Pure Nash Equilibrium Market Recommendations

Chris Parker
Darden School of Business, University of Virginia, Charlottesville, VA parkerc@darden.virginia.edu, https://www.chrisparker.io

Shai Vardi
Daniels School of Business, Purdue University, West Lafayette, IN, svardi@purdue.edu, https://www.shaivardi.com

New business models help sellers make better decisions by communicating information about market prices. However, sellers' inability to coordinate greatly reduces the efficacy of this information and can lead to market failures. We study the feasibility and benefits of providing equilibria as recommendations in economies where sellers face price uncertainty, information scarcity, and an inability to coordinate.

We describe a general model in which sellers wish to sell some quantity of a good at one of several markets with elastic prices, and show that while a pure Nash equilibrium may not necessarily exist, an approximate equilibrium always exists under mild assumptions on the market concentration. We describe an algorithm that a market planner can use to recommend approximate equilibria to sellers, and compare the recommended (approximate) equilibria to selling strategies identified from a survey of onion and potato farmers in India. Using real data on India's agricultural markets, we show that the recommended strategy outperforms the other strategies in terms of seller welfare, geographic price dispersion, market volume concentration, and several metrics of individual farmer welfare improvements. We use parameterized data to show that the framework is robust to imperfect market penetration, is fair to smaller farmers, and captures nearly all of the welfare from a welfare-maximizing allocation.

Key words: Market Design; Weighted Congestion Games; Game Theory; Price of Stability; Simulation; Auctions and Mechanism Design

1. Introduction

Sellers in various economies often need to make decisions in the face of uncertainty. Ridesharing drivers, for example, face a daily decision of where to begin their shift. A driver who lives in a remote location can begin their shift near home, where both demand and supply are low, or drive to a nearby town, where both are high. Farmers in India face a similar challenge: lacking timely and reliable information about prevailing market prices, they do not know the price they will receive for their produce before they arrive at each market, and large inter-market distances make switching to another market infeasible. Both of these scenarios are examples of weighted *congestion games*—non-cooperative games used to model congestion externalities that arise when agents compete for resources.

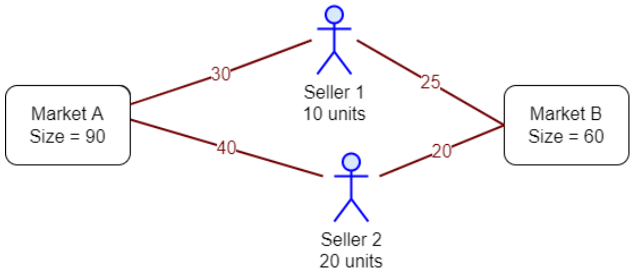


Figure 1: Simple example with 2 sellers and 2 markets.

		Seller 2: A	Seller 2: B
Seller 1	A	$(0, 20)$	$(60, 40)$
	B	$(35, 50)$	$(-5, 20)$

Table 1: Payoff matrix for the example in Figure 1.

To gain insight into the complexities of decisions in congestion games, consider the toy example, consisting of two sellers and two markets, illustrated in Figure 1. Markets A and B have sizes 90 and 60 respectively. Seller 1 has 10 units of a good, and is at distance 30 and 25 from Markets A and B respectively. Seller 2 has 20 units, and is at distance 40 and 20 from Markets A and B respectively. The price per unit at each market is simply its size divided by the total volume at the market.[1] The sellers' utilities are their revenue (volume times price per unit volume) minus the distance from the market they choose. The payoff matrix of this example is given in Table 1. How should the sellers decide which market to travel to? If both sellers choose to travel to their nearest market, they will both go to Market B; if they choose the largest market, they will both go to Market A. In both cases, they will realize low utilities, with Seller 1 possibly even making a loss. Ideally, the sellers would coordinate, and choose either (A, B) or (B, A). Empirical evidence, however, shows that sellers in such markets do not typically coordinate, and instead make suboptimal, sometimes ad-hoc, decisions (e.g., Ashkrof et al. 2020, Camerer et al. 1997).

One way market planners can overcome coordination issues is to use (non-enforceable) recommendations. In the example in Figure 1, a market planner could recommend (A, B) or (B, A). Either of these would lead to an allocation of farmers to markets in which both farmers are better off than either (A, A) or (B, B). In addition to a higher overall and individual welfare, these recommendations lead to other desiderata: there is less price variability,[2] less market concentration and, as both (A, B) and (B, A) are equilibria, neither seller regrets their market choice. This example illustrates the efficacy of such recommendations in simpler contexts and sets the stage for the primary focus of our research: investigating the applicability and effectiveness of similar strategies in more intricate and realistic environments, such as large markets and digital platforms.

