Markets for Price Risk*

Baiyang Han[†], Junyi Hu[‡], Anthony Lee Zhang[§]

January 2026

Abstract

Financial derivatives, such as futures, options, and swaps, are not contracts on exogenous states of the world, as in Arrow (1964): their payoffs depend on the endogenous market prices of certain goods. How well do *markets for price risk* approximate the richer state-contingent contracts analyzed by Arrow? We solve analytically for equilibrium trades, asset prices, and welfare in price-linked derivative contract markets, illustrating the role that these derivatives play in risk sharing.

*We thank David Baqaee, Doug Diamond, Darrell Duffie, Piotr Dworczak, Arpit Gupta, David Hirshleifer, Alex Imas, Baiyun Jing, Michael Johannes, Ralph Koijen, Dan Luo, Liming Ning, Hrishikesh Relekar, Yijing Ren, Milena Wittwer, Xiaobo Yu and seminar participants at Columbia University and UChicago Booth for helpful comments. We are grateful to Jingxi Li for excellent research assistance.

[†]Stanford University Graduate School of Business; byhan@stanford.edu

[‡]USC Marshall School of Business; junyihu@marshall.usc.edu

[§]University of Chicago, Booth School of Business; anthony.zhang@chicagobooth.edu.

1 Introduction

Arrow (1964) modelled financial assets as state-contingent contracts. But oil futures, cattle options, and interest rate swaps are not contracts on exogenous states of the world: their payments depend on the market prices of commodities such as oil, cattle and bonds. Financial markets in reality are thus much simpler than Arrow's state-contingent contracts: they are largely low-dimensional markets for price risk. This paper asks what equilibrium outcomes in such markets look like, and how well they approximate the risk-sharing possibilities of complete Arrow markets.

To address this question, we build an analytically tractable model of risk sharing in incomplete markets using price derivatives. Under our functional form assumptions, we solve analytically for equilibrium outcomes. We show that price derivative markets can be thought of as "markets for Greeks": futures contracts allow consumers to trade linear "delta" exposure to goods prices, and variance swaps allow consumers to trade quadratic "gamma exposure". Contract trade is driven by differences in consumers' Greeks, and risk premia arise when society in aggregate is long or short Greeks. Our model reconciles disparate arguments in the literature for how trade and equilibrium prices are determined in futures markets, and introduces a qualitatively new argument for the role that variance swaps play in socially efficient risk-sharing.

Our model has two goods: "money", and a real commodity, representing a good such as oil, wheat, or steel. Agents have a concave production technology which converts the commodity into money-equivalents: in other words, conditional on shock realization, agents' utility is quasilinear over the commodity and money. Agents are risk-averse over monetary wealth, with CARA utility with potentially different risk aversions. The only source of uncertainty in the economy is agents' random endowments of the commodity, which we call inventory shocks.

Our social planner faces a two-stage problem. In the second stage, conditional on any realization of inventory shocks, market outcomes should be *allocatively efficient*: society's commodity endowment must be allocated to agents in a way that maximizes money-equivalent wealth. In the first stage, *risk sharing* should be optimal: the shocks to social wealth induced by the random aggregate commodity endowments should be divided proportionally across agents according to their risk aversions, following Borch (1962) and Wilson (1968).

There is a simple and classical Arrow (1964) implementation of the first-best outcome, which can be thought of as a backward induction process. Allocative efficiency is achieved through *spot markets*, which open after inventory shocks are realized, in which money and commodities are traded for each other, maximizing society's aggregate money-equivalent value of the commodity. Risk sharing is achieved through perfect *financial markets*: when agents can trade contingent claims on states of the world – the entire vector of inventory shocks – then financial markets allow agents to reallocate spot market wealths efficiently across states of the world, decentralizing the social planner's first-best solution.

In the absence of financial markets, spot markets achieve full allocative efficiency, optimally converting goods into money, but fail to efficiently distribute money between agents across states of the world. We analyze three risk-sharing distortions induced by spot markets. First, exposures to inventory volatility are distributed inefficiently across agents: "market makers" in spot markets have *positive* volatility exposure, leaving "liquidity takers" with more negative exposure than society at large. Second, high or low realizations of agents' inventory shocks represent wealth shocks, which spot markets do not distribute across agents. Third, expected commodity trades generate directional price risk – sellers (buyers) have positive (negative) spot price exposures – and spot markets fail to efficiently net these risks across agents.

We proceed towards analyzing equilibrium with *price-linked derivative contracts*, developing two main technical results.

First, a key source of complexity in our setting is that consumers face both common uncertainty from spot good prices, and idiosyncratic uncertainty from inventory shocks. By *projecting* inventory shocks onto spot prices, we derive a "price-contingent-certainty-equivalent" representation, in which consumers have modified wealths, but *only* face common price risk. This reduces our baseline incomplete-markets setting into a complete-markets problem, where the random spot price is the only source of uncertainty.

Second, classically, consumers' optimal financial asset positions must equate "risk-neutral" – that is, marginal-utility-weighted – expected returns of all securities to their prices. Under our functional form assumptions, all consumers' risk-neutral distributions over spot prices are Gaussian, with means and variances "tilted" relative to the physical spot price distributions. The consumer optimally purchases derivatives until her tilted expectations of derivative payoffs are equal to market prices; we can thus derive consumers' financial asset demands from formulas for moments of the Gaussian distribution. Optimal contract positions depend on contract prices, and the linear and quadratic exposures of consumer i's wealth to spot prices; by analogy to "Greeks" in option pricing, we call these first- and second-order wealth exposures "delta" and "gamma".

Futures contracts are markets for *delta*, allowing consumers to exchange linear wealth exposures to spot prices. A consumer's spot market delta is driven by three components. The first is *expected spot purchases*: consumers who expect to be net sellers (buyers) tend to benefit (suffer) when spot prices are high. The second is *inventory shock variances*: consumers' inventories are correlated with spot prices, generating net negative exposure to high prices. The third is *basis risk*: the negative effects of idiosyncratic shocks on conditional-certainty-equivalent wealths are larger when prices are greater in absolute value, generating additional negative exposures to high prices.

All three forces can generate dispersion in consumers' deltas, and thus trade in futures markets. As in the classical futures models, expected inventory shocks generate trade: if spot good sellers take short futures positions and spot buyers take long positions, each party can neutralize the others' directional exposure to spot prices. But trade can also be generated by other forces: consumers with highly variable inventory shocks for valuable goods naturally long futures to hedge inventory risk, generalizing observations in the classic literature on optimal hedging (McKinnon, 1967; Rolfo, 1980), and heterogeneous exposures to basis risk can also generate futures trade.

Society has negative aggregate delta, if the good is valuable, since high spot prices imply that society's aggregate inventory is low. Society thus has aggregate demand to be long the futures contract to hedge wealth risk. Financial market clearing then implies that the equilibrium futures risk premium is negative: the futures price is higher than the expected spot price, and long futures positions make negative expected returns.

Interestingly, generalizing a finding of Hirshleifer (1990), futures demand from expected spot trades never contributes to futures risk premia in our model. Spot markets always clear, so society as a whole does not expect to buy or sell the spot good on average. Expected commodity sellers want to short exactly as many futures contracts as expected buyers want to be long; buyers and sellers can thus always perfectly neutralize each others' spot price exposures, without generating any aggregate demand pressure on futures risk premia.

A *variance swap* pays consumers based on the realized squared deviation of prices from their mean; variance swaps allow consumers to exchange second-order spot price exposure, or *gamma*. We demonstrate a novel intuition for how price variance markets enable welfare-improving risk transfers. *Liquidity takers* in spot markets – for example, oil consumers who inelastically demand random quantities of oil – have natural short exposures to price volatility: extreme shocks cause them to pay more in price impact costs, lowering their expected wealth. *Market makers* in spot markets – for example, oil producers who supply oil in response to consumers' demand shocks – have natural *long* volatility exposure: their profits are higher the more extreme consumers' shocks are. Just as futures contracts allow expected buyers and sellers to neutralize their offsetting directional exposures to price risk, variance swap markets allow liquidity takers and market makers to neutralize their offsetting exposures to price *variance* risk, lowering wealth variance and increasing expected utilities for both parties.

In our model, society as a whole is always short gamma. Society's aggregate wealth is concave in prices, since society's marginal utility for the good is downwards-sloping: negative inventory shocks harm society more than positive inventory shocks benefit. In incomplete markets, price variance has an additional negative effect on aggregate wealth through basis risk: more variable prices increase the wealth cost of unhedgeable inventory shocks. Society thus wants to be long variance swaps in the aggregate to hedge wealth risk. Financial market clearing thus leads the variance risk premium to be negative: variance swaps trade at greater than their fair price under the physical probability measure $\mathbb{P}$, and thus long positions in variance swaps make negative returns in equilibrium. In equilibrium, variance markets reallocate society's net gamma exposure to market participants according to their risk aversions.

Futures and variance swaps are "optimal" price-contingent transfer tools in our model, in the sense that a social planner who is limited to transfers across agents that are functions of spot market prices, cannot improve on the allocation generated in equilibrium by futures and variance swaps. Technically, agents' $\mathbb{Q}$-measures are Gaussian in our model, and a normal distribution is fully characterized by its mean and variance, implying that first and second-order price derivatives fully align consumers' entire conditional-certainty-equivalent pricing kernels (marginal utilities), which is exactly the condition for optimal risk sharing in the sense of Borch (1962).

1.1 Related Literature

We contribute to literatures on futures contracts, variance swap markets, and equilibrium in incomplete markets.

Futures. Suppose a farmer produces a constant x units of crop, and sells at a random price p; clearly, the optimal hedge is to short x futures contracts, which fully neutralizes price risk. Hedging is more complicated when x is random, and may be related to p through market clearing. Solving for optimal hedging and equilibrium futures market outcomes in such an *incomplete markets* setting is a technically difficult problem which has been studied in a sizable literature (Hirshleifer, 1977; Baesel and Grant, 1982; Britto, 1984; Weller and Yano, 1987; Hirshleifer, 1988*a*,*b*, 1989*a*,*b*; Ekeland, Lautier and Villeneuve, 2019). To maintain tractability, existing models typically restrict heterogeneity to a small number of agent types and states; no existing model delivers closed-form expressions for equilibrium outcomes – trade, contract prices, and welfare – under rich heterogeneity. Another set of papers analyze complete-market models of futures trading (Black, 1976; Breeden, 1980; Hirshleifer, 1990); these approaches deliver predictions about futures pricing, but imply trivial or degenerate equilibrium trading volume. Moreover, since they assume market completeness, these models cannot speak to how closely futures markets approximate the first-best setting of Arrow's state-contingent contracts.

Our contribution to this literature is a fully analytic incomplete-markets model of futures markets. With an arbitrary finite number of heterogeneous agents and continuously distributed inventory shocks, our model delivers closed-form and economically intuitive expressions for trade volume, risk premia, and welfare gains relative to the first-best benchmark. This allows us to unify and generalize many insights from the existing literature, such as the relationship between optimal hedges and demand elasticity discussed in McKinnon (1967) and Rolfo (1980).

Options. The pricing of options is a cornerstone of modern asset pricing theory (Black and Scholes, 1973) and the subject of an extensive theoretical and applied literature. Yet the literature has largely failed to address a simple question: what drives trade in options markets, and do they serve any legitimate risk-sharing purpose? The answer to this question is intuitively clear for *futures* contracts: a classic literature has recognized that goods buyers and sellers have offsetting spot price exposures, which they can mutually offset through the use of futures contracts. It is less clear whether any realistic economic agent has natural price *variance* exposures, which can be traded in the market for second-order price risks.

Our paper delivers a simple qualitative answer to this question. Walrasian equilibrium in spot markets causes "market makers" to be naturally long volatility, and "liquidity takers" to be naturally short. Consistent with classic narratives regarding the role of risk-sharing markets in society (Bernstein, 1996; Shiller, 2012), price variance derivatives enable a *division of labor* between the activities of physical production and financial risk-bearing: commodity producers can optimally process inventory in spot markets, and then use variance swaps to offload the resultant financial risks to professional speculators.

General Equilibrium in Incomplete Markets (GEI). We also contribute to a classic literature on general equilibrium in incomplete markets, surveyed in Geanakoplos (1990) and Magill and Quinzii (2002). Classic results in the literature show that there may be significant equilibrium multiplicity (Hart, 1975), and equilibria may not be constrained-efficient (Stiglitz, 1982; Geanakoplos and Polemarchakis, 1986).

Our model simply removes income effects, assuming preferences between goods and money are quasilinear, avoiding most of these pathologies. In the spirit of Marshall (1920), we think of our setting as a *partial equilibrium model of risk sharing.* Our functional form assumptions – CARA utility, quadratic production technologies, and normal shocks – then deliver economically intuitive and analytically tractable equilibrium outcomes. We believe that these technical properties can be generalized beyond the commodity derivative setting, to analyze risk sharing in incomplete financial markets more broadly.

1.2 Paper Summary

The paper proceeds as follows. Section 2 introduces the model. Section 3 characterizes first-best outcomes. Section 4 analyzes outcomes in spot markets, together with complete Arrow securities, and then in the absence of financial markets. Section 5 develops the "conditional certainty equivalent" methodology we use to analyze derivative markets. We analyze futures contracts in Section 6, and variance swaps in Section 7. We discuss our model assumptions and relationship to other literature in Section 8, and conclude in Section 9. We relegate some longer and less informative proofs and derivations to Appendix A.

2 Model

Notationally, we will use bold symbols for vectors, for example writing $\boldsymbol{x}$ to mean the vector $(x_1 \ldots x_N)$.

There are N "types" of consumers indexed by i, with a representative consumer of each type who behaves competitively, ignoring price impact.[1] For expositional simplicity, we will refer to the representative consumer of type i as simply "consumer i". Consumers have CARA utility over monetary wealth, with possibly different risk aversions α_i:

$$U_i\left(W_i\right) = -e^{-\alpha_i W_i} \tag{1}$$

There are two goods: money, and a single commodity. i is endowed with an initial constant amount m_i of money, and all consumers can hold infinitely large positive or negative positions in goods and money. Each consumer i has a quadratic "production technology", which converts any positive or negative quantity y_i of goods into wealth:

$$W_i = m_i + \underbrace{\psi y_i - \frac{y_i^2}{2\kappa_i}}_{\textit{Production Technology}} \tag{2}$$

[1] This is equivalent to assuming there is a unit measure of identical atomistic consumers of each type, who behave competitively because their trades are too small to move prices.

CARA utility implies that money endowments m_i have no effect on i's behavior, since m_i simply scales $U_i(W_i)$ in (1) by a constant factor; thus, we proceed to set $m_i = 0$ for all i, so we can write W_i simply as a function of y_i:

$$W_i(y_i) = \psi y_i - \frac{y_i^2}{2\kappa_i} \tag{3}$$

Wealth $W_i(y_i)$ consists of a linear component ψy_i, which pays the consumer ψ per unit of the commodity; and a quadratic "inventory cost" component $\frac{y_i^2}{2\kappa_i}$, which implies that the marginal monetary value of the good is decreasing in the amount of the good held. Consumers with higher κ_i have lower inventory costs, and thus more elastic demand for the good. We will allow the edge case of $\kappa_i = 0$: we interpret such a consumer as having no capacity to hold the commodity, so she has perfectly inelastic demand for exactly $y_i = 0$ units of the commodity, and attains $-\infty$ wealth under any other value of y_i.

