ((theta_qL)^beta)*((theta_tH)^(1-beta))...
   -((theta_qL)^beta)*((theta_tL)^(1-beta));
delta_HL_LL = ((theta_qH)^beta)*((theta_tL)^(1-beta))...
   -((theta_qL)^beta)*((theta_tL)^(1-beta));
%**************************************************%

K1 = gama*((theta_qH)^beta)*((theta_tH)^(1-beta));
K2 = gama*delta_HH_LH;
K3 = gama*delta_HH_HL;
K4 = gama*delta_HH_LL;

 Q1 = gama*((theta_qL)^beta)*((theta_tH)^(1-beta));
Q2 = gama*delta_LH_LL;

R1 = gama*((theta_qH)^beta)*((theta_tL)^(1-beta));
R2 = gama*delta_HL_LL;

S1 = gama*((theta_qL)^beta)*((theta_tL)^(1-beta));
```

```
T1 = gama*(plh*delta_HH_LH + phl*delta_HH_HL + phh*pll*delta_HH_LL);

%****************************************************%
%            Lambdas wrt q                   %
%****************************************************%

syms qhh

thh = 6;

 lambda1 = (plh * (plh * (gama*((qhh*thh)^(gama-1))*thh*delta_HH_LH) + ...
            phl * (gama*((qhh *thh)^(gama-1))*thh*delta_HH_HL)+ ...
            phh*pll*(gama*((qhh *thh)^(gama-1))*thh*delta_HH_LL)))/...
            ((phh - (gama*((qhh *thh)^(gama-1))*thh*delta_HH_HL)*phl ...
            + (gama*((qhh *thh)^(gama-1))*thh*delta_HH_LL)*phh*phl - ...
            (gama*((qhh*thh)^(gama-1))*thh*delta_HH_LH)*plh +...
            (gama*((qhh *thh)^(gama-1))*thh*delta_HH_LL)*phh*plh)* pll);

lambda2 = (plh * (phl * (gama*((qhh*thh)^(gama-1))*thh*delta_HH_LH) + ...
            phl * (gama*((qhh *thh)^(gama-1))*thh*delta_HH_HL)+ ...
            phh*pll*(gama*((qhh *thh)^(gama-1))*thh*delta_HH_LL)))/...
            ((phh - (gama*((qhh *thh)^(gama-1))*thh*delta_HH_HL)*phl ...
            + (gama*((qhh *thh)^(gama-1))*thh*delta_HH_LL)*phh*phl - ...
            (gama*((qhh*thh)^(gama-1))*thh*delta_HH_LH)*plh +...
            (gama*((qhh *thh)^(gama-1))*thh*delta_HH_LL)*phh*plh)* pll);

 lambda3 = (phh*pll - (gama*((qhh *thh)^(gama-1))*thh*delta_HH_HL)*phl*(1-phh) - ...
        (gama*((qhh*thh)^(gama-1))*thh*delta_HH_LH)*plh*(1-phh))/...
        ((phh - (gama*((qhh *thh)^(gama-1))*thh*delta_HH_HL)*phl ...
            + (gama*((qhh *thh)^(gama-1))*thh*delta_HH_LL)*phh*phl - ...
            (gama*((qhh*thh)^(gama-1))*thh*delta_HH_LH)*plh +...
            (gama*((qhh *thh)^(gama-1))*thh*delta_HH_LL)*phh*plh)* pll);


%****************************************************%
%      Numerical computation of qhh              %
%****************************************************%

x= (((qhh*thh)^(gama-1))*thh);

qhh = vpasolve(A*alpha*(qhh^(alpha - 1))*(thh^(1 - alpha)) == x*K1 ...
    + ((lambda1*pll + plh)/phh)*(x*K2)...
    + ((lambda2*pll + phl)/phh)*(x*K3)...
    + ((lambda3*pll)/phh)*(x*K4),qhh, [0,0.1]);

 qlh = qhh;
 qhl = qhh;