[1] Our model allows for more general price elasticity.

[2] Technically, we had not defined the price at a market whose volume is zero; we can define it as some arbitrary high value or assume that all markets are seeded with some initial volume.

Alternatively, one could simply provide information about market prices, which allows sellers to make more informed decisions. For instance, numerous services do or have provided farmers in India with access to previous days' market prices such as Reuters Market Light (RML)[3] and the Government of India's e-National Agriculture Market (e-NAM). With knowledge of (previous days') prices, farmers move demand to markets with high prices and, in doing so, reduce geographic price dispersion by 12% over and above access to mobile phones (Parker et al. 2016). One might think that an increase in information is always desirable: as market participants become more informed, they make decisions that lead to increased welfare and reduced price dispersion. However, this is not necessarily the case. Assume that without any information, both sellers in Figure 1 go to Market A. They now both receive information about the previous day's prices, and both (correctly) deduce that if they had gone to Market B, their utility would have been higher. The next day, they both travel to Market B, which leads to even lower utility. If this process is repeated, the sellers simply oscillate between the markets. Of course, this conclusion is hyperbolic, but it serves to illustrate the difficulty of arriving at an equilibrium without coordination.

1.1. Research Objectives

The primary aim of this study is to explore whether insights from the toy example remain valid in more complex real-world scenarios. That is, can similar recommendations be made in large markets such as Indian agricultural markets and ridesharing platforms, and would they be effective? To address these questions, we take an interdisciplinary approach with analytical modeling combined with data-informed simulation.

In many of these markets, there is an inherent mistrust of the market designer. For example, Indian farmers possess an inherent mistrust of the government (Frayer 2021), and by extension, any authority. In such an environment, sellers would be hesitant to follow any recommendation. Nevertheless, an information service can gain the users' trust by coupling recommendations with verifiable price information. This way, users can reason about whether the recommendation was optimal for them: If the best-response to the recommendation in hindsight was to follow it, the user should be more likely to follow subsequent recommendations. Note that it is crucial that the equilibrium is a *pure* equilibrium, as that ensures that it is an *ex-post* one; recommendations that lead to ex-ante equilibria, such as correlated or mixed equilibria, are unsuitable for economies where trust is fragile. This is because ex-ante equilibria might not yield an ex-post equilibrium; given the ex-post realizations of allocations and prices of ex-ante equilibria, the participants would often (correctly) deduce that, in hindsight, they would have done better had they *not* followed the recommendation.

[3] RML was a business that provided technology and data analytics solutions to over 3 million farmers; its impacts were analyzed by Parker et al. (2016).

We study the feasibility and benefits of providing pure Nash equilibria market recommendations in congestion games such as those above. In particular, we focus on the following questions: (i) Does a pure equilibrium always exist, and if so, can it be computed efficiently? (iii) Do pure equilibrium recommendations yield improvements in seller welfare, price dispersion, and market volume concentration? (iii) How much welfare is lost by demanding that the strategies form an equilibrium? (iii) How robust and fair are these recommendation systems?

1.2. Overview of Results

1.2.1. Existence of pure equilibria. Our setting is modeled as a weighted congestion game. In this game, heterogeneous sellers have a certain amount of goods they wish to sell at a market. The price at each market is decreasing in the supply, and the sellers incur some travel cost. It was previously shown that a pure equilibrium always exists in two special cases of our setting: when all travel costs are zero (Fotakis et al. 2005) and when the sellers have identical volumes (Milchtaich 1996). On the other hand, a more general setting does not necessarily have a pure equilibrium (Milchtaich 2009). Our first result is a negative one: we show that there are instances of our model in which a pure equilibrium does not exist. As a result, it may not be possible to recommend an (exact) equilibrium. Nevertheless, we show that under some mild conditions, there always exists an *equilibrium in the large*.[4] That is, for any ϵ, if the economy is sufficiently large, there is always a strategy profile such that no agent can improve their utility by more than ϵ by deviating.