We call $W_i(y_i)$ a "production technology" because it is intuitive to think of y_i being literally transformed into units of consumable wealth. After "transformation" of y_i, the economy reduces to a single-good problem: each consumer has some amount of produced wealth, which can be redistributed across consumers arbitrarily, since money is tradable and consumers have deep pockets. Of course, it is isomorphic to think of $W_i(y_i)$ as a preference function for y_i rather than a production technology; in these terms, consumer i gets utility equivalent to having $W_i(y_i)$ extra dollars from having y_i units of the commodity.

Uncertainty in the model arises from *inventory shocks*: consumer i begins with a random endowment x_i of the commodity. We assume the x_i are independent normal random variables, with means μ_i and variances σ_i^2 that may vary across consumers. If i receives inventory shock x_i, and purchases q_i of the commodity at price p per unit, her final wealth is:

$$W_i = \psi(x_i + q_i) - \frac{(x_i + q_i)^2}{2\kappa_i} - pq_i \tag{4}$$

We will impose the normalization that the mean of the average inventory shock is 0:

$$E\left[\sum_{i=1}^{N} x_i\right] = \sum_{i=1}^{N} \mu_i = 0 \tag{5}$$

Appendix A.1 shows that (5) is purely a normalization and has no substantive content, because any nonzero average in inventory shocks can be absorbed into the definition of ψ.

There are two periods. The first period is a market for *risk*: consumers may trade *financial securities* which alter their endowments of goods or money in future states of the world. In the following sections, we will analyze three financial market structures: complete financial markets with Arrow securities; no financial markets; and commodity price-linked derivatives contracts, which we will show constitute an incomplete financial market. We ignore consumption in the first period, so all asset trades in the first period transfer consumption across future states of the world.

Before the second period begins, consumers' inventory shocks x_i are realized. The second period is a market for *goods*, or in traditional terms, a *spot market*: conditional on financial securities trades in period 1 and inventory shock realizations $x_1 \ldots x_N$, consumers trade money for the commodity.

3 The First-Best Outcome

Conditional on any vector of inventory shocks $\boldsymbol{x}$, the social planner can freely reallocate commodities across consumers; that is, the social planner chooses functions $y_1(\boldsymbol{x})$ to $y_N(\boldsymbol{x})$, satisfying, pointwise in $\boldsymbol{x}$, the aggregate resource constraint:

$$\sum_{i=1}^{N} y_i(\boldsymbol{x}) = \sum_{i=1}^{N} x_i \tag{6}$$

It is of course equivalent to assume that the social planner chooses net trade amounts $q_i(\boldsymbol{x})$ rather than final inventories $y_i(\boldsymbol{x})$. The social planner can also freely reallocate wealth across agents, pointwise in $\boldsymbol{x}$. Conditional on the planner's choice of final inventories, society's aggregate wealth is:

$$W(\boldsymbol{y}(\boldsymbol{x})) \equiv \sum_{i=1}^{N} W_i(y_i(\boldsymbol{x})) = \sum_{i=1}^{N} \psi y_i(\boldsymbol{x}) - \frac{(y_i(\boldsymbol{x}))^2}{2\kappa_i} \tag{7}$$

The social planner thus chooses final monetary wealths of agents, which we will call $G_i(\boldsymbol{x})$, subject to the constraint, pointwise in $\boldsymbol{x}$, that:

$$\sum_{i=1}^{N} G_i(\boldsymbol{x}) \leq W(\boldsymbol{y}(\boldsymbol{x})) \tag{8}$$

Thus, in sum, the social planner chooses commodity allocations $y_i(\boldsymbol{x})$ and money allocations $G_i(\boldsymbol{x})$, satisfying (6) and (8). An allocation is *Pareto efficient* if it is not expected-utility dominated by some other allocation; formally, under our assumption of CARA utility, $\tilde{G}_i(\boldsymbol{x})$

Pareto-dominates $G_i(\boldsymbol{x})$ if:

$$E\left[-e^{-\alpha_i \tilde{G}_i(\boldsymbol{x})}\right] \geq E\left[-e^{-\alpha_i G_i(\boldsymbol{x})}\right] \quad \forall i, \quad \text{and } E\left[-e^{-\alpha_i \tilde{G}_i(\boldsymbol{x})}\right] > E\left[-e^{-\alpha_i G_i(\boldsymbol{x})}\right] \quad \text{for some } i \tag{9}$$

To handle probability-zero edge cases, we will additionally strengthen this definition by saying that $\tilde{G}_i(\boldsymbol{x})$ Pareto-dominates $G_i(\boldsymbol{x})$ if $\tilde{G}_i(\boldsymbol{x}) \geq G_i(\boldsymbol{x})$ for all i and all realizations of $\boldsymbol{x}$, and $\tilde{G}_i(\boldsymbol{x}) > G_i(\boldsymbol{x})$ for some i and $\boldsymbol{x}$, even if the set of $\boldsymbol{x}$ values on which the inequality is strict has measure zero.

Notice that, while commodity allocations $y_i(\boldsymbol{x})$ do not explicitly enter into (9), they matter because they constrain money allocations $G_i(\boldsymbol{x})$ through the wealth constraint (8).

Proposition 1. *Pareto-efficient commodity allocations $y_i^*(\boldsymbol{x})$ and money allocations $G_i^*(\boldsymbol{x})$ are characterized by two conditions:* spot market allocative efficiency*, and* optimal risk-sharing*. Spot market allocative efficiency requires that commodity allocations $y_i^*(\boldsymbol{x})$ satisfy:*

$$y_i^*(\boldsymbol{x}) = \frac{\kappa_i}{\sum_{j=1}^N \kappa_j} \sum_{j=1}^N x_j \tag{10}$$

In any efficient spot market allocation, society's aggregate wealth is:

$$W^*(\boldsymbol{x}) \equiv W(\boldsymbol{y}^*(\boldsymbol{x})) = \psi \sum_{i=1}^N x_i - \frac{\left(\sum_{i=1}^N x_i\right)^2}{2\sum_{i=1}^N \kappa_i} \tag{11}$$

Optimal risk-sharing requires that wealth is shared as:

$$G_i^*(\boldsymbol{x}) = C_i + \frac{\alpha_i^{-1}}{\sum_{j=1}^N \alpha_j^{-1}} W^*(\boldsymbol{x}), \tag{12}$$

where $\sum_{i=1}^N C_i = 0$.

Intuitively, in any Pareto-efficient allocation, spot market commodity allocations $y_i^*(\boldsymbol{x})$ must be efficient, in the sense that commodities are distributed in a way which optimally produces money, given consumers' heterogeneous production technologies $W_i(y_i)$. If this were not the case for any realization $\boldsymbol{x}$, society could simply reallocate goods to generate more wealth, and redistribute this wealth to increase all consumers' money-metric utility in state $\boldsymbol{x}$. Expression (10) states that, since all consumers have quadratic inventory costs, the aggregate endowment $\sum_{j=1}^N x_j$ is simply divided among consumers proportional to their inventory capacities κ_i; higher-κ_i consumers have more elastic demand, suffering lower costs for absorbing inventory, and thus absorb a larger fraction of aggregate inventory shocks in equilibrium.

Through the optimal spot market allocations, society simply transforms commodities $\boldsymbol{x}$ into some total monetary wealth $W^*(\boldsymbol{x})$, characterized by (11). Intuitively, when spot markets function optimally, the N consumers' wealth is equivalent to a single representative consumer with inventory capacity:

$$K \equiv \sum_{i=1}^{N} \kappa_i$$

Conditional on spot market optimal allocations, society then faces a simple one-good risk-sharing problem: there is some random total monetary wealth $W^*(\boldsymbol{x})$ which is to be divided amongst risk-averse consumers. Then, Pareto efficiency requires the equalization of the ratio of marginal utility across states. Under our assumption of CARA utility, the classic results of Borch (1962) and Wilson (1968) imply that any Pareto-efficient allocations redistribute risks in wealth, driven by uncertainty in $\boldsymbol{x}$, affinely according to consumers' risk aversions, as in (12).

Proposition 1 shows that Pareto efficiency is a very restrictive criterion in our model: consumers' spot market outcomes are fully pinned down, and wealths are pinned down across states up to consumer-specific constants. Thus, with slight abuse of terminology, we will occasionally refer to the outcomes described in Proposition 1 as "the first-best outcome" in singular form, implicitly ignoring the constant terms in (12).

We now formally prove Proposition 1.

Proof. After the realization of shocks, consumers' utility is quasilinear in money, implying that all Pareto-efficient outcomes must maximize the sum of consumers' monetary-equivalent values of goods; any non-maximizing allocation is Pareto-dominated with transfers. That is, efficient allocations must solve:

$$\max_{y_i} \sum_{i=1}^{N} W_i(y_i) = \max_{y_i} \sum_{i=1}^{N} \psi y_i - \frac{y_i^2}{2\kappa_i}$$

$$s.t. \ \sum_{i=1}^{N} y_i = \sum_{i=1}^{N} x_i$$

The Lagrangian is:

$$\Lambda = \max_{y_i} \left[\sum_{i=1}^{N} \psi y_i - \frac{y_i^2}{2\kappa_i}\right] - \lambda \left(\sum_{i=1}^{N} y_i - \sum_{i=1}^{N} x_i\right)$$

The first-order condition is:

$$0 = \frac{\partial \Lambda}{\partial y_i} = \psi - \frac{y_i}{\kappa_i} - \lambda$$

Implying simply that consumers' marginal rate of substitution between wealth and goods, $\frac{\partial W_i}{\partial y_i} = \psi - \frac{y_i}{\kappa_i}$, must be equated:

$$\frac{y_i}{\kappa_i} = \psi + \lambda$$

Combining this with the resource constraint (6), we get (10), which uniquely characterizes the allocations $y_i^* (\boldsymbol{x})$ which maximize aggregate wealth, conditional on any inventory shock realization $\boldsymbol{x}$. Plugging (10) into consumers' production technology (3) and summing, we then get (11).

To show that (10) is necessary for Pareto efficiency, suppose $y_i (\boldsymbol{x})$ does not satisfy (10) for some i and $\boldsymbol{x}$. For any realization $\boldsymbol{x}$ where (10) is violated, replacing $\boldsymbol{y} (\boldsymbol{x})$ with $\boldsymbol{y}^* (\boldsymbol{x})$ increases total social wealth $W (\boldsymbol{y} (\boldsymbol{x}))$ and loosens the constraint (8). We can thus increase $G_i (\boldsymbol{x})$ for all i, leading to a Pareto improvement. We can also conclude from the above analysis that (8) must be binding.

Suppose now that $y_i (\boldsymbol{x})$ does satisfy (10). The optimal risk-sharing condition is simply the result of Borch (1962) in our setting. Pareto efficiency requires agents to equate the ratio of their marginal utilities across all states:

$$\frac{w_j}{w_i} = \frac{U_i' (G_i (\boldsymbol{x}))}{U_j' (G_j (\boldsymbol{x}))} = \frac{\alpha_i e^{-\alpha_i G_i (\boldsymbol{x})}}{\alpha_j e^{-\alpha_j G_j (\boldsymbol{x})}},$$

where w_i is the weight for i's utility. Combining this with the binding constraint (8), we get (12), following Wilson (1968). Hence, the wealth allocation is uniquely characterized up to agent-specific, state-independent constants. $\square$

4 Spot Market Equilibrium

We next solve for equilibrium in spot markets, and illustrate why spot markets fail to implement the first-best outcome.

After inventory shocks x_i are realized, there is no remaining uncertainty, and our setting reduces to a simple quasilinear-utility Walrasian equilibrium, where the only goods are money and the commodity. From (4), i's wealth is:

$$W_i = C_i + \psi (x_i + q_i) - \frac{(x_i + q_i)^2}{2\kappa_i} - p q_i \tag{13}$$

where q_i is the amount of the commodity i purchases at market price p, and C_i is any monetary endowment i may have attained through financial asset trade in the first period. Differentiating (13) with respect to q_i, we have:

$$\frac{\partial W_i}{\partial q_i} = \psi - \frac{x_i + q_i}{\kappa_i} - p \tag{14}$$

(14) depends on x_i and q_i, but not C_i: preferences are quasilinear, so there are no income effects. Thus, arbitrary wealth transfers in financial markets have no effect on spot market demand, and thus equilibrium prices and quantities.

W_i is concave in q_i; thus, setting (14) to zero and solving for q_i, we obtain consumers' demand for the good as a function of the spot price p, in terms of money:

$$q_i(p) = -x_i - \kappa_i (p - \psi) \tag{15}$$

Hence, the inventory shock x_i determines the intercept of the demand curve, and inventory capacity κ_i determines the slope.

Spot market equilibrium is characterized by a market-clearing scalar price p. Summing consumers' demand and setting to zero, we require:

$$\sum_{i=1}^{N} \left[x_i + \kappa_i (p - \psi)\right] = 0$$

The spot market clearing price is thus simply a function of consumers' inventory shocks:

$$p^{Spot}(\boldsymbol{x}) - \psi = -\frac{\sum_{i=1}^{N} x_i}{\sum_{i=1}^{N} \kappa_i} \tag{16}$$

Intuitively, the equilibrium price deviation from ψ is simply the aggregate inventory shock $\sum_{i=1}^{N} x_i$, divided by the aggregate "inventory capacity", or alternatively the slope of aggregate demand, $\sum_{i=1}^{N} \kappa_i$. Plugging (16) into consumer demand (15), we can calculate consumers' equilibrium inventories:

$$x_i + q_i^{Spot}(\boldsymbol{x}) = \frac{\kappa_i}{\sum_{j=1}^{N} \kappa_j} \sum_{j=1}^{N} x_j \tag{17}$$

That is, i ends up holding a fraction $\frac{\kappa_i}{\sum_{j=1}^{N} \kappa_j}$ of the aggregate inventory shock $\sum_{j=1}^{N} x_j$, implementing the first-best outcome (10). Intuitively, conditional on the realization of inventory shocks, the two-good money-and-commodities market is trivially complete, and the welfare theorems hold. Spot market competitive equilibria are thus *allocatively efficient*, in the sense of always allocating commodities in a way which maximizes society's aggregate monetary wealth.

Note also that our normalization in (5) implies that:

$$\sum_{i=1}^{N} E\left[x_i\right] = \sum_{i=1}^{N} \mu_i = 0$$

This conveniently implies from (16) that the mean of the spot price is equal to ψ, and also that i's expected spot market purchase quantity is equal to $-\mu_i$, since:

$$E\left[q_i^{Spot}\left(\boldsymbol{x}\right)\right] = -E\left[x_i\right] + \frac{\kappa_i}{\sum_{j=1}^{N}\kappa_j}\sum_{j=1}^{N} E\left[x_j\right] = -\mu_i \tag{18}$$

We can calculate the wealth distribution induced by competitive equilibria in spot markets, by plugging equilibrium quantities (17) and prices (16) into consumers' wealth, (13). In Appendix A.2, we show that this simplifies to:

$$W_i^{Spot} = \psi x_i + \frac{\kappa_i}{\sum_{j=1}^{N}\kappa_j}\frac{\left(\sum_{j=1}^{N} x_j\right)^2}{2\sum_{j=1}^{N}\kappa_j} - x_i\frac{\sum_{j=1}^{N} x_j}{\sum_{j=1}^{N}\kappa_j} \tag{19}$$

where we have omitted the constant money term, C_i, for convenience. Alternatively, substituting (16) for $p^{Spot}\left(\boldsymbol{x}\right)$ into (19) and rearranging, we have:

$$W_i^{Spot} = p^{Spot}\left(\boldsymbol{x}\right) x_i + \frac{\kappa_i}{2}\left(p^{Spot}\left(\boldsymbol{x}\right) - \psi\right)^2 \tag{20}$$

where for convenience we do not explicitly write the dependence of W_i^{Spot} on $\boldsymbol{x}$.