%****************************************************%
%     Computing qlh, given tlh and lambda        %
%****************************************************%

tlh = 6;
```

```
qlh = (((1/(A*alpha))*(Q1 + lambda1*(pll/plh)*Q2))^(1/(alpha - gama)))*...
   tlh^((gama + alpha - 1 )/(alpha - gama));

%**************************************************%
%    Computing qhl, iven tlh and lambda           %
%**************************************************%

thl = 6;

qhl = (((1/(A*alpha))*(R1 + lambda2*(pll/phl)*R2))^(1/(alpha - gama)))*...
   thl^((gama + alpha - 1 )/(alpha - gama));

% %**************************************************%
% %              Computing qll                      %
% %**************************************************%

tll = 6;

qll = (((1/(A*alpha))*S1)^(1/(alpha - gama)))*...
   tll^((gama + alpha - 1 )/(alpha - gama));
```

For Case 2 (Weak Correlation):

```
%**************************************************%
A = 5;
alpha = 0.5;
beta = 0.5;
gama = 0.6;

thh = 6;
tlh = 6;
thl = 6;
tll = 6;
%**************************************************%
pll = 0.3;
plh = 0.25;
phl = 0.19;
phh = 0.26;
%**************************************************%
theta_qL = 2;
theta_qH = 4;
theta_tL= 4;
theta_tH= 8;
%**************************************************%
delta_HH_LH = ((theta_qH)^beta)*((theta_tH)^(1-beta))...
   -((theta_qL)^beta)*((theta_tH)^(1-beta));
delta_HH_HL = ((theta_qH)^beta)*((theta_tH)^(1-beta))...
   -((theta_qH)^beta)*((theta_tL)^(1-beta));
delta_HH_LL = ((theta_qH)^beta)*((theta_tH)^(1-beta))...
   -((theta_qL)^beta)*((theta_tL)^(1-beta));
delta_LH_LL = ((theta_qL)^beta)*((theta_tH)^(1-beta))...
   -((theta_qL)^beta)*((theta_tL)^(1-beta));
```

```
delta_HL_LL = ((theta_qH)^beta)*((theta_tL)^(1-beta))...
   -((theta_qL)^beta)*((theta_tL)^(1-beta));

%**************************************************%

K1 = gama*((theta_qH)^beta)*((theta_tH)^(1-beta));
K2 = gama*delta_HH_LH;
K3 = gama*delta_HH_HL;
K4 = gama*delta_HH_LL;

Q1 = gama*((theta_qL)^beta)*((theta_tH)^(1-beta));
Q2 = gama*delta_LH_LL;

R1 = gama*((theta_qH)^beta)*((theta_tL)^(1-beta));
R2 = gama*delta_HL_LL;

S1 = gama*((theta_qL)^beta)*((theta_tL)^(1-beta));

T1 = gama*(plh*delta_HH_LH + phl*delta_HH_HL + phh*pll*delta_HH_LL);
%**************************************************%
%  Finding the lambda that satisfies eq. A40      %
%**************************************************%
%
 init_guess=[0 1];
% %
 syms lambda
% %
 vpasolve( ((( ((((1/(A*alpha))*(K1 + ((lambda*pll + plh)/phh)* K2 + ...
      (((1 - lambda)*pll + phl)/phh)*K3))...
       ^(1/(alpha - gama)))*thh^((gama - 1 + alpha)/(alpha - gama))) * thh)^(gama))*delta_HH_LH) +
(((( ((((1/(A*alpha))*(Q1 + (lambda * pll/plh)*Q2))...
       ^(1/(alpha - gama)))*tlh^((gama - 1 + alpha)/(alpha - gama))) *tlh)^(gama))*delta_LH_LL)   ==
(((( ((((1/(A*alpha))*(K1 + ((lambda*pll + plh)/phh)* K2 + ...
        (((1 - lambda)*pll + phl)/phh)*K3))...
        ^(1/(alpha - gama)))*thh^((gama - 1 + alpha)/(alpha - gama))) *thh)^(gama))*delta_HH_HL) +
(((( ((((1/(A*alpha))*(R1 + ((1-lambda) * pll/phl)*R2))...
        ^(1/(alpha - gama)))*thl^((gama - 1 + alpha)/(alpha - gama))) *thl)^(gama))*delta_HL_LL),
lambda, init_guess)