1.2.2. Understanding Farmer Strategies. In order to determine whether equilibrium recommendations would lead to better economic outcomes than the status quo, we must first understand how the agents currently make their decisions. To better understand the decision-making process, we conducted a survey of 502 onion and potato farmers in Maharashtra, India. Some of the main insights from the survey are (i) farmers appear to mostly make ad-hoc decisions, based on intuition and simple heuristics for choosing a market, and (ii) the problem shown in Example 2—that some markets become congested as a result of neighbors making similar decisions—is one that occurs frequently: approximately 85-90% of farmers state that they go to the same market as neighboring farmers. The survey responses combined with one author's conversations with Indian farmers and agricultural experts lead us to define four baseline strategies of increasing sophistication: (i) traveling to the nearest market, (ii) traveling to the welfare-maximizing market based on information on market size alone, (iii) best-response to the latest information (e.g., yesterday's market prices and volumes), and (iv) iterative best response, where farmers iteratively best-respond until either equilibrium or a cycle is reached.

[4] This terminology was chosen to align with analogous concepts found in economics literature, such as 'strategy-proofness in the large' (Azevedo and Budish 2019).

1.2.3. Economic advantages of recommendations. We compare outcomes from agents following our computed equilibrium recommendations to those from baseline strategies using real data about about different crops sold through India's agricultural markets. We compare the outcomes using seven metrics that collectively capture: (i) individual and total seller welfare, (ii) geographic price dispersion, whose reduction leads to net (producer plus consumer) welfare gains (Jensen 2007), and (iii) market volume concentration. In addition, we perform additional analyses using parameterized synthetic data, which allows us to cleanly identify the sensitivity of outcomes to some key variables in the analysis.

For the real data, both price dispersion and market concentration are better under the recommendations for the most part. For the synthetic data, however, we get more mixed results. When the price is more elastic and the seller volumes are larger, market outcomes are better in both of these metrics (as well as total welfare). When the price is less elastic and the seller volumes are smaller, an equilibrium leads to lower price dispersion than some strategies. A deeper analysis shows that whenever there is a lower price dispersion, there is a higher market concentration: we typically see that the low price dispersion is because the sellers are concentrated in only one or two markets. In instances when the market does not become concentrated, the equilibrium recommendation indeed leads to lower price dispersion in the synthetic data as well.

We also investigate whether the improvement is equitable: do some farmers benefit more than others? The answer is mixed. While all sellers are likely to see an improvement in equilibrium over other strategies, large sellers are more likely to see an improvement than smaller ones. We believe this is because the travel costs affect large and small sellers equally, but the price does not: if both travel cost and price increase by a small amount, this has a more positive effect on the large sellers. However, conditioned on improving their welfare, large and small sellers see a similar percentage improvement distribution with smaller sellers achieving a slightly higher improvement.

Following an equilibrium recommendation almost always leads to higher total seller welfare. In the synthetic data, an equilibrium led to higher welfare in 100% of the simulations. In the real data, we observe that equilibrium recommendations occasionally resulted in marginally lower welfare. However, this only occurred when a large proportion of farmers were such that their cost of travel exceeded their potential revenue. In practice, smaller farmers often sell their product through (or to) aggregators, as approximately 22.5% of surveyed farmers do. While our simulations do not explicitly account for the role of aggregators, we gleaned significant insights into their impact in the market. Aggregators have often been viewed as predatory, utilizing their superior market information to extract rents from farmers (e.g., Abebe et al. 2016). However, we find that their presence can lead to Pareto improvements: when travel costs are relatively high, aggregators can make a profit and farmers can reduce their travel costs and hence increase their profit, without negatively affecting market prices.

In addition to comparing the total seller welfare, we perform comparisons on an individual level. We find that the welfare gain from following the equilibrium recommendation is almost always shared across a large number of sellers. Additionally, the baseline strategies rarely result in a pure equilibrium; typically, most of the sellers regret their choice in hindsight. This contrasts starkly with the recommendation which is designed to be a pure equilibrium and hence no sellers regret their choice.