4.1 Arrow-Debreu Securities and the First-Best Outcome

In the first period, suppose agents can trade *Arrow securities*, denominated in units of wealth, which fully span the state space. Markets are then complete, the welfare theorems hold, and market equilibrium implements the first-best outcome.

Working with Arrow securities in high-dimensional state spaces is somewhat technically involved; while we will carefully define these objects here, note that we will not work with Arrow securities outside this subsection. Since our state space $\boldsymbol{x}$ is continuous, Arrow security prices constitute a *state price density* (Duffie, 2010, ch. 2), which we will refer to as $\pi\left(\boldsymbol{x}\right)$.[2]

[2]A subtle difference between our model and the canonical setting is that, since we assume there is no first-period consumption, there is no natural numeraire in our setting. Thus, we leave $\pi\left(\boldsymbol{x}\right)$ defined only up

Let $\theta_i(\boldsymbol{x})$ denote the *security demand function* of consumer i; that is, consumer i purchases securities paying her a net amount $\theta_i(\boldsymbol{x})$ in state $\boldsymbol{x}$. Since there is no first-stage consumption, agents trade money across states of the world by buying Arrow securities in some states and selling them in other states. i's budget constraint is that her total expenditures must integrate to 0 across states:

$$\int \pi(\boldsymbol{x})\,\theta_i(\boldsymbol{x})\,d\boldsymbol{x} = 0 \;\forall i \tag{21}$$

Note that, following tradition in the literature, we absorb the physical probability density $f(\boldsymbol{x})$ into the definition of the state price density $\pi(\boldsymbol{x})$.

Agents purchase Arrow securities to maximize expected utility subject to (21). We will require *market clearing* pointwise in $\boldsymbol{x}$; since Arrow securities are financial assets in zero net supply, asset demands must sum to zero across agents:

$$\sum_{i=1}^{N} \theta_i(\boldsymbol{x}) = 0 \;\forall \boldsymbol{x} \tag{22}$$

Equilibrium is described by a state price density $\pi(\boldsymbol{x})$ and security demands $\boldsymbol{\theta}(\boldsymbol{x})$, such that all consumers are maximizing utility, and markets for Arrow securities clear.

Proposition 2. *When Arrow securities are available, the unique equilibrium state price density is:*

$$\pi(\boldsymbol{x}) = C \cdot \exp\left(-\frac{W^*(\boldsymbol{x})}{\sum_{j=1}^{N} \alpha_j^{-1}}\right) \cdot f(\boldsymbol{x}) \tag{23}$$

where C is an arbitrary positive constant. Agents' asset demands are:

$$\theta_i(\boldsymbol{x}) = W_i^*(\boldsymbol{x}) - W_i^{Spot}(\boldsymbol{x}) \tag{24}$$

where:

$$W_i^*(\boldsymbol{x}) = \frac{\mathbb{E}\left[\exp\left(-\frac{W^*}{\sum_{j=1}^{N}\alpha_j^{-1}}\right)\cdot\left(W_i^{Spot} - \frac{\alpha_i^{-1}}{\sum_{j=1}^{N}\alpha_j^{-1}}W^*\right)\right]}{\mathbb{E}\left[\exp\left(-\frac{W^*}{\sum_{j=1}^{N}\alpha_j^{-1}}\right)\right]} + \frac{\alpha_i^{-1}}{\sum_{j=1}^{N}\alpha_j^{-1}}W^*(\boldsymbol{x}) \tag{25}$$

The equilibrium with Arrow securities is Pareto-efficient.

Technically, Proposition 2 is simply the classical welfare theorems, applied to our setting: if financial markets are complete, and the first-best outcome described in Proposition 1 is

to scale. An equivalent alternative approach would be to choose some arbitrary value of $\boldsymbol{x}$ as the numeraire good.

attainable through financial asset trading, then the first-best outcome must be an equilibrium.

Intuitively, spot markets and financial markets bring us from autarky to first-best efficiency in two stages, corresponding to the two efficiency conditions in Proposition 1. Spot market equilibrium implements the allocative efficiency condition (10), reallocating goods across consumers in each state of the world to optimally convert goods to wealth. This reduces the two-good problem to a one-good problem, where each consumer is effectively endowed with $W_i^{Spot}(\boldsymbol{x})$ dollars in state $\boldsymbol{x}$. However, spot markets do not efficiently distribute this wealth across agents.

Financial market equilibrium implements the risk-sharing condition (12), optimally sharing state-dependent shocks to aggregate wealth $W^*(\boldsymbol{x})$ across consumers. Note that equilibrium wealth with Arrow securities, (25), takes the form of first-best wealth (12), dividing aggregate wealth across agents according to their inverse risk aversions. Equilibrium Arrow security demands (24) are very simple: agents simply purchase the difference between their first-best wealths $W_i^*(\boldsymbol{x})$ and their spot-equilibrium wealths $W_i^{Spot}(\boldsymbol{x})$.

Financial markets are needed because spot market wealths $W_i^{Spot}(\boldsymbol{x})$ generally differ from first-best outcomes $W_i^*(\boldsymbol{x})$. In particular, risk aversions α_i influence first-best wealth allocations $W_i^*(\boldsymbol{x})$, but not spot market outcomes $W_i^{Spot}(\boldsymbol{x})$; clearly, spot market outcomes cannot implement first-best outcomes in general. Spot market outcomes occur after all inventory shock uncertainty $\boldsymbol{x}$ is realized, and thus clearly cannot allow consumers to share $\boldsymbol{x}$-related risk.[3]

We now discuss the differences between $W_i^*(\boldsymbol{x})$ and $W_i^{Spot}(\boldsymbol{x})$ in two examples.

4.2 Market Makers and Liquidity Takers

Suppose we have two types of agents, a "liquidity taker" (T) and a "market maker" (M). T has risk aversion α_T, receives an inventory shock $x_T \sim N(0, \sigma_T)$, and has no capacity to absorb inventory, $\kappa_T = 0$. We might think of T as a group of oil consumers, who have perfectly inelastic demand for a noisy quantity x_T of oil. M has risk aversion α_M, receives no inventory shock, $x_M = 0$, and has positive capacity $\kappa_M > 0$. We might think of M as a group of oil suppliers, who face no uncertainty, but has an upwards-sloping production cost curve, allowing them to supply increasing amounts of oil as prices increase.

[3] Another way to see this is that, in a two-good spot market after $\boldsymbol{x}$ is realized, utility is ordinal rather than cardinal: a consumer's preferences are fully described by indifference curves between money and goods, which are traced out by (13). Expression (19) for $W_i^{Spot}(\boldsymbol{x})$ is thus valid regardless of what consumers' preferences over wealth are: (19) holds for any choice of CARA-utility risk aversions, or indeed any other classes of risk-averse preferences over wealth.

In equilibrium, the market maker simply buys the entire inventory shock from the liquidity taker:

$$q_T = -x_T,\ q_M = x_T$$

From (12), efficient wealths are:

$$W_T^* = C_T + \frac{\alpha_T^{-1}}{\alpha_T^{-1} + \alpha_M^{-1}} \left(\psi x_T - \frac{1}{2} \frac{x_T^2}{\kappa_M} \right) \tag{26}$$

$$W_M^* = C_M + \frac{\alpha_M^{-1}}{\alpha_T^{-1} + \alpha_M^{-1}} \left(\psi x_T - \frac{1}{2} \frac{x_T^2}{\kappa_M} \right) \tag{27}$$

From (19), equilibrium wealths are:

$$W_T^{Spot} = \psi x_T - \frac{x_T^2}{\kappa_M},\ W_M^{Spot} = \frac{1}{2} \frac{x_T^2}{\kappa_M} \tag{28}$$

Equilibrium wealths feature two distortions relative to first-best wealth distributions.

Wealth Shock Sharing. In our model, the commodity can have some average value – ignoring the concave term, a unit of the commodity is marginally worth ψ dollars – so inventory shocks serve as wealth shocks. When $\psi > 0$, in the first-best outcome, society is wealthier when the endowment x_T is higher, so this shock is split between T and M according to their inverse risk aversions, as indicated by the coefficients $\frac{\alpha_T^{-1}}{\alpha_T^{-1}+\alpha_M^{-1}}$ and $\frac{\alpha_M^{-1}}{\alpha_T^{-1}+\alpha_M^{-1}}$ on x_T in (26) and (27). In contrast, in spot market equilibrium (28), T keeps the entire linear term ψx_T, leaving M no linear exposure to the asset. Somewhat surprisingly, in Section 6, we will show that futures trading can allow M and T to better share endowment shock risks.

Volatility Exposures. In the social optimum, all agents have concave exposure to aggregate inventory shocks: the coefficients on x_T^2 in (26) and (27) are respectively:

$$-\frac{1}{2\kappa_M} \frac{\alpha_T^{-1}}{\alpha_T^{-1} + \alpha_M^{-1}}, -\frac{1}{2\kappa_M} \frac{\alpha_M^{-1}}{\alpha_T^{-1} + \alpha_M^{-1}}$$

which add up to society's total exposure $-\frac{1}{2\kappa_M}$. Intuitively, agents receive constant shares of society's money-equivalent wealth, which is a concave function of aggregate inventory x_T. Interestingly, in spot market equilibrium, the sign of M's volatility exposure is wrong: M actually *benefits* from larger values of x_T^2. Conversely, T has negative volatility exposure of $\frac{1}{\kappa_M}$, which is greater than society's total exposure $\frac{1}{2\kappa_M}$. Liquidity takers essentially pay for volatility both due to the concavity of society's wealth, and also from a "liquidity tax" paid to market makers.

Realized volatility x_T^2 is a random variable that causes W_T^{Spot} and W_M^{Spot} to be risky.

The offsetting volatility exposures in (28) thus seem clearly inefficient: if M could pay $\frac{x_T^2}{2\kappa_M} - E\left[\frac{x_T^2}{2\kappa_M}\right]$ to T, T's wealth risk would decrease and M's risk would be eliminated entirely, improving expected utility for both. We will show how this can be accomplished with variance swaps in Section 7.

Physical Capacity, Risk Aversion, and the "Division of Labor". In our model, an agent's demand slope κ_i for the spot good is distinct from their risk aversion α_i over wealth. We view this as realistic, and in fact crucial to the value of financial markets. High κ_i agents may be commodity producers: entities like oil producers or crop farmers, who have made large infrastructure investments to produce or store physical goods, but may be fairly risk-averse over monetary wealth. Low α_i agents may be professional speculators: entities like hedge funds, who have high financial risk tolerance, and are willing to harvest risk premia in various financial markets, but may have no capacity to produce and trade physical commodities.

Authors such as Bernstein (1996) and Shiller (2012) have argued that financial markets enable a *division of labor* among these two groups of firms, allowing commodity producers to specialize in physical processing, offloading the resultant wealth risks to professional speculators. This division of labor is clearly visible in the social planner's outcome, and in idealized financial markets: high-κ_i agents physically bear a large share of aggregate inventory fluctuations, but the resultant financial risks are shifted onto low-α_i agents. In later sections, we will illustrate how it is accomplished in more realistic markets for price risk.

4.3 Buyers and Sellers

Next, we add to T, M two additional consumers, called "buyer" (B) and "seller" (S). B and S have nonrandom inventories, $x_B = -1$, and $x_S = 1$, and have no ability to absorb capacity, $\kappa_B = \kappa_S = 0$. Intuitively, B and S respectively want to inelastically buy and sell a unit of the commodity, and their net demands cancel out.

In spot market equilibrium, goods allocations are:

$$q_T = -x_T,\ q_M = x_T,\ q_B = 1,\ q_S = -1$$

That is, T sells her entire endowment x_T to M, and S sells a unit to B. This allocation generates identical societal wealth to the last example, so in the first-best, this wealth is divided between agents according to their risk aversions, for example:

$$W_S^* = C_S + \frac{\alpha_S^{-1}}{\alpha_T^{-1} + \alpha_M^{-1} + \alpha_B^{-1} + \alpha_S^{-1}} \left(\psi x_T - \frac{1}{2}\frac{x_T^2}{\kappa_M}\right) \tag{29}$$

and likewise for B, T, M. Prices are exactly the same as in the previous example, so the wealths of M and T are unchanged from (28). Equilibrium wealths for the buyer and seller are now:

$$W_B^{Spot} = -\psi + \frac{x_T}{\kappa_M},\ W_S^{Spot} = \psi - \frac{x_T}{\kappa_M} \tag{30}$$

Expression (30) illustrates two distortions. First, neither S nor B have any exposure to inventory risk ψx_T and volatility risk $\frac{x_T^2}{2\kappa_M}$. Second, in spot market equilibrium, S inelastically sells a unit of the commodity to B, at a price $-\frac{x_T}{\kappa_M}$ determined by T's inventory shock and M's inventory capacity. Commodity price randomness thus creates variance in W_B^{Spot}, W_S^{Spot}, lowering both parties' ex-ante utility. This is clearly inefficient: these wealth risks are exactly offsetting, so S and B should commit to making a net transfer $\frac{x_T}{\kappa_M}$ to neutralize price risks; equivalently, S and B could simply contract in advance to trade at a fixed price. The classical role of futures contracts is to allow S and B to offset these price risks; we will show in Section 6 how this is accomplished in our model.

5 Conditional Certainty Equivalents and Risk-Neutral Pricing

Building on our analysis of autarkic spot markets, we proceed towards the main goal of the paper: the analysis of realistically incomplete financial markets, consisting of financial derivatives linked to the spot good's price. This section will develop two useful technical results towards this end.

We have thus far worked with inventory shocks x_i; henceforth, we will instead use the demeaned representation:

$$x_i \equiv \mu_i + \epsilon_i$$

where μ_i are constants and ϵ_i are independent mean-zero normal random variables with standard deviations σ_i. From (16), the spot price is then:

$$p^{Spot}(\boldsymbol{x}) - \psi = -\frac{\sum_{i=1}^N (\mu_i + \epsilon_i)}{\sum_{i=1}^N \kappa_i} = -\frac{\sum_{i=1}^N \epsilon_i}{\sum_{i=1}^N \kappa_i} \tag{31}$$

where we used the normalization in (5) that $\sum_{i=1}^N \mu_i = 0$, that is, the mean inventory shock sums to 0 across agents.[4] Thus, $p^{Spot}(\boldsymbol{x}) - \psi$ is normally distributed, with mean 0 and

[4]Recall that $\sum_{i=1}^N \mu_i = 0$ has no substantive content: Appendix A.1 shows that any nonzero average in μ_i can simply be absorbed into the definition of ψ.

variance:

$$\sigma_p^2 \equiv \frac{\sum_{i=1}^N \sigma_i^2}{\left(\sum_{i=1}^N \kappa_i\right)^2} \tag{32}$$

We will also renormalize prices, defining:

$$z \equiv p^{Spot}(\boldsymbol{x}) - \psi \tag{33}$$

where we suppress z's dependence on the vector of inventory shocks $\boldsymbol{x}$ (equivalently, $\boldsymbol{\epsilon}$) for notational convenience. With slight abuse of terminology, we will refer to z as the "spot price" or "normalized spot price". z and ϵ_i are convenient because $(z, \epsilon_1 \ldots \epsilon_N)$ is a vector of mean-0 jointly normal random variables. Substituting z into (20), we can write i's spot equilibrium wealth as:

$$W_i^{Spot} = (\psi + z)(\mu_i + \epsilon_i) + \frac{\kappa_i}{2} z^2 \tag{34}$$

5.1 z-Certainty Equivalents

Consumer i's spot wealth in expression (34) depends on both the common spot price z, which determines the payoffs of spot-price-linked derivatives such as futures and variance swaps; and the idiosyncratic inventory shock ϵ_i. In our model, it turns out we can simply "integrate out" the idiosyncratic ϵ_i terms from utility, transforming the full $(z, \epsilon_1 \ldots \epsilon_N)$ problem into the much simpler problem of risk sharing conditional on a single source of uncertainty, the spot price z, which is common to all agents.