%**************************************************%
%        Computing output levels                 %
%**************************************************%

lambda = 0.5681818181818181724747619198759 6;

%**************************************************%
%  Output levels for the inefficient (HH) type     %
%**************************************************%
%
  qhh = ((((1/(A*alpha))*(K1 + ((lambda*pll + plh)/phh)* K2 + ...
              (((1 - lambda)*pll + phl)/phh)*K3))...
              ^(1/(alpha - gama)))*thh^((gama - 1 + alpha)/(alpha - gama)));

% %**************************************************%
```

```
% %  Output levels for middle (LH) type            %
% %**************************************************%
%
  qlh = ((((1/(A*alpha))*(Q1 + (lambda * pll/plh)*Q2))...
         ^(1/(alpha - gama)))*tlh^((gama - 1 + alpha)/(alpha - gama)));
% %**************************************************%
% %  Output levels for middle (HL) type            %
% %**************************************************%
%
    qhl = ((((1/(A*alpha))*(R1 + ((1-lambda) * pll/phl)*R2))...
        ^(1/(alpha - gama)))*thl^((gama - 1 + alpha)/(alpha - gama)));
% %**************************************************%
% %  Output levels for efficient (LL) type            %
% %**************************************************%

qll = (((1/(A*alpha))*S1)^(1/(alpha - gama)))*tll^((gama - 1 + alpha)/(alpha - gama));
```

For Case 3 (Asymmetry towards the LH type):

```
%**************************************************%
A = 5;
alpha = 0.5;
beta = 0.5;
gama = 0.6;
%**************************************************%
pll = 0.08;
plh = 0.74;
phl = 0.1;
phh = 0.08;
%**************************************************%
theta_qL = 2;
theta_qH = 4;
theta_tL= 4;
theta_tH= 8;
%**************************************************%
delta_HH_LH = ((theta_qH)^beta)*((theta_tH)^(1-beta))...
   -((theta_qL)^beta)*((theta_tH)^(1-beta));
delta_HH_HL = ((theta_qH)^beta)*((theta_tH)^(1-beta))...
   -((theta_qH)^beta)*((theta_tL)^(1-beta));
delta_HH_LL = ((theta_qH)^beta)*((theta_tH)^(1-beta))...
   -((theta_qL)^beta)*((theta_tL)^(1-beta));
delta_LH_LL = ((theta_qL)^beta)*((theta_tH)^(1-beta))...
   -((theta_qL)^beta)*((theta_tL)^(1-beta));
delta_HL_LL = ((theta_qH)^beta)*((theta_tL)^(1-beta))...
   -((theta_qL)^beta)*((theta_tL)^(1-beta));
%**************************************************%

K1 = gama*((theta_qH)^beta)*((theta_tH)^(1-beta));
K2 = gama*delta_HH_LH;
K3 = gama*delta_HH_HL;
K4 = gama*delta_HH_LL;

Q1 = gama*((theta_qL)^beta)*((theta_tH)^(1-beta));
Q2 = gama*delta_LH_LL;
```

```
R1 = gama*((theta_qH)^beta)*((theta_tL)^(1-beta));
R2 = gama*delta_HL_LL;

S1 = gama*((theta_qL)^beta)*((theta_tL)^(1-beta));

T1 = gama*(plh*delta_HH_LH + phl*delta_HH_HL + phh*pll*delta_HH_LL);


%**************************************************%
%              Computing qhh                       %
%**************************************************%

thh = 6;

qhh = (((1/(A*alpha))*(K1 + ((pll + plh)/phh)*K2 + ...
   (phl/phh)*K3))^(1/(alpha - gama)))* thh^((gama - 1 + alpha)/(alpha - gama));