1.2.4. Robustness and fairness All the advantages of the equilibrium recommendation discussed thus far hold under the assumption that all of the sellers follow the recommendation. We investigate the robustness of the results to imperfect market penetration, i.e., when not all sellers follow the recommendation. This is important as some of the sellers might not use the system or follow its recommendations, particularly during the implementation period. In order to gain the trust of the sellers, the recommendations should be robust to uncertainty: even when not all sellers follow the recommendations, each recommendation should be best-response to the realized choices of all of the sellers (or at least yield comparable utility to the best-response). If this is the case, sellers would then be more likely to follow recommendations in the future. We model the uncertainty using two parameters: service penetration (the percentage of sellers who follow the recommendation) and non-subscriber volume uncertainty (how well the service predicts the total volume that will be realized by the non-subscribing sellers, modeled as the standard deviation in the distribution of the sum of non-subscribed seller volume that is only realized after the sellers reach the market). The results are robust to both of these parameters: even with only 25% penetration and uncertainty about seller volume that is three times the seller's expected size, sellers that follow the recommendation obtain over 99% of both their anticipated welfare (the welfare the recommendation system expected them to obtain) and their best-response welfare (the welfare they could have attained by best-responding to the realized volumes of the remaining sellers). We note that volume uncertainty can also be thought of as subscribed sellers not adhering to the recommendation for any reason. As such, the recommendation is robust to relatively imprecise estimates of volumes from non-subscribed sellers and/or sellers that choose not to follow the recommendation.

1.2.5. Comparison to Welfare-Maximizing Allocations. While our results show that following an equilibrium recommendation leads to better results than those of the baseline strategies, it remains unclear whether better solutions exist. In particular, we consider the following questions: (i) Could we achieve a higher total welfare if we do not insist that the recommendation is an equilibrium? (ii) What are the potential impacts of choosing a suboptimal equilibrium? To answer the first question, we show that the Price of ϵ-Stability[5] is one, when the market is large.

[5] The *Price of ϵ-Stability* is the ratio of the utility of the best ϵ-equilibrium relative to the optimal solution.

That is, under the same conditions as those under which an ϵ-equilibrium exists, requiring that the recommendation is an ϵ-equilibrium leads to no loss of efficiency, as the optimal allocation is an equilibrium itself. A small bound on the Price of Anarchy, the ratio of the worst equilibrium and the optimal solution, would provide an analytical answer to the second question. However, this appears to be much more challenging than computing the Price of Stability. We therefore compare the total welfare of the recommended equilibrium to that of the welfare-maximizing solution using simulations. We do this subject to computational limitations, as it is NP-hard to compute the optimal welfare (Fotakis et al. 2009). We show that our approach captures 99.84% of the optimal farmer welfare on problems where we are able to compute the welfare maximizing strategy. Thus, if the recommended equilibrium is not optimal, it is typically almost optimal; further, when the optimal strategy is not an equilibrium, the recommendation still has the advantage of the sellers not regretting their market choice. All of our results are robust to a wide range of parameters, including market elasticities and seller volume parameters, demonstrating that the proposed system could be used for a broad range of economies.

2. Related Literature

This study is related to several research streams, including decision making and equilibria in congestion markets, and the welfare impacts of equilibria.

2.1. Decision Making in Congestion Markets

There is a large range of strategies that sellers can use to make their decisions. Parker et al. (2016) utilize two strategies to determine whether RML data is actionable and useful for farmers. First, farmers simply try to minimize transportation cost by going to their nearest market and second, slightly more strategic customers who take their produce to whichever market had the highest price on the previous day. They find that utilizing the previous day's price can lead to a higher price for farmers, but ignore that multiple farmers making the same decision will lead to a glut of volume at the market which could drive prices much lower than they were yesterday.

Sellers could also randomize their strategy. When sellers are risk-averse, however, this is unlikely (e.g., Rabin 2000). In Indian produce markets, less than 50% of farmers we surveyed sell their crops more than once a month, and each decision is critical, hence it is likely the farmers are strongly risk-averse. There is evidence that both ride-sharing drivers (Ashkrof et al. 2020) and traditional taxi drivers (Camerer et al. 1997) *do* randomize their strategy (at least in some sense). However, even when sellers randomize, they are likely to reach sub-optimal outcomes: Hsee et al. (2021) recently showed that sellers in similar scenarios suffer from biases that lead to poor overall outcomes. Without coordination or at least additional information, it is unclear how the sellers could reach a mutually beneficial outcome; the current paper provides a framework for inducing (implicit) coordination.

2.2. Equilibria in Congestion Games

Our model can be viewed as a generalization of a congestion game, a non-cooperative game used to model congestion externalities that arise when players compete for resources. In his seminal work, Rosenthal (1973) showed every congestion game has a pure equilibrium, by defining a potential function, whose value decreases whenever an agents decreases their delay. As the potential is bounded from below and the strategy space is finite, this implies the existence of a pure equilibrium, and furthermore shows that such equilibrium can be reached by best-response dynamics.[6] This technique has been extended to many generalizations of this setting (e.g., Cominetti et al. 2022, Gairing et al. 2011). Most relevant to this work, Fotakis et al. (2005) and Milchtaich (1996) show that a pure equilibrium always exists in cases that are special cases of our setting - when all travel costs are zero and when the sellers have identical volumes respectively. On the other hand, Milchtaich (2009) shows that a game that generalizes our setting does not necessarily have a pure equilibrium. While his example does not conform to our setting, we adapt it to demonstrate that our setting also does not always have an equilibrium.