From (31), (z, ϵ_i) is a bivariate normal distribution for any i. We can thus linearly "project" ϵ_i onto z, as follows.

Definition 1. We can write:

$$\epsilon_i = \beta_i z + \eta_i \tag{35}$$

where $\beta_i z$ is the conditional expectation $E[\epsilon_i \mid z]$, and η_i is uncorrelated with z. By computing covariances and variances of z and ϵ_i through (16), we show in Appendix A.4 that:

$$\beta_i = -\left(\sum_{j=1}^N \kappa_j\right) \frac{\sigma_i^2}{\sum_{j=1}^N \sigma_j^2}, \ \nu_i \equiv Var(\epsilon_i \mid z) = \sigma_i^2 \left(\frac{1 - \sigma_i^2}{\sum_{j=1}^N \sigma_j^2}\right) \tag{36}$$

Intuitively, the inventory shock ϵ_i is correlated with the price z, since any individual consumer receiving a high inventory predicts slightly lower spot prices, through the spot market clearing condition (31). (35) simply separates the inventory shock into a z-predictable component, $\beta_i z$, and a "fully idiosyncratic" orthogonal residual η_i. β_i is larger for agents with comparatively high inventory shock variances, σ_i^2, since the price is driven by these agents' inventory shocks to a greater degree; correspondingly, the residual variance fraction $\frac{\nu_i}{\sigma_i^2}$ is smaller for these agents, since prices "explain" a larger share of their inventory shocks.

A useful identity implied by (36), which we will use in later calculations, is:

$$\sum_{i=1}^{N} \beta_i = -\left(\sum_{j=1}^{N} \kappa_j\right)\left(\sum_{i=1}^{N} \frac{\sigma_i^2}{\sum_{j=1}^{N} \sigma_j^2}\right) = -\sum_{i=1}^{N} \kappa_i \tag{37}$$

Intuitively, when we project the aggregate inventory shock $\sum_{i=1}^{N} \epsilon_i$ onto spot prices, we simply recover the slope of aggregate market demand, $-\sum_{i=1}^{N} \kappa_i$.

Substituting (35) into (34) and rearranging, we can express W_i^{Spot} in terms of a z-spanned component, and an idiosyncratic residual:

$$W_i^{Spot} = \underbrace{(\psi + z)(\mu_i + \beta_i z) + \frac{\kappa_i}{2} z^2}_{\textit{Spanned by } z} + \underbrace{(\psi + z)\,\eta_i}_{\textit{Orthogonal to } z} \tag{38}$$

Now, we evaluate the expected utility of an agent i who receives spot wealth W_i^{Spot} plus some arbitrary wealth transfer $t(z)$ which is a function of z (equivalently, of the spot price p^{Spot}).

Claim 1. If consumer i receives $\boldsymbol{\epsilon}$-measurable random wealth $W_i^{Spot} + t(z)$, her expected utility is:

$$\int_{-\infty}^{\infty} -\exp\left[-\alpha_i\left[(\psi + z)(\mu_i + \beta_i z) + \frac{\kappa_i}{2} z^2 + t(z) - \frac{\alpha_i}{2}\nu_i (\psi + z)^2\right]\right] dF(z) \tag{39}$$

Since (39) has no dependence on η_i, i can be thought of as a CARA-utility consumer with wealth that is a z-measurable random variable:

$$(\psi + z)(\mu_i + \beta_i z) + \frac{\kappa_i}{2} z^2 - \frac{\alpha_i}{2}\nu_i (\psi + z)^2 + t(z) \tag{40}$$

Proof. Intuitively, conditional on z, the only randomness in the consumer's utility function is $(\psi + z)\,\eta_i$; the classical mean-variance property of the CARA-normal system means that the consumer's certainty equivalent for this randomness is its variance $\nu_i (\psi + z)^2$ multiplied by half i's risk aversion, $\frac{\alpha_i}{2}$.

Formally, applying the law of iterated expectations, a consumer with wealth $W_i^{Spot} + t(z)$ has expected utility:

$$E\left[U_i\left(W_i^{Spot} + t(z)\right)\right] = E\left[E\left[U_i\left(W_i^{Spot} + t(z)\right) \mid z\right]\right]$$

This expands into:

$$= \int \int - \exp\left(-\alpha_i \left[(\psi + z)(\mu_i + \beta_i z) + \frac{\kappa_i}{2} z^2 + t(z) + (\psi + z)\eta_i\right]\right) dF(\eta_i \mid z)\, dF(z)$$

$$= \int - \exp\left(-\alpha_i \left[(\psi + z)(\mu_i + \beta_i z) + \frac{\kappa_i}{2} z^2 + t(z)\right]\right) \left[\int \exp(-\alpha_i (\psi + z)\eta_i)\, dF(\eta_i \mid z)\right] dF(z)$$

From (35) and (36) in Definition 1, the conditional distribution $F(\eta_i \mid z)$ is normal with mean 0 and variance ν_i, implying:

$$\int \exp(-\alpha_i (\psi + z)\eta_i)\, dF(\eta_i \mid z) = \exp\left(\frac{\alpha_i^2}{2} \nu_i (\psi + z)^2\right)$$

Hence, i's expected utility is equal to (39). (40) is then immediate from inspection. □

Claim 1 "projects out" the idiosyncratic ϵ_i terms onto the common z term. In doing so, we show that an agent who receives W_i^{Spot}, which is affected by both common price risk z and idiosyncratic risk η_i, is equivalent to a consumer who *only* faces common price risk z, with the modified wealth in (40). The basis risk terms, η_i, collapse into z-contingent "certainty-equivalent" penalties $\frac{\alpha_i}{2} \nu_i (\psi + z)^2$: intuitively, this is just the variance of the residual $(\psi + z)\eta_i$ term in (38), multiplied by the classic $\frac{\alpha_i}{2}$ risk aversion adjustment factor in CARA-normal models. This representation holds conditional on arbitrary z-contingent transfers, whether they are from a social planner, or financial asset trading.

We will proceed to analyze financial market trading behavior using the representation in (40), which we will call consumers' *z-certainty-equivalent* wealths.

Definition 2. We define i's *z-certainty-equivalent wealth* $\bar{W}_i$, or *z-CE wealth* in short, as (40), ignoring the $t(z)$ term:

$$\bar{W}_i = A_i + \Delta_i z + \Gamma_i z^2 \tag{41}$$

where:

$$\Delta_i = \mu_i + \psi(\beta_i - \alpha_i \nu_i) \tag{42}$$

$$\Gamma_i = \frac{\kappa_i}{2} + \beta_i - \frac{\alpha_i \nu_i}{2} \tag{43}$$

$\bar{W}_i$ is a z-measurable random variable.

By analogy to option pricing practice, Δ_i and Γ_i are "Greeks": the linear ("delta") and quadratic ("gamma") exposures of z-CE wealth $\bar{W}_i$ to the spot price deviation z. Claim 1 and Definition 2, in eliminating idiosyncratic risk, effectively convert the incomplete-market

(z, η_i) problem into a complete-market problem with modified wealths, where z is the only source of uncertainty.

5.2 Risk-Neutral Measures and Derivative Pricing

Using z-CEs, we can take a "risk-neutral distribution" approach to solving for derivatives demand. Suppose, in the first period, an agent can commit to buying h_i units of a derivative contract; for each unit purchased, i pays a constant price C in all states of the world, and receives a z-contingent payoff $f(z)$ in return. Since i's derivative payoff is z-contingent, Claim 1 applies, and i's effective wealth is her z-CE plus her derivative payoffs:

$$\bar{W}_i + h_i\left(f(z) - C\right) = A_i + \Delta_i z + \Gamma_i z^2 + h_i\left(f(z) - C\right) \tag{44}$$

i's expected utility is thus:

$$E_{\mathbb{P}}\left[U\left(\bar{W}_i + h_i\left(f(z) - C\right)\right)\right] \tag{45}$$

Where $\mathbb{P}$ denotes the physical probability measure over the spot price z, which has mean 0 and variance (32). Differentiating (45), i's optimal choice of contract position h_i must satisfy the first-order condition:

$$E_{\mathbb{P}}\left[U'\left(\bar{W}_i + h_i\left(f(z) - C\right)\right)\left(f(z) - C\right)\right] = 0 \tag{46}$$

Rearranging, we have:

$$\frac{E_{\mathbb{P}}\left[U'\left(\bar{W}_i + h_i\left(f(z) - C\right)\right) f(z)\right]}{E_{\mathbb{P}}\left[U'\left(\bar{W}_i + h_i\left(f(z) - C\right)\right)\right]} = C \tag{47}$$

Under appropriate regularity conditions, the LHS of (47) can be interpreted as a h_i-dependent *change of measure*, which uses $U'\left(\bar{W}_i + h_i\left(f(z) - C\right)\right)$ as the pricing kernel, or Radon-Nikodym derivative:

$$\frac{d\mathbb{Q}_i^{h_i}}{d\mathbb{P}} = \frac{U'\left(\bar{W}_i + h_i\left(f(z) - C\right)\right)}{E_{\mathbb{P}}\left[U'\left(\bar{W}_i + h_i\left(f(z) - C\right)\right)\right]} \tag{48}$$

where we write $\mathbb{Q}_i^{h_i}$ to emphasize that the measure is i-specific, and depends on i's contract purchase quantity h_i. This leads to a simple and classical characterization of i's contract demand.

Claim 2. At i's optimal contract position h_i, we must have:

$$E_{\mathbb{Q}_i^{h_i}}\left[f\left(z\right)\right] = C \tag{49}$$

with $\mathbb{Q}_i^{h_i}$ defined through (48).

Intuitively, consumer i chooses her derivative position h_i, which determines her state-contingent payoffs (44), thus her utility (45) and marginal utility (46), and thus her risk-neutral measure $\mathbb{Q}_i^{h_i}$ through the Radon-Nikodym derivative (48). i adjusts h_i until the classical risk-neutral pricing equation (49) is satisfied: in words, until i's personal risk-neutral probability measure sets the $\mathbb{Q}_i^{h_i}$-expectation of the derivative contract's payoff $f\left(z\right)$ is equal to its price C. We will proceed in the following two sections to use this representation to solve for equilibrium in futures and variance swap markets.

6 Futures Markets

We now analyze *commodity futures contracts*. We will analyze a single *cash-settled futures contract* on the commodity. Cash-settled and "physical delivery" contracts are outcome-equivalent in our model; we focus on cash settlement because contracts which transfer wealth instead of goods across states of the world are more intuitive in our setting.[5]

A consumer holding a long (short) contract position promises to pay (receive) some fixed p^c in the second period, and receives (pays) the uncertain spot price p; equivalently, the consumer pays some fixed $z^c \equiv p^c - \psi$ and is paid the normalized spot price z. We thus call z^c the *contract price*. Let c_i be i's net contract position, and let $c_i\left(z^c\right)$ denote i's contract demand as a function of z^c. Contract markets clear through z^c adjusting until aggregate

[5] Kyle (2007) shows that cash-settled and physical-delivery derivatives are exactly equivalent whenever a set of "microstructure fungibility" conditions hold; our model satisfies these conditions. Essentially, since spot market outcomes are modelled simply through competitive equilibria, it is equivalent to consumer i's budget set whether she receives x units of the commodity through a physical-delivery contract, or xp dollars through a cash-settled contract. Trade quantities vary – if the consumer receives an endowment of x units of the commodity, she must adjust her net trade quantity in spot markets by x – but market clearing in zero-net-supply futures markets imply that these net quantity adjustments cancel out across consumers, leaving net wealths and spot market prices unchanged. This is shown formally in Section 5.4 of Zhang (2022), in a related model to ours. The expositional advantage of using cash-settled contracts in our setting is that, since they only transfer wealth across states of the world, they do not influence consumers' spot market bids, so spot market outcomes are unchanged from the previous section. With physical-delivery contracts, spot wealth outcomes would be identical state-by-state, but consumers' goods trading patterns would be more notationally cumbersome, since consumers would have to adjust spot market bids to account for futures contract "deliveries".

demand for long and short contract positions is equal:

$$\sum_i c_i\left(z^c\right) = 0 \tag{50}$$

Futures contracts only induce wealth transfers; quasilinearity of spot goods preferences imply that the spot market outcomes we analyzed in Section 4 are unaffected by futures market outcomes. Since futures payments are z-measurable, Claim 1 applies, and i's effective z-measurable wealth when purchasing c_i net contracts is simply her z-certainty-equivalent spot market wealth plus her futures payoffs:

$$\bar{W}_i + \left(z - z^c\right) c_i \tag{51}$$

with $\bar{W}_i$ defined in (41). Recall that we defined Δ_i as the linear exposure of i's z-CE wealth $\bar{W}_i$ to the spot price z. From (51), buying c_i contracts adjusts this linear exposure to:

$$\Delta_i + c_i = \mu_i + \psi\left(\beta_i - \alpha_i \nu_i\right) + c_i \tag{52}$$

Market clearing (50) then implies that these delta adjustments sum to zero across consumers, leaving society's aggregate delta unchanged:

$$\sum_i \left(\Delta_i + c_i\right) = \sum_i \Delta_i$$

Futures contracts are thus effectively a way for consumers to trade delta risk.

Building on Claim 2, we first characterize i's risk-neutral distribution when purchasing c_i contracts, which we call $\mathbb{Q}_i^{c_i}$.

Claim 3. Suppose the following regularity condition holds:

$$\Gamma_i > -\frac{1}{2\alpha_i \sigma_p^2} \tag{53}$$

Then, for any contract position c_i and price z^c, i's risk-neutral distribution $\mathbb{Q}_i^{c_i}$ is Gaussian, with mean and variance:

$$E_{\mathbb{Q}_i^{c_i}}(z) = -\alpha_i\left(\Delta_i + c_i\right) V_i \tag{54}$$

$$V_i \equiv Var_{\mathbb{Q}_i^{c_i}}(z) = \frac{1}{\frac{1}{\sigma_p^2} + 2\alpha_i \Gamma_i} \tag{55}$$

with Δ_i, Γ_i defined in (42) and (43) of Definition 2.