%**************************************************%
%              Computing qlh                       %
%**************************************************%
tlh = 6;
  qlh = (((1/(A*alpha))*(Q1 + (pll/plh)*Q2))...
   ^(1/(alpha - gama)))*tlh^((gama - 1 + alpha)/(alpha - gama));

%**************************************************%
%              Computing qhl                       %
%**************************************************%

thl = 6;

qhl = (((1/(A*alpha))*(R1))^(1/(alpha - gama)))...
   *thl^((gama - 1 + alpha)/(alpha - gama));

%**************************************************%
%              Computing qll                       %
%**************************************************%

tll = 6;
%
 qll = (((1/(A*alpha))*S1)^(1/(alpha - gama)))*tll^((gama - 1 + alpha)/(alpha - gama));
```

For Case 4 (Asymmetry towards the HL type):

```
%**************************************************%
A = 5;
alpha = 0.5;
beta = 0.5;
gama = 0.6;
%**************************************************%
pll = 0.125;
plh = 0.25;
phl = 0.5;
```

```
phh = 0.125;
%**************************************************%
theta_qL = 2;
theta_qH = 4;
theta_tL= 4;
theta_tH= 8;
%**************************************************%
delta_HH_LH = ((theta_qH)^beta)*((theta_tH)^(1-beta))...
   -((theta_qL)^beta)*((theta_tH)^(1-beta));
delta_HH_HL = ((theta_qH)^beta)*((theta_tH)^(1-beta))...
   -((theta_qH)^beta)*((theta_tL)^(1-beta));
delta_HH_LL = ((theta_qH)^beta)*((theta_tH)^(1-beta))...
   -((theta_qL)^beta)*((theta_tL)^(1-beta));
delta_LH_LL = ((theta_qL)^beta)*((theta_tH)^(1-beta))...
   -((theta_qL)^beta)*((theta_tL)^(1-beta));
delta_HL_LL = ((theta_qH)^beta)*((theta_tL)^(1-beta))...
   -((theta_qL)^beta)*((theta_tL)^(1-beta));
%**************************************************%

K1 = gama*((theta_qH)^beta)*((theta_tH)^(1-beta));
K2 = gama*delta_HH_LH;
K3 = gama*delta_HH_HL;
K4 = gama*delta_HH_LL;

Q1 = gama*((theta_qL)^beta)*((theta_tH)^(1-beta));
Q2 = gama*delta_LH_LL;

R1 = gama*((theta_qH)^beta)*((theta_tL)^(1-beta));
R2 = gama*delta_HL_LL;

S1 = gama*((theta_qL)^beta)*((theta_tL)^(1-beta));

T1 = gama*(plh*delta_HH_LH + phl*delta_HH_HL + phh*pll*delta_HH_LL);


%**************************************************%
%              Computing qhh                     %
%**************************************************%

thh = 6;

qhh = (((1/(A*alpha))*(K1 + (plh/phh)*K2 + ...
   ((pll+phl)/phh)*K3))^(1/(alpha - gama)))* thh^((gama - 1 + alpha)/(alpha - gama));

%**************************************************%
%              Computing qlh                     %
%**************************************************%
tlh = 6;

qlh = (((1/(A*alpha))*(Q1))...
   ^(1/(alpha - gama)))*tlh^((gama - 1 + alpha)/(alpha - gama));

%**************************************************%
%              Computing qhl                     %
```

```
%**************************************************%

thl = 6;

qhl = (((1/(A*alpha))*(R1+(pll/phl)*R2))^(1/(alpha - gama)))...
    *thl^((gama - 1 + alpha)/(alpha - gama));

%**************************************************%
%              Computing qll                     %
%**************************************************%

tll = 6;
%
 qll = (((1/(A*alpha))*S1)^(1/(alpha - gama)))*tll^((gama - 1 + alpha)/(alpha - gama));
```