We use a potential function to prove the existence of an approximate, or ϵ-equilibrium. The classical argument by (Monderer and Shapley 1996) shows that in certain games, any myopic improvement by a player leads to an increase in the potential function. In our proof, we also consider unilateral deviations; however, we do not require that the deviation is profitable for the deviating agent, only that it increases the potential. Potential functions have been used in the computation of approximate equilibria in other works. Caragiannis et al. (2015) approximate games that do not have a potential function with ones that do and use them to compute approximate equilibria in the games that do not. Candogan et al. (2013) study games that are close to potential games and show that best (or better) response dynamics converge to approximate equilibria in these games. Hansknecht et al. (2014) use approximate potential functions, where the potential decreases if a player decreases their cost by at least some α.

Even in cases where an equilibrium is known to exist, it may not be easy to find. Fotakis et al. (2009) show that when the sellers do not necessarily have identical volumes, it is NP-hard to find the *best* (social-welfare maximizing) equilibrium, even when all travel costs are zero. Their proof applies to our setting; that is, even when an equilibrium is known to exist, it is NP-hard to compute the best one.

[6] In best-response dynamics, agents take turns best-responding to the others' strategies, which remain fixed.

2.3. Welfare Impacts of Equilibria

Since Aumann's seminal work on correlated equilibria (Aumann 1974), hundreds of papers have studied the advantages of such recommendations, from theoretical, practical and empirical perspectives (e.g., Moreno and Wooders 1998, Papadimitriou and Roughgarden 2008). While correlated equilibria can lead to a higher (expected) welfare than Nash equilibria, they often do not yield ex-post equilibria. Furthermore, Cason and Sharma (2007) showed that when there is a lack of trust, players frequently reject third-party correlated equilibria recommendations.

Two measures for quantifying the efficiency loss from equilibria are the Price of Anarchy (Koutsoupias and Papadimitriou 1999) and the Price of Stability (Schulz and Moses 2003).[7] The Price of Anarchy(PoA) measures how much worse off a system can be when players reach an equilibrium by themselves, while the Price of Stability (PoS) quantifies how much worse the system will be if the *best* equilibrium is reached; the latter presupposes a market planner that can suggest an equilibrium. Our focus in this paper is different; our focus is on how much *better off* the system can be when an equilibrium is suggested as opposed to the efficiency loss compared to the optimum. Nevertheless, the Price of (ϵ)-Stability in our setting is 1 under some assumptions, which shows that the recommended equilibrium can also yield the optimal seller welfare. While we were not able to compute the PoA for our setting, we remark that a low PoA would imply that any equilibrium, and in particular the one that our algorithm computes, guarantees a high total seller welfare.

The efficiency loss of interventions, in particular ones that lead to equilibria, has been studied in many related settings in the Operations Research literature: Scarsini et al. (2018) study dynamic congestion games with atomic players and a single source-destination pair and compute the PoS and PoA on some special cases. Paccagnan and Gairing (2024) study the problem of minimizing social cost in atomic congestion games and show that efficiently computed taxation mechanisms yield good outcomes. Cominetti et al. (2009) study a a routing game, and show that a pricing mechanism can elicit more coordination from the players, and reduce the inefficiency of equilibria.

Similar games with different notions of incomplete information have also been studied. Wu et al. (2021) study a routing game in an uncertain environment, and consider the effects of asymmetric and incomplete information on the travelers' equilibrium route choices and costs. Cominetti et al. (2019) study an atomic congestion game with incomplete information in which each agent participates with some known probability p. They compute the PoA and PoS when the costs are affine.

[7] The terms Price of Anarchy and Price of Stability were introduced by Papadimitriou (2001) and Anshelevich et al. (2008), respectively.

Several papers consider asymptotic games. Fournier and Scarsini (2019) study a game where a finite number of retailers choose a location, given that their potential consumers are distributed on a network. They prove asymptotic bounds on the PoA and PoS. Similarly to this paper, they show that asymptotically, the PoS in their game is one. Cominetti et al. (2023) study the similarities between Wardrop and Nash equilibria