Proof. The wealth in (51) implies a risk-neutral distribution:

$$f_{\mathbb{Q}_i^{c_i}}(z) \propto U_i'\left(\bar{W}_i + (z - z^c)\, c_i\right) f_{\mathbb{P}}(z) \tag{56}$$

where we have used simply that, by the definition of Radon-Nikodym derivative, the density of $\mathbb{Q}_i^{c_i}$ is proportional to the density of $\mathbb{P}$ multiplied by $\frac{d\mathbb{Q}_i^{c_i}}{d\mathbb{P}}$, which is the consumer's marginal utility. Now, for $f_{\mathbb{P}}(z)$, (16) implies that z is normally distributed with mean 0 and variance σ_p^2. From (41) and our assumption of CARA utility, marginal utility is:

$$U_i'\left(\bar{W}_i + (z - z^c)\, c\right) = \alpha_i \exp\left[-\alpha_i\left(A_i + \Delta_i z + \Gamma_i z^2 + (z - z^c)\, c\right)\right]$$

The RHS of (56) is thus:

$$f_{\mathbb{Q}_i^{c_i}}(z) \propto \exp\left(-\frac{z^2}{2\sigma_p^2} - \alpha_i\left(\Delta_i z + \Gamma_i z^2 + c_i (z - z^c)\right)\right) \tag{57}$$

Whenever (53) holds, (57) is a *Gaussian* density, with mean and variance (54) and (55) respectively. When (53) does not hold, the coefficient on z^2 in (57) is nonnegative. Thus, the $\mathbb{P}$-expectation of marginal utility is infinite; since under CARA utility $U_i'(W_i) = -\alpha_i U_i(W_i)$, this implies expected utility is $-\infty$.

□

Claim 3 represents the core analytical "trick" in our paper, so we pause to summarize the steps we have taken so far. First, under the quadratic goods-money preferences in (13), spot wealth W_i^{Spot} (20), or (34) in terms of z, is a quadratic function of spot prices z and an idiosyncratic η_i term. Second, from Claim 1, CARA utility and joint normality of z, η_i implies that W_i^{Spot} collapses into the z-contingent "certainty equivalent"$\bar{W}_i$, which is a quadratic function of z, with no idiosyncratic term. Third, CARA utility implies that any z-quadratic wealths combine with the Gaussian physical probability measure $\mathbb{P}$ to create a Gaussian risk-neutral distribution $\mathbb{Q}_i^{c_i}$, whose mean and variance depend on the $\mathbb{P}$-variance σ_p^2; the polynomial "delta" and "gamma" coefficients of $\bar{W}_i$, Δ_i, Γ_i; i's risk aversion α_i; and any contracts c_i that i purchases.[6] This convenient property allows us to straightforwardly solve for futures contract demand.

Proposition 3. *i's contract demand at price z^c is:*

$$c_i(z^c) = -\mu_i + \psi\left(\nu_i \alpha_i - \beta_i\right) - \frac{z^c}{\alpha_i V_i} \tag{58}$$

[6]The "completing the square" approach in (57) is also isomorphic to Bayesian inference with normal priors, as well as the Kalman filter.

with V_i defined in (55). The equilibrium contract price is:

$$z_{eq}^c = \psi \frac{\sum_i \nu_i \alpha_i + \sum_i \kappa_i}{\sum_i \frac{1}{\alpha_i V_i}} \tag{59}$$

Proof. The futures contract pays $f(z) = z$. Thus, (49) of Claim 2 implies:

$$E_{\mathbb{Q}_i^{c_i}}[z] = z^c$$

Using (54) to substitute for $E_{\mathbb{Q}_{c_i}}[z]$, we require:

$$-\alpha_i (\Delta_i + c_i) V_i = z^c \tag{60}$$

Solving for c_i and substituting for Δ_i using (42), we attain (58).

To solve for the equilibrium futures price z_{eq}^c, we sum futures demand across consumers and impose market clearing. From (60), it is instructive to write the market clearing condition as:

$$z_{eq}^c \sum_i \frac{1}{\alpha_i V_i} = -\sum_i \Delta_i \tag{61}$$

Intuitively, z_{eq}^c depends on $\sum_i \Delta_i$, which can be thought of as the economy's aggregate "delta", or linear exposure to the spot price z, divided by the harmonic sum of the product of risk aversions α_i and $\mathbb{Q}_{c_i}$-variances V_i, defined in (55). Substituting for Δ_i using (42), we have:

$$z_{eq}^c \sum_i \frac{1}{\alpha_i V_i} = -\sum_i (\mu_i + \psi(\beta_i - \alpha_i \nu_i)) \tag{62}$$

In (5), we normalized $\sum_i \mu_i = 0$, and from (37) we have $\sum_i \beta_i = -\sum_i \kappa_i$. Thus, we have:

$$z_{eq}^c \sum_i \frac{1}{\alpha_i V_i} = \psi \left(\sum_i \nu_i \alpha_i + \sum_i \kappa_i \right)$$

which gives (59). □

Expression (58) shows that futures demand consists of three terms: μ_i, which reflects i's expected trade volume in spot markets; $\psi\beta_i$, which measures correlations between i's inventory shock and the spot price; and $\psi\nu_i\alpha_i$, which is the product of i's risk aversion α_i and the variance of the idiosyncratic component of her inventory shock, ν_i. We proceed to discuss how each term affects futures trade and prices.

6.1 Expected Spot Market Trades

Suppose $\psi = 0$. From (42), we have:

$$\Delta_i = \mu_i$$

that is, i's natural price exposure is just her expected inventory shock μ_i, which from (18) is negative her expected spot market purchase quantity. From (59), there is no equilibrium futures risk premium, $z_{eq}^c = 0$ and $p^c = \psi = 0$; from (58), we have:

$$c_i\left(z_{eq}^c\right) = -\mu_i$$

Futures trade is thus driven by the classical force of hedging expected trade volumes. In the buyer-seller example of Subsection 4.3, the buyer goes long a unit of the futures and the seller goes short a unit. This is welfare-improving because it eliminates both sides' directional exposure to spot prices, which shows up as the μ_i term in (42).

An idea dating to Keynes (1930) is that "hedging pressure" from expected spot market trades drive futures risk premia: for example, farmers expect to sell crops, hedge by shorting futures, and thus drive futures prices below expected spot prices, a phenomenon termed "normal backwardation". Hirshleifer (1990) shows that this idea does not work in general equilibrium: in a complete-markets model where futures span the entire state space, hedging pressure alone does not generate futures risk premia.

Our results support the view of Hirshleifer, and demonstrate a new intuition behind this result. Individual agents' deltas – and thus their contract demands – are driven by expected spot purchases: sellers have natural long exposures to prices and buyers have natural short exposure. Expected spot purchases thus generate *trade* in futures markets. But, since $\sum_i \mu_i = 0$,[7] these directional exposures always net to zero across agents. The intuition of Keynes (1930) cannot hold in general equilibrium because spot markets always clear, so society as a whole cannot be a buyer or seller of the spot good in expectation. Expected spot trades thus contribute nothing to society's aggregate delta position, and have no effect on futures risk premia in our model.[8]

[7] Once again, while we assumed $\sum_i \mu_i = 0$ in (5), this assumption is purely a normalization. Appendix A.1 shows that any model in which aggregate inventory shocks are not mean-zero can be written as a model with mean-zero inventory shocks, where we simply redefine ψ to absorb the average inventory shock.

[8] As Hirshleifer observes, hedging pressure can create futures risk premia when there are participation constraints; exogenous participation constraints are used to create futures risk premia in many models (Hirshleifer, 1988*a*,*b*; Gorton, Hayashi and Rouwenhorst, 2013; Acharya, Lochstoer and Ramadorai, 2013; Goldstein, Li and Yang, 2014; Goldstein and Yang, 2022).

6.2 Inventory Shock Hedging

When $\psi \neq 0$, consumers' contract demand contains a $-\beta_i \psi$ term. Recall that β_i is the projection coefficient in (35) of Definition 1, capturing the correlation between i's inventory shock ϵ_i and the spot price, determined by (36):

$$\beta_i = -\left(\sum_j \kappa_j\right) \frac{\sigma_i^2}{\sum_j \sigma_j^2}$$

β_i is thus strictly negative, and $-\beta_i \psi$ always has the same sign as ψ. Intuitively, this term captures the fact that the spot price is informative about i's inventory shock, so i can partially hedge inventory risk using futures contracts.

Consider the maker-taker example of Subsection 4.2. Expected shocks μ_T, μ_M are both 0, and basis risk variances ν_T, ν_M are zero because the spot price is perfectly correlated with ϵ_T. The only driver of contract trade is the $\psi \beta_i$ term; we have:

$$\beta_T = -\kappa_M, \ \beta_M = 0$$

Plugging (59) into (58), each agent's equilibrium linear price exposures, $c_i + \Delta_i$, satisfy:

$$c_i\left(z_{eq}^c\right) + \Delta_i = c_i\left(z_{eq}^c\right) + \beta_i \psi = -\psi \kappa_M \left(\frac{\frac{1}{\alpha_i V_i}}{\sum_i \frac{1}{\alpha_i V_i}}\right) \tag{63}$$

Hence, in equilibrium, the liquidity taker T is long futures and the market maker M is short.

We noted in Subsection 4.2 that spot market outcomes leave T holding her entire linear "inventory risk" $x_T \psi$, whereas the first-best outcome involves T sharing this term with M. Somewhat surprisingly, this example shows that futures markets allow T and M to share this inventory risk: T's long futures position pays out when prices are high and x_T is low, reducing her exposure to her own inventory, and correspondingly transferring some exposure to M. Exposure transfers are particularly straightforward in this example because prices are a perfect signal for T's inventory shock.

The $\psi \beta_i$ components of consumers' deltas sum to a simple aggregate delta position of $\psi \sum_i \kappa_i$. Technically, $\sum_i \beta_i$ reflects the coefficient from regressing the aggregate inventory shock $\sum_i \epsilon_i$ on the spot price; from (16), this is equal to the slope of aggregate demand, $\sum_i \kappa_i$. Intuitively, society's aggregate delta, $\sum_i \Delta_i$, represents the marginal relationship of social aggregate wealth to the spot price; an increase of 1 in spot prices implies an increase of $\sum_i \kappa_i$ in aggregate inventory, which is marginally worth $\psi \sum_i \kappa_i$ dollars.

Expression (59) shows that this aggregate delta creates a futures risk premium. If we ignore the $\nu_i \alpha_i$ term – as in the M, T example where ν_T, ν_M are both zero – we can write (59) as:

$$\frac{p^c_{eq} - \psi}{\psi} = \frac{\sum_i \kappa_i}{\sum_i \frac{1}{\alpha_i V_i}} \tag{64}$$

That is, the futures price p^c contains a proportional markup over the $\mathbb{P}$-fair price ψ, which depends on the aggregate delta $\sum_i \kappa_i$ divided by the harmonic sum of social risk-aversion-weighted $\mathbb{Q}$-variance $\sum_i \alpha_i V_i$. Intuitively, since long futures positions hedge negative aggregate inventory shocks, society in the aggregate would like to be long futures at the fair price ψ; the markup (64) must be positive to lower futures demand sufficiently to clear markets. The denominator of (64), $\sum_i \frac{1}{\alpha_i V_i}$, is the futures *demand slope*: it is driven by risk aversion and $\mathbb{Q}$-variances, as in the familiar problem of optimal portfolio choice in normal-exponential mean-variance models.

Returning to the M, T example, substituting (16) into first-best wealth (26), we can write:

$$W^*_T = C_T + \frac{\alpha_T^{-1}}{\alpha_T^{-1} + \alpha_M^{-1}} \left(-\psi \kappa_M \left(p - \psi\right) - \frac{1}{2} \kappa_M \left(p - \psi\right)^2 \right) \tag{65}$$

The linear price exposure in first-best wealth (65) looks quite similar to contract-induced wealth (63); the difference is that there is a $\frac{1}{\alpha_i V_i}$ weighting term in futures equilibrium, versus a $\frac{1}{\alpha_i}$ term in the first-best. In the following section, we will show that variance swaps "fix" this issue, aligning equilibrium weights with the first-best outcome.

6.2.1 Farmers, Inventory Risk, and Optimal Hedging

By combining the μ_i and $\psi \beta_i$ effects, we can address a classic question regarding optimal crop hedging. Suppose a group of farmers, F, receive a random harvest, which they inelastically sell to a market with some aggregate demand slope. Suppose the crop is valuable on average, so $\psi > 0$. We model the farmers as having $\mu_F > 0$, since they are expected sellers; this force tends to push farmers to short the futures contract. We assume all inventory shocks in the economy accrue to farmers, so their variance share is $\frac{\sigma_F}{\sum_i \sigma_i} = 1$, and basis risk is $\nu_F = 0$. Our discussion above showed, counterintuitively, that these inventory shocks lead farmers to have high β_F, which pushes towards farmers being *long* the futures contract.

How do these two forces interact with each other? For simplicity, we ignore contract market equilibrium, and instead assume contracts are $\mathbb{P}$-fairly priced, so $z^c = 0$, or in non-normalized terms, $p^c = \psi$. From (58), farmers' contract demand is then simply:

$$c_F(0) = -\mu_F - \psi\beta_F \tag{66}$$

Now, $\frac{\sigma_F}{\sum_i \sigma_i} = 1$ implies $\beta_F = -\sum_{i=1}^N \kappa_i$. Thus, the sign of c_F is positive if:

$$\frac{\psi}{\mu_F}\sum_{i=1}^N \kappa_i > 1 \tag{67}$$

Recall that $\sum_{i=1}^N \kappa_i$ is the slope of spot market demand; $\psi = p^c$ is the contract price; and μ_i is the expected quantity farmers sell. The LHS of (67) can thus be interpreted as the *demand elasticity* for crops that F faces; farmers are long or short futures depending on whether this elasticity is greater than 1.

Expression (67) is a classic result in the literature on optimal hedging with futures (McKinnon, 1967; Rolfo, 1980): from the perspective of the farmer, (67) simply says that, if demand is sufficiently elastic, revenue decreases when inventory is low despite high prices, so farmers should counterintuitively hedge using net long futures positions. Our model offers a new perspective on this classic result, as driven by the tension between the μ_i-driven, "expected-trade-hedging" role of futures, and the β_i-driven, "inventory-shock-sharing" role. In this light, farmers' long positions can actually improve social efficiency, as we showed in the maker-taker example above. Our model also shows how (67) can be generalized to settings where, for example, inventory shocks accrue to agents other than farmers, or farmers have nonzero demand slopes.

6.3 Basis Risk

The final term in contract demand (58) involves *basis risk*, $\psi\nu_i\alpha_i$. Since ν_i and α_i are both strictly positive, this term always shares the sign of ψ. Technically, spot wealth (38) has an idiosyncratic $(\psi + z)\,\eta_i$ term; Claim 1 translates this into a z-CE wealth penalty proportional to $(\psi + z)^2\,\nu_i\alpha_i$. Intuitively, inventory shocks create wealth risk which is greater when the spot price $(\psi + z)$ is large in absolute value; when ψ is positive, high spot prices increase wealth risk and thus decrease consumers' z-CE wealth.

As in the previous examples, heterogeneity in consumers' natural Δ_i exposures leads to gains from futures trade. Consider a simple example where $\psi > 0$, $\nu_i = \nu$ and $\beta_i = \beta$ are constants, and all agents have $\mu_i = 0$, but agents differ in their risk aversions α_i. Contract demand is then simply:

$$c_i\left(z^c\right)=\psi\left(\nu\alpha_i-\beta\right)-\frac{z^c}{\alpha_i V_i}$$

Intuitively, all agents suffer greater wealth losses from basis risk when prices are higher, generating implicit long delta exposure to spot prices. Agents with larger risk aversion α_i suffer more, in z-CE terms. In equilibrium, higher α_i agents thus take long futures positions, selling some of their wealth risk to lower-α_i agents, who are short futures in equilibrium.

The $\nu_i\alpha_i$ term is always positive, so it does not net out across agents. We can in general write the contract pricing condition, (59), as:

$$\frac{p_{eq}^c-\psi}{\psi}=\frac{\sum_i \nu_i\alpha_i+\sum_i \kappa_i}{\sum_i \frac{1}{\alpha_i V_i}} \tag{68}$$

Intuitively, futures risk premia can in general be thought of as resulting from the aggregate *wealth betas* of spot prices. This has two components: correlations between spot prices and aggregate inventory, captured by the $\sum_i \kappa_i$ term, and the fact that higher prices amplify basis risk and lower agents' z-CEs, captured by the $\sum_i \nu_i\alpha_i$ term. Both terms are always positive, implying that the futures risk premium always has the same sign as ψ.

If markets were complete and agents were able to eliminate the basis risk term $\sum_i \nu_i\alpha_i$, the numerator of (68) would be smaller; in our model, market incompleteness magnifies society's aggregate delta position, because high prices amplify the negative effects of inventory shocks. Note that the effect on the futures risk premium is ambiguous, since market incompleteness also influences agents' $\mathbb{Q}$-variances V_i, by changing Γ_i in (55).

To summarize, futures contracts allow agents to trade their natural directional exposure to spot prices, which we call Δ_i. Optimal risk sharing always involves sharing Δ_i according to agents' risk aversions, and their $\mathbb{Q}$-variances V_i in (55). Agents' Δ_i has three terms – the expected spot trade term, μ_i; the inventory-price correlation term, $\psi\beta_i$; and the spot-price-basis-risk correlation term, $\psi\nu_i\alpha_i$; in general, futures trade can be driven by all three of these forces.

7 Variance Swaps

Next, we add *variance swaps.* A consumer who purchases a unit of a variance swap pays some fixed price q^c in all future states, and receives $z^2=(p-\psi)^2$.[9] A consumer who purchases c_i

[9] Clearly, any noncentered variance swap is payoff-equivalent to some combination of a centered variance swap and a futures contract position.

futures contracts and d_i variance swaps at respective prices z^c, q^c, applying Claim 1 and (41), i attains effective z-measurable wealth:

$$\bar{W}_i + (z - z^c)\, c_i + \left(z^2 - q^c\right) d_i$$

with $\bar{W}_i$ defined in (41). Analogously to the futures case, quadratic derivatives function by adjusting agents' quadratic exposure to prices, to:

$$\Gamma_i + d_i = \frac{\kappa_i}{2} + \beta_i - \frac{\alpha_i \nu_i}{2} + d_i$$

Variance swaps thus allow consumers to trade gamma risk – second-order price risk – leaving society's aggregate gamma unchanged. We first solve for i's $\mathbb{Q}_i^{c_i,d_i}$-measure.

Claim 4. For any futures and variance swap positions c_i, d_i and prices z^c, q^c, suppose d_i satisfies:

$$\Gamma_i + d_i > -\frac{1}{2\alpha_i \sigma_p^2} \tag{69}$$

Then i's risk-neutral distribution $\mathbb{Q}_i^{c_i,d_i}$ is Gaussian, with mean and variance:

$$m_i \equiv E_{\mathbb{Q}_i^{c_i,d_i}}(z) = -\alpha_i (\Delta_i + c_i) V_i \tag{70}$$

$$V_i \equiv Var_{\mathbb{Q}_i^{c_i,d_i}}(z) = \frac{1}{\frac{1}{\sigma_p^2} + 2\alpha_i (\Gamma_i + d_i)} \tag{71}$$

with Δ_i, Γ_i defined in (42) and (43) of Definition 2.

Proof. Analogous to the proof of Claim 3, combining CARA utility with the normal distribution of the spot price z, the density of $\mathbb{Q}_i^{c_i,d_i}$ satisfies:

$$f_{\mathbb{Q}_i^{c_i,d_i}}(z) \propto \exp\left(-\frac{z^2}{2\sigma_p^2} - \alpha_i \left(\Delta_i z + \Gamma_i z^2 + c_i (z - z^c) + d_i \left(z^2 - q^c\right)\right)\right) \tag{72}$$

This again describes a Gaussian density with mean and variance (70) and (71). Intuitively, the $\mathbb{Q}_i^{c_i,d_i}$-density is normal as long as marginal utility is at most a second-order polynomial in z^2; variance swaps simply adjust the z^2 exposure of marginal utility, and thus the normality of $\mathbb{Q}_i^{c_i,d_i}$ is preserved. Analogous to (53) of Claim 3, (69) guarantees that expected marginal utility is finite. □

We then use risk-neutral pricing to solve simultaneously for futures and variance swap demand and equilibrium prices.

Proposition 4. *Let*

$$V\left(q^c, z^c\right) \equiv q^c - \left(z^c\right)^2 \tag{73}$$

be the "price of spot price variance" implied by q^c, z^c. i's contract demand is:

$$d_i\left(q^c, z^c\right) = \frac{1}{2\alpha_i}\left(\frac{1}{V} - \frac{1}{\sigma_p^2}\right) - \Gamma_i \tag{74}$$

$$c_i\left(q^c, z^c\right) = -\mu_i + \psi\left(\nu_i\alpha_i - \beta_i\right) - \frac{z^c}{\alpha_i V} \tag{75}$$

with Δ_i, Γ_i defined in (42) and (43) of Definition 2. Equilibrium prices q^c_{eq}, z^c_{eq} satisfy the system of equations:

$$\left(\frac{1}{V\left(q^c_{eq}, z^c_{eq}\right)} - \frac{1}{\sigma_p^2}\right)\sum_i \frac{1}{\alpha_i} = -\sum_i \left(\kappa_i + \alpha_i\nu_i\right) \tag{76}$$

$$z^c_{eq} = \psi\frac{\sum_i \nu_i\alpha_i + \sum_i \kappa_i}{\frac{1}{V\left(q^c_{eq}, z^c_{eq}\right)}\sum\frac{1}{\alpha_i}} \tag{77}$$

Proof. With two derivatives, (49) must hold for both simultaneously, implying:

$$E_{\mathbb{Q}_i^{c_i,d_i}}\left(z\right) = z^c, E_{\mathbb{Q}_i^{c_i,d_i}}\left(z^2\right) = q^c$$

Now, the $\mathbb{Q}_i^{c_i,d_i}$-expectation of the ψ-centered "variance swap", z^2, can be written as:

$$E_{\mathbb{Q}_i^{c_i,d_i}}\left(z^2\right) = \left(E_{\mathbb{Q}_i^{c_i,d_i}}\left(z\right)\right)^2 + Var_{\mathbb{Q}_i^{c_i,d_i}}\left(z\right) = \left(z^c\right)^2 + Var_{\mathbb{Q}_i^{c_i,d_i}}\left(z\right) \tag{78}$$

where the $\left(z^c\right)^2$ correction term arises because z has $\mathbb{P}$-mean 0, but not $\mathbb{Q}_i^{c_i,d_i}$-mean 0. Substituting for $Var_{\mathbb{Q}_i^{c_i,d_i}}\left(z\right)$ in (78) using (71) and rearranging, we have:

$$\frac{1}{\frac{1}{\sigma_p^2} + 2\alpha_i\left(\Gamma_i + d_i\right)} = q^c - \left(z^c\right)^2 \tag{79}$$

That is, the $\mathbb{Q}_i^{c_i,d_i}$-implied variance on the LHS must equal the q^c, z^c-implied price of variance on the RHS, which we define as $V\left(q^c, z^c\right)$ in (73). Solving for d_i, we attain (74). Note that, as long as the $\mathbb{Q}_i^{c_i,d_i}$-variance $V\left(q^c, z^c\right)$ is positive, we have:

$$d_i\left(q^c, z^c\right) + \Gamma_i = \frac{1}{2\alpha_i}\left(\frac{1}{V\left(q^c, z^c\right)} - \frac{1}{\sigma_p^2}\right) > -\frac{1}{2\alpha_i\sigma_p^2}$$

Hence the regularity condition (69) in Claim 4 holds. Intuitively, i purchases contracts until her $\mathbb{Q}$-variance, which is (71), is equal to the market "price of variance" $\frac{1}{V(q^c,z^c)}$; i's optimized $\mathbb{Q}$-variance is thus finite whenever the market price is positive. While certain choices of d_i deliver unboundedly negative utility to i, i would never optimally purchase these values of d_i. The assumption that $V\left(q^c, z^c\right)$ is positive is needed for the consumer's optimization problem to be well-posed; we show in Appendix A.5 that, if $V\left(q^c, z^c\right)$ is negative, an arbitrage opportunity exists and contract demand is thus unbounded.

The derivation of linear contract demand is analogous to Proposition 3. The $\mathbb{Q}_i^{c_i,d_i}$-mean of z is:

$$-\alpha_i\left(\Delta_i + c_i\right) V_i$$

with V_i defined in (71) of Claim 4 as the $\mathbb{Q}_i^{c_i,d_i}$-variance of z. But (79) implies that this $\mathbb{Q}_i^{c_i,d_i}$-variance is just equal to V as defined in (73); hence, futures demand with variance swaps, (75), is simply futures demand without variance swaps, (58) of Proposition 3, with V in place of V_i.

Imposing market clearing on variance swap demand in (74), we find that the equilibrium price of variance must satisfy:

$$\left(\frac{1}{V\left(q^c_{eq}, z^c_{eq}\right)} - \frac{1}{\sigma_p^2}\right)\sum_i \frac{1}{2\alpha_i} = \sum_i \Gamma_i \tag{80}$$

Intuitively, the variance risk premium $\frac{1}{V} - \frac{1}{\sigma_p^2}$ depends on the aggregate gamma position $\sum_i \Gamma_i$, and risk aversion $\sum_i \frac{1}{2\alpha_i}$. Substituting for Γ_i using expression (43), we have:

$$\left(\frac{1}{V\left(q^c_{eq}, z^c_{eq}\right)} - \frac{1}{\sigma_p^2}\right)\sum_i \frac{1}{2\alpha_i} = \sum_i \left(\beta_i + \frac{\kappa_i}{2} - \frac{\alpha_i}{2}\nu_i\right)$$

Applying (37) and simplifying, we attain (76). Since all terms on the RHS of (76) are negative, the equilibrium $V\left(q^c_{eq}, z^c_{eq}\right)$ is positive and larger than σ_p^2. The linear contract price (77) follows analogously to the proof of Proposition 3.

Expressions (76) and (77) show that z^c-centered variance V turns out to be analytically simpler than the ψ-centered variance q^c, because z^c is the endogenous common mean of consumers' risk-neutral distributions; this is why we state Proposition 4 in terms of $V\left(q^c, z^c\right)$. To solve for equilibrium prices q^c_{eq}, z^c_{eq}, we would calculate the equilibrium risk-neutral variance $V\left(q^c_{eq}, z^c_{eq}\right)$ using (76), calculate the risk-neutral mean z^c_{eq} using (77), and then solve for the implied ψ-centered swap price q^c_{eq} through the $\mathbb{Q}$-variance identity (73). □

7.1 Market Makers and Liquidity Takers

In the real world, variance is traded in *options markets*. Option pricing is one of the foundational results in asset pricing (Black and Scholes, 1973) and the subject of a large and ongoing literature. Yet there is a simple question the literature has not addressed: what role do options markets play in improving social welfare through risk-sharing? There is a straightforward narrative for the role of futures contracts in allowing expected buyers and sellers to mutually hedge risk; but in what circumstances would agents ever have offsetting price *variance* exposures to trade in options markets?

We give a simple answer by applying Proposition 4 to the example of Subsection 4.2, where the "liquidity taker" has natural short exposure to inventory variance, and the "market maker" is naturally long variance. Through (36) and (43), this shows up in their gamma exposures to price variance:

$$\Gamma_T = \beta_T = -\kappa_M, \ \Gamma_M = \frac{\kappa_M}{2} \tag{81}$$

where we have used that, since z is perfectly correlated with T's inventory, $\nu_T = \nu_M = 0$. Exactly analogous to how buyers and sellers neutralize their offsetting *directional* exposures to price risk through futures contracts, M and T can neutralize their offsetting exposures to price *variance* risk using variance swaps.

In the simplest case, suppose T is long $\frac{\kappa_M}{2}$ variance swaps, and M is correspondingly short. M then perfectly eliminates her price variance exposure: her short volatility position pays off in states of the world where the aggregate inventory shock is near 0, so her liquidity provision profits are low. T eliminates half of her negative exposure, since her long swap position pays off when inventory shocks are large, partially offsetting her spot market wealth losses.

M and T cannot fully offset their price variance exposures because, from (81), society's aggregate gamma is negative and equal to half the market maker's demand slope, $\frac{\kappa_M}{2}$. Analogous to the futures case, variance swap markets redistribute society's aggregate gamma risk according to agents' risk capacities. Combining (74) and (76), in market equilibrium, each agent's final gamma position $d_i + \Gamma_i$ satisfies:

$$d_T\left(q^c_{eq}, z^c_{eq}\right) + \Gamma_T = \frac{\alpha_T^{-1}}{\alpha_T^{-1} + \alpha_M^{-1}} \frac{\kappa_M}{2}, \ d_M\left(q^c_{eq}, z^c_{eq}\right) + \Gamma_M = \frac{\alpha_M^{-1}}{\alpha_T^{-1} + \alpha_M^{-1}} \frac{\kappa_M}{2} \tag{82}$$

Substituting inventory shocks for prices using (16), (82) are exactly the efficient quadratic exposures described in (26) and (27). Variance swap markets thus allow T and M to efficiently share price variance risk.

In this example, variance swaps accomplish the division of labor between physical production and financial risk-bearing discussed in Bernstein (1996) and Shiller (2012). It is socially efficient for T to sell her entire inventory to M in spot markets; this causes M to be naturally long volatility, and T to be naturally short. But variance swap markets fix this, by first neutralizing any offsetting volatility exposures, and then allocating the residual price volatility risk to agents according to their inverse risk aversions.

7.2 Markets for Gamma

In the general case, (74) can be interpreted as saying that each consumer sells her natural Γ_i exposure, then purchases an amount of contracts depending on the gap between the $\mathbb{P}$-fair price of a variance swap, σ_p^2, and the market price V of the swap, as well as i's risk aversion α_i. (80) shows that the equilibrium *variance risk premium*, $\frac{1}{V} - \frac{1}{\sigma_p^2}$, depends on society's aggregate Γ_i and aggregate risk aversion. This simplifies substantially, due to the identity (37): the β_i terms in (43) cancel, leaving us with the RHS of (76), $\sum_i (\kappa_i + \alpha_i \nu_i)$.

Intuitively, in our model, society has negative gamma in the aggregate – concave exposure to the spot price – for two reasons. First, demand for the spot good slopes downwards – society in the aggregate has declining marginal utility for the good. Social dollar-valued wealth is thus concave in the aggregate inventory shock $\sum_i \epsilon_i$ – negative inventory shocks hurt more than positive shocks help – and thus is also concave in prices, which are linear in inventory.[10]

Second, in (40) of Claim 1, basis risk enters consumers z-CE wealth multiplicatively with the squared price $(\psi + z)^2$; thus, basis risk is more costly when price variance is higher. This term contributes positively to all consumers' gamma exposures through (43), and thus increases society's aggregate gamma exposure and thus the equilibrium variance risk premium.

[10]Note that, from (11), utility is:

$$W^*(\boldsymbol{x}) = \psi \sum_{i=1}^{N} x_i - \frac{\left(\sum_{i=1}^{N} x_i\right)^2}{2\sum_{i=1}^{N} \kappa_i}$$

Substituting for $\sum_{i=1}^{N} x_i$ using (16), we have:

$$W^*(\boldsymbol{x}) = \psi \left(\sum_{i=1}^{N} \kappa_i\right)(p - \psi) - \frac{\left(\sum_{i=1}^{N} \kappa_i\right)(p - \psi)^2}{2}$$

Hence, society's aggregate wealth exposure to $(p - \psi)^2$ is in fact exactly $\left(\sum_{i=1}^{N} \kappa_i\right)$.

7.3 Basis Risk-Driven Variance Trading

Since basis risk $\alpha_i \nu_i$, also appears in agents' spot gamma exposures, basis risk can also drive variance swap trading. Suppose we have N agents, with $\mu_i = 0$, $\psi = 0$, and identical κ and inventory shock variance σ^2. From (36), we have:

$$\beta_i = \beta = -\kappa,\ \nu_i = \nu = \sigma^2 \left(\frac{N-1}{N} \right)$$

And thus the Greeks are:

$$\Delta_i = 0, \Gamma_i = -\frac{\kappa}{2} - \frac{\alpha_i}{2} \sigma^2 \left(\frac{N-1}{N} \right)$$

Thus, high α_i agents have comparatively high gammas, through the $\alpha_i \nu_i$ term in (43), whereas optimal risk-sharing requires these consumers to hold comparatively small shares of total gamma, as in (82) of the example above. Society's total gamma exposure is:

$$\sum_{i=1}^{N} \Gamma_i = -\frac{N\kappa}{2} - \frac{\sum_{i=1}^{N} \alpha_i}{2} \sigma^2 \left(\frac{N-1}{N} \right)$$

Applying (74) and (76), equilibrium total gamma exposures are:

$$d_i \left(q^c_{eq}, z^c_{eq} \right) + \Gamma_i = \frac{\frac{1}{\alpha_i}}{\sum_i \frac{1}{\alpha_i}} \left(-\frac{N\kappa}{2} - \frac{\sum_{i=1}^{N} \alpha_i}{2} \sigma^2 \left(\frac{N-1}{N} \right) \right)$$

Equilibrium thus involves high-alpha consumers "selling variance" to low-alpha consumers, again moving from consumers' natural spot market Greeks to their efficient Greeks. Note that, in this case, variance swaps do not eliminate basis risk; rather, they allow agents to trade the price-dependent impact of basis risk – the fact that the z-CE cost of basis risk in (40) is proportional to α_i and the price z^2 – thus loading this implicit squared price risk preferentially onto less risk-averse agents.

7.4 Futures Trade and Pricing with Variance Swaps

Variance swap trading influences futures demand by changing agents' $\mathbb{Q}$-variances. This moves the price elasticity of futures demand – the z^c coefficient in (75). This does not qualitatively change futures market outcomes: the basic patterns in the examples we analyzed in the previous section persist when variance swaps are traded alongside futures. However, quantitatively, from (75) and (77) of Proposition 4, and using the definition of Δ_i in (42), agents' equilibrium delta exposures satisfy:

$$c_i\left(q_{eq}^c, z_{eq}^c\right) + \Delta_i = \left(\psi \sum_j \Delta_j\right)\left(\frac{\frac{1}{\alpha_i}}{\sum_j \frac{1}{\alpha_j}}\right)$$

Consider again the maker-taker example of Subsection 4.2; analogous to our derivations in Subsection 6.2 in the futures-only case, we find:

$$c_i\left(q_{eq}^c, z_{eq}^c\right) + \Delta_i = c_i\left(q_{eq}^c, z_{eq}^c\right) - \beta_i \psi = \psi \kappa_M \left(\frac{\frac{1}{\alpha_i}}{\sum_i \frac{1}{\alpha_i}}\right) \tag{83}$$

Expression (83) is exactly first-best wealth, as we wrote in (65). Intuitively, equilibrium outcomes with only linear futures contracts differ from first-best risk-sharing because agents' $\mathbb{Q}$-variances do not agree, adding an extra $\mathbb{Q}$-variance weight, $\frac{1}{\alpha_i V_i}$, in delta-sharing through futures contracts in equilibrium. But variance swap trading perfectly aligns agents' $\mathbb{Q}$-variances: this is exactly the equilibrium condition (76) in Proposition 4. Thus, variance swap trading causes the common V_i term to drop out of (83), leading linear contracts to also implement exactly the inverse-risk-aversion formula that characterizes first-best risk-sharing. We will generalize this result in Proposition 5 below, showing that futures and variance swaps implement "constrained-efficient" risk sharing.

An interesting technical result is that, while variance swaps influence futures trade, they happen not to change equilibrium futures pricing.

Claim 5. The futures risk premium is identical with or without variance swaps, and is equal to:

$$z_{eq}^c = \psi \frac{\sum_i \kappa_i + \sum_i \nu_i \alpha_i}{\frac{1}{\sigma_p^2} \sum \frac{1}{\alpha_i} - \left(\sum \kappa_i + \alpha_i \nu_i\right)}$$

Proof. See Appendix A.6. □

We currently view this result as a convenient mathematical accident, without particularly deep economic content.

7.5 Constrained Optimality

The wealth distribution across states of the world induced by financial market equilibrium with futures and variance swaps is *constrained optimal*, in a sense we formalize in the following proposition.

Proposition 5. *Suppose a social planner begins at the wealth allocation generated in equilibrium by futures and variance swaps. Suppose the planner implements z-contingent wealth transfers:*

$$t_1(z), t_2(z) \dots t_N(z)$$

satisfying budget balance:

$$\sum_i t_i(z) = 0$$

No such choice of $t_i(z)$ can generate a Pareto improvement over the equilibrium wealth allocations in Proposition 4.

Proof. By Claim 1, the CE $\bar{W}_i$ applies to any z-measurable transfers by the planner. Let

$$\bar{W}_i^{FinEq}(z) \equiv \bar{W}_i + \left(z - z_{eq}^c\right) c_i\left(z_{eq}^c\right) + \left(z^2 - q_{eq}^c\right) d_i\left(q_{eq}^c\right)$$

denote financial market equilibrium wealths for i, accounting for equilibrium futures and variance swap positions; clearly $\bar{W}_i^{FinEq}(z)$ is z-measurable. The result of Borch (1962) applies: Pareto-optimal transfer functions $t_i(z)$ must fully equalize the ratio of marginal utilities across states:

$$\frac{U_i'\left(\bar{W}_i^{FinEq}(z) + t_i(z)\right)}{U_j'\left(\bar{W}_j^{FinEq}(z) + t_j(z)\right)} = \frac{\lambda_j}{\lambda_i}$$

Equivalently, we require that marginal utilities are proportional with probability 1:

$$U_i'\left(\bar{W}_i^{FinEq}(z) + t_i(z)\right) = \frac{\lambda_j}{\lambda_i} U_j'\left(\bar{W}_j^{FinEq}(z) + t_j(z)\right) \tag{84}$$

Thus, all agents' pricing kernels (Radon-Nikodym derivatives), $\frac{d\mathbb{Q}_i^{c_i,d_i}}{d\mathbb{P}}$ in (48), must be equal $\mathbb{P}$-almost-surely.

We simply show that (84) holds without transfers, that is, with $t_i(z) = 0$ for all i $\mathbb{P}$-a.s. In financial market equilibrium with futures and variance swaps, in Proposition 4, requires all agents i to have equal $\mathbb{Q}_i^{c_i,d_i}$-means and variances of p. From Claim 4, all agents' $\mathbb{Q}_i^{c_i,d_i}$-measures are normal, and a normal distribution is fully characterized by its mean and variance. Thus, all agents have identical $\mathbb{Q}_i^{c_i,d_i}$-measures, hence (84) holds $\mathbb{P}$-a.s. without transfers, proving the result. □

As the proof demonstrates, Proposition 5 holds because our distributional assumptions generate Gaussian $\mathbb{Q}$-measures, and the Gaussian distribution is a two-parameter family. Thus, linear and quadratic derivatives, in aligning agents' $\mathbb{Q}$-means and variances, in fact fully align agents' risk-neutral distributions over z, immediately implying that no z-measurable transfers can further improve risk sharing. Proposition 5 thus illustrates that our model – in staying within the two-parameter Gaussian family – essentially gives up on analyzing risk beyond second moments; the upside is that we attain a very tractable model of risk up to second-order effects.

Note that the classic definition of constrained efficiency in Geanakoplos and Polemarchakis (1986) requires that the planner can only reallocate existing financial assets. Our constraint in Proposition 5 is different: our planner can choose arbitrary transfer functions, but transfers must be fully contingent on the spot price z. We thus do not use "constrained efficiency" in the statement of Proposition 5 to avoid confusion with Geanakoplos and Polemarchakis.

8 Discussion of Model Assumptions

Formally, we have analyzed a model of *general equilibrium with incomplete markets*, sometimes referred to as GEI. Our key simplification is that we assume consumers' preferences over goods and money are *quasilinear*: in (3) of Section 2, consumers' wealth is the sum of money and a concave function of commodity holdings. Thus, in spot markets, we have a *partial-equilibrium* supply-demand model, in the spirit of Marshall (1920). Equilibria in spot markets are unique; equilibrium prices and quantities are unaffected by the distribution of "money" across agents; and equilibria increase society's money-metric welfare by an amount equal to the classic consumer and producer surplus triangle areas.

By further assuming consumers have concave utility over money-metric wealth, we generalize partial equilibrium analysis to settings with *risk*. This drives a clean separation in our model between the roles of goods markets and risk markets. Spot markets facilitate *allocative efficiency*, maximizing the sum of money-metric wealth $W_i(y_i)$ across consumers. Financial markets facilitate *risk-sharing*, redistributing this wealth depending on consumers' risk preferences; but quasilinearity implies there are no income effects, so financial market trade has no effect on spot market outcomes. Without quasilinearity, income effects imply that financial market trades influence spot market behavior in possibly complex ways.

Income effects also drive the classic pathologies in the GEI literature: there may be multiple equilibria, and equilibria need not be efficient or even constrained-efficient (Magill and Quinzii, 2002). We thus avoid these pathologies: financial market equilibria in our setting are unique and constrained-efficient in the sense of Geanakoplos and Polemarchakis (1986), and separately in the sense of implementing the optimal price-contingent outcomes, as we showed in Proposition 5.

In summary, we think of our setting as a *partial equilibrium incomplete-market model*. Its main benefits are that it is intuitive and tractable relative to general GEI models. Like classic partial-equilibrium models, our setting is less appropriate for studying goods that are "large" enough that income effects on demand are important.

In addition to quasilinearity, we assume that inventory shocks are normally distributed, production technologies are quadratic, and utility is CARA. These are core to our analytical solutions for futures and variance swaps. They imply that basis risk can be "projected" out onto prices, allowing us to construct z-certainty-equivalents in Claim 1. They also imply that z-CE wealth is ultimately a quadratic function of spot prices, which implies that agents' risk-neutral distributions are normal, which we use to solve for demand and equilibrium for futures and variance swaps.

Differences from CARA-Normal Models. There are subtle but important differences between our model and the standard "CARA-normal" setting. In many models, y_i is thought of as a *financial asset*, which has risky direct monetary payoffs ψy_i, where ψ is a random variable. This generates a risk-sharing motive for trade in financial markets: prior to the realization of ψ, holding y_i is risky, so consumers trade the financial asset according to their risk aversions. However, there is no spot market in these models: after ψ is realized, the financial asset is equivalent to money, and there is no further motive for trade.

Our model is different because y_i represents not a financial asset, but a generalized *commodity*, which is distinct from money even after uncertainty is realized. The good is *converted* to money (or money-equivalents), through the quadratic agent-dependent production technology in (2). After inventory shocks x_i are realized, even though there is no residual uncertainty, consumers trade to improve *allocative efficiency*, moving goods into the hands of those who are most effective at converting goods into money. In reality, our model could be thought of as representing commodities such as oil or wheat being sold from producers and intermediaries to consumers. For simplicity, we assume ψ is constant, so all price risk is generated by inventory shocks x_i.

The distinction between spot good capacity and risk aversion plays a crucial role in our model. Implicit in this separation is the assumption that there is no fundamental link between the ability to efficiently and elastically supply commodities, captured by κ_i, and the ability to tolerate large fluctuations in wealth, captured by α_i^{-1}. We believe this is a realistic assumption; it is also the basis for the classical narrative that the role of derivative markets is to enable the division of labor between commodity processors, who handle physical commodities, and professional speculators, who handle price risks (Bernstein, 1996; Shiller, 2012).

Our model also abstracts entirely from common values, information frictions, and other forces analyzed in the rational expectations equilibrium literature following Grossman and Stiglitz (1980). We assume futures are traded before agents receive any signals, so there is no scope for adverse selection. The only friction in our setting is market incompleteness: Arrow-style state-contingent contracts exist, but fail to span the full state space.

We also abstract from other theoretical forces studied in other literatures, such as dynamics, market power, moral hazard, and imperfect rationality. Combining such elements with our framework may be potentially interesting directions for future work.

9 Conclusion

Modern financial markets are to a large extent *markets for price risk*. This paper constructs and solves a partial-equilibrium risk-sharing model of these markets. Spot market equilibria, in optimizing the allocation of goods across agents, naturally endow agents with certain risk exposures to spot prices. The role of financial markets is to allow agents to trade these exposures, optimally sharing the components of wealth risk that are spanned by spot prices. But price derivatives do not span the state space, and leave agents exposed to unhedgeable, idiosyncratic risks that are orthogonal to spot market prices. Financial asset pricing, in equilibrium, is driven by society's aggregate exposure to price risks, adjusted for the effects of the incompleteness of price risk markets.

We have introduced a set of functional form assumptions – CARA utility, quadratic production technologies, and normal shocks – which together result in surprisingly tractable equilibrium outcomes. While this paper focuses on price derivatives, we believe our core theoretical results – z-certainty-equivalents, Gaussian risk-neutral measures, and the resultant "market for Greeks" characterization of equilibrium in financial markets – can be generalized to the analysis of other financial assets. Our technical contribution to the literature is thus an analytically solvable model of financial asset trade in incomplete markets, which resolves some of the classic analytical difficulties in the incomplete markets literature (Hart, 1975; Geanakoplos and Polemarchakis, 1986). We hope this contribution stimulates further research in developing tractable applied-theory models of risk sharing in incomplete financial markets.

References

Acharya, Viral V., Lars A. Lochstoer, and Tarun Ramadorai. 2013. "Limits to arbitrage and hedging: Evidence from commodity markets." *Journal of Financial Economics*, 109(2): 441–465.

Arrow, Kenneth J. 1964. "The role of securities in the optimal allocation of risk-bearing." *The Review of Economic Studies*, 31(2): 91–96.

Baesel, Jerome, and Dwight Grant. 1982. "Equilibrium in a futures market." *Southern Economic Journal*, 49(2): 320–329.

Bernstein, Peter L. 1996. *Against the Gods: The Remarkable Story of Risk.* Wiley New York.

Black, Fischer. 1976. "The pricing of commodity contracts." *Journal of Financial Economics*, 3(1–2): 167–179.

Black, Fischer, and Myron Scholes. 1973. "The pricing of options and corporate liabilities." *Journal of Political Economy*, 81(3): 637–654.

Borch, Karl. 1962. "Equilibrium in a reinsurance market." *Econometrica*, 30(3): 424–444.

Breeden, Douglas T. 1980. "Consumption risk in futures markets." *The Journal of Finance*, 35(2): 503–520.

Britto, Ronald. 1984. "The simultaneous determination of spot and futures prices in a simple model with production risk." *The Quarterly Journal of Economics*, 99(2): 351–365.

Duffie, Darrell. 2010. *Dynamic asset pricing theory.* Princeton University Press.

Ekeland, Ivar, Delphine Lautier, and Bertrand Villeneuve. 2019. "Hedging pressure and speculation in commodity markets." *Economic Theory*, 68(1): 83–123.

Geanakoplos, John. 1990. "An introduction to general equilibrium with incomplete asset markets." *Journal of Mathematical Economics*, 19(1–2): 1–38.

Geanakoplos, John D., and Heraklis M. Polemarchakis. 1986. "Existence, regularity, and constrained suboptimality of competitive allocations when the asset market is incomplete." In *Uncertainty, Information and Communication: Essays in Honor of Kenneth J. Arrow.* , ed. Walter P. Heller, Ross M. Starr and David A. Starrett, 65–95. Cambridge University Press.

Goldstein, Itay, and Liyan Yang. 2022. "Commodity financialization and information transmission." *The Journal of Finance*, 77(5): 2613–2667.

Goldstein, Itay, Yan Li, and Liyan Yang. 2014. "Speculation and hedging in segmented markets." *The Review of Financial Studies*, 27(3): 881–922.

Gorton, Gary B., Fumio Hayashi, and K. Geert Rouwenhorst. 2013. "The fundamentals of commodity futures returns." *Review of Finance*, 17(1): 35–105.

Grossman, Sanford J., and Joseph E. Stiglitz. 1980. "On the impossibility of informationally efficient markets." *American Economic Review*, 70(3): 393–408.

Hart, Oliver D. 1975. "On the optimality of equilibrium when the market structure is incomplete." *Journal of Economic Theory*, 11(3): 418–443.

Hirshleifer, David. 1988*a*. "Residual risk, trading costs, and commodity futures risk premia." *The Review of Financial Studies*, 1(2): 173–193.

Hirshleifer, David. 1988*b*. "Risk, futures pricing, and the organization of production in commodity markets." *Journal of Political Economy*, 96(6): 1206–1220.

Hirshleifer, David. 1989*a*. "Determinants of hedging and risk premia in commodity futures markets." *Journal of Financial and Quantitative Analysis*, 24(3): 313–331.

Hirshleifer, David. 1989*b*. "Futures trading, storage, and the division of risk: A multiperiod analysis." *The Economic Journal*, 99(397): 700–719.

Hirshleifer, David. 1990. "Hedging pressure and futures price movements in a general equilibrium model." *Econometrica*, 58(2): 411–428.

Hirshleifer, Jack. 1977. "The theory of speculation under alternative regimes of markets." *The Journal of Finance*, 32(4): 975–999.

Keynes, John Maynard. 1930. *A Treatise on Money.* Macmillan & Company.

Kyle, Albert S. 2007. "Cash settlement, price manipulation, and the Modigliani-Miller theorem." Working paper.

Magill, Michael, and Martine Quinzii. 2002. *Theory of Incomplete Markets.* Vol. 1, MIT Press.

Marshall, Alfred. 1920. *Principles of Economics.* . 8 ed., Macmillan.

McKinnon, Ronald I. 1967. "Futures markets, buffer stocks, and income stability for primary producers." *Journal of Political Economy*, 75(6): 844–861.

Rolfo, Jacques. 1980. "Optimal hedging under price and quantity uncertainty: The case of a cocoa producer." *Journal of Political Economy*, 88(1): 100–116.

Shiller, Robert J. 2012. *Finance and the Good Society.* Princeton University Press.

Stiglitz, Joseph E. 1982. "The inefficiency of the stock market equilibrium." *The Review of Economic Studies*, 49(2): 241–261.

Weller, Paul, and Makoto Yano. 1987. "Forward exchange, futures trading, and spot price variability: A general equilibrium approach." *Econometrica*, 55(6): 1433–1450.

Wilson, Robert. 1968. "The theory of syndicates." *Econometrica*, 36(1): 119–132.

Zhang, Anthony Lee. 2022. "Competition and manipulation in derivative contract markets." *Journal of Financial Economics*, 144(2): 396–413.

Internet Appendix

A Omitted Proofs and Derivations

A.1 Justification of Mean-Zero Inventory Shocks

It is without loss of generality to assume (5) – that the aggregate inventory shock, $\sum_{i=1}^{N} x_i$, has mean 0 – because of a redundancy in the way we specify consumers' wealth W_i: the linear term ψ and the inventory shock x_i can be renormalized in a way that keeps consumer utility unchanged. We state this in the following simple claim, which we prove in Appendix A.1.1 below.

Claim 6. For any constant A, define:

$$\tilde{\psi} \equiv \psi + A, \;\; \tilde{x}_i \equiv x_i + \kappa_i A \tag{85}$$

Then consumer i's wealth – ignoring the price term $-pq_i$, which is unaffected – can be written as:

$$W_i = \tilde{\psi}\left(\tilde{x}_i + q_i\right) - \frac{\left(\tilde{x}_i + q_i\right)^2}{2\kappa_i} + C_i\left(A\right) \tag{86}$$

where $C_i\left(A\right)$ is a constant that does not depend on q_i or x_i.

Claim 6 implies that we can "renormalize" the constant term ψ, increasing it by any constant A across all consumers, as long as we correspondingly renormalize inventory shocks x_i. Intuitively, since ψx_i is simply a linear component of preferences, increasing ψ by A can be offset by shifting each x_i by $\kappa_i A$, up to an additive constant in wealth. Since the scaling in (85) is linear in A, this immediately implies that, for any set of original inventory shocks x_i which do not have 0 mean across consumers, we can find some A to normalize ψ and inventory shocks, which leads the resultant inventory shocks to have zero mean across consumers. This choice of A is simply:

$$\sum_{i=1}^{N} E\left[\tilde{x}_i\right] = \sum_{i=1}^{N} E\left[x_i\right] + A \sum_{i=1}^{N} \kappa_i = 0$$

$$\implies A = -\frac{\sum_{i=1}^{N} E\left[x_i\right]}{\sum_{i=1}^{N} \kappa_i}$$

As a result, it is completely without loss of generality – that is, it is simply a renormalization of agents' utility functions – to assume that the expected sum of inventory shocks across consumers is 0, as we do in (5). As we show in (18) of Section 4, (5) is a natural normalization, because it implies that μ_i is equal to negative i's expected trade volume in spot markets.

A.1.1 Proof of Claim 6

Note that (85) implies:

$$x_i + q_i = (\tilde{x}_i + q_i) - \kappa_i A$$

Substituting for $(x_i + q_i)$, we can write W_i as:

$$W_i = \psi\left((\tilde{x}_i + q_i) - \kappa_i A\right) - \frac{\left((\tilde{x}_i + q_i) - \kappa_i A\right)^2}{2\kappa_i}$$

$$= \psi\left(\tilde{x}_i + q_i\right) - \psi\kappa_i A - \frac{\left(\tilde{x}_i + q_i\right)^2}{2\kappa_i} + A\left(\tilde{x}_i + q_i\right) - \frac{\kappa_i A^2}{2}$$

$$= (\psi + A)\left(\tilde{x}_i + q_i\right) - \frac{\left(\tilde{x}_i + q_i\right)^2}{2\kappa_i} - \kappa_i\left(\psi A + \frac{A^2}{2}\right)$$

This is the RHS of (86), with the constant:

$$C_i(A) \equiv -\kappa_i\left(\psi A + \frac{A^2}{2}\right)$$

A.2 Derivation of Spot Market Equilibrium Wealth (19)

Copying (13), consumers' wealth is:

$$W_i = \psi\left(x_i + q_i\right) - \frac{\left(x_i + q_i\right)^2}{2\kappa_i} - pq_i \tag{87}$$

Rearranging slightly, we have:

$$W_i = \psi x_i - \frac{\left(x_i + q_i\right)^2}{2\kappa_i} - (p - \psi)\, q_i$$

Substituting equilibrium quantities (17) and prices (16), we have:

$$W_i = \psi x_i - \frac{\left(\frac{\kappa_i}{\sum_{j=1}^N \kappa_j} \sum_{j=1}^N x_j\right)^2}{2\kappa_i} - \left(-\frac{\sum_{j=1}^N x_j}{\sum_{j=1}^N \kappa_j}\right)\left(\frac{\kappa_i}{\sum_{j=1}^N \kappa_j} \sum_{j=1}^N x_j - x_i\right)$$

Simplifying, we attain (19).

A.3 Proof of Proposition 2

Spot market equilibrium endows agent i with $W_i^{Spot}(\boldsymbol{x})$ wealth in state $\boldsymbol{x}$. When agents can trade Arrow securities, markets are trivially complete, so the first welfare theorem implies that equilibrium allocations are Pareto-efficient. We use $W_i^*(\boldsymbol{x})$ to denote i's total equilibrium wealth in state $\boldsymbol{x}$: this is the sum of spot wealth $W_i^{Spot}(\boldsymbol{x})$ and any Arrow security payoffs $\theta_i(\boldsymbol{x})$. Agents' FOC for optimal security demand implies that the state price density is determined by agents' marginal utilities at $W_i^*(\boldsymbol{x})$:

$$\frac{\pi(\boldsymbol{x})}{\pi(\boldsymbol{x}')} = \frac{m(\boldsymbol{x}) \cdot f(\boldsymbol{x})}{m(\boldsymbol{x}') \cdot f(\boldsymbol{x}')} = \frac{U_i'(W_i^*(\boldsymbol{x})) \cdot f(\boldsymbol{x})}{U_i'(W_i^*(\boldsymbol{x}')) \cdot f(\boldsymbol{x}')}.$$

Using the representation of first-best wealth allocations in (12) of Proposition 1, we have:

$$\frac{U_i'(W_i^*(\boldsymbol{x})) \cdot f(\boldsymbol{x})}{U_i'(W_i^*(\boldsymbol{x}')) \cdot f(\boldsymbol{x}')} = \frac{\exp\left(-\frac{W^*(\boldsymbol{x})}{\sum_{j=1}^N \alpha_j^{-1}}\right) \cdot f(\boldsymbol{x})}{\exp\left(-\frac{W^*(\boldsymbol{x}')}{\sum_{j=1}^N \alpha_j^{-1}}\right) \cdot f(\boldsymbol{x}')} \tag{88}$$

which gives (23). Notice that, under CARA utility, all Pareto efficient allocations imply the same state-price density: the constant terms in (12) fall out of the ratio in (88).

To calculate equilibrium Arrow security demands, note that spot market equilibrium endows i with wealth $W_i^{Spot}(\boldsymbol{x})$ in state $\boldsymbol{x}$, and Pareto-efficient wealth allocations $W_i^*(\boldsymbol{x})$ have the form in (12) of Proposition 1. In order for $\theta_i(\boldsymbol{x})$ to induce Pareto-efficient wealth allocations, we must have, for each i:

$$\theta_i(\boldsymbol{x}) = W_i^*(\boldsymbol{x}) - W_i^{Spot}(\boldsymbol{x}) = C_i + \frac{\alpha_i^{-1}}{\sum_{j=1}^N \alpha_j^{-1}} W^*(\boldsymbol{x}) - W_i^{Spot}(\boldsymbol{x}). \tag{89}$$

for some C_i. We can find C_i using the budget constraint (21), substituting (23) and (89):

$$C \int \left(C_i + \frac{\alpha_i^{-1}}{\sum_{j=1}^N \alpha_j^{-1}} W^*(\boldsymbol{x}) - W_i^{Spot}(\boldsymbol{x})\right) \cdot \exp\left(-\frac{W^*(\boldsymbol{x})}{\sum_{j=1}^N \alpha_j^{-1}}\right) f(\boldsymbol{x})\, d\boldsymbol{x} = 0$$

Solving, we have:

$$C_i = \frac{\mathbb{E}\left[\exp\left(-\frac{W^*}{\sum_{j=1}^N \alpha_j^{-1}}\right) \cdot \left(W_i^{Spot} - \frac{\alpha_i^{-1}}{\sum_{j=1}^N \alpha_j^{-1}} W^*\right)\right]}{\mathbb{E}\left[\exp\left(-\frac{W^*}{\sum_{j=1}^N \alpha_j^{-1}}\right)\right]}$$

This gives (25).

A.4 Derivation of Projection Coefficients in Definition 1

Expression (16) for prices gives us:

$$Var\left(z\right)=\frac{\sum_i\sigma_i^2}{\left(\sum_i\kappa_i\right)^2},Cov\left(\epsilon_i,z\right)=-\frac{\sigma_i^2}{\sum_j\kappa_j}$$

The coefficient β_i in (35) is just the OLS regression coefficient of ϵ_i on z:

$$\beta_i=\frac{Cov\left(\epsilon_i,z\right)}{Var\left(z\right)}=-\left(\sum_j\kappa_j\right)\frac{\sigma_i^2}{\sum_j\sigma_j^2}$$

And the residual variance can be calculated by subtracting the variance of $\beta_i z$ from the variance of ϵ_i:

$$\nu_i\equiv Var\left(\epsilon_i\mid z\right)=Var\left(\epsilon_i\right)-\frac{\left(Cov\left(\epsilon_i,z\right)\right)^2}{Var\left(z\right)}$$

$$\nu_i=\sigma_i^2\left(\frac{\sum_{j\neq i}\sigma_j^2}{\sum_j\sigma_j^2}\right)$$

A.5 $V\left(q^c,z^c\right)$ Must Be Positive

Consider a portfolio which is long 1 unit of the variance swap and short $2z^c$ units of the futures contract. The payoff of this portfolio is:

$$\underbrace{\left(z^2-q^c\right)}_{Variance\ Swap}-\underbrace{\left(2z^cz-2\left(z^c\right)^2\right)}_{Futures}$$

$$=\underbrace{\left(z-z^c\right)^2}_{Payoff}-\underbrace{\left(q^c-\left(z^c\right)^2\right)}_{Price} \tag{90}$$

The random payoff $\left(z-z^c\right)^2$ of this position is positive with probability 1. Thus, the constant price $q^c-\left(z^c\right)^2$, which we defined as $V\left(q^c,z^c\right)$ must also be positive; if $q^c<\left(z^c\right)^2$, then this portfolio is an arbitrage – it has positive payoffs net of its price with probability 1 – and demand for the portfolio is unbounded.

A.6 Proof of Claim 5

Rearranging (74), optimal variance swap purchasing implies:

$$\frac{1}{\alpha_i V} = \frac{1}{\alpha_i \sigma_p^2} + 2\left(\Gamma_i + d_i\right) \tag{91}$$

Plugging (91) into futures demand (75) and summing, we can write the futures market clearing condition as:

$$0 \quad = \quad \sum_{i=1}^{N} c_i\left(q^c, z^c\right) \quad = \quad \sum_{i=1}^{N}\left[-\mu_i + \psi\left(\nu_i\alpha_i - \beta_i\right) - z^c\left(\frac{1}{\alpha_i\sigma_p^2} + 2\left(\Gamma_i + d_i\right)\right)\right] \tag{92}$$

Now, variance swaps market clearing requires $\sum_i d_i = 0$, so variance swap positions simply drop out of (92), leaving us with the market clearing condition:

$$\sum_{i=1}^{N}\left[-\mu_i + \psi\left(\nu_i\alpha_i - \beta_i\right) - z^c\left(\frac{1}{\alpha_i\sigma_p^2} + 2\Gamma_i\right)\right] = 0$$

which is exactly the market clearing condition for futures contracts without variance swaps, (62) in the proof of Proposition 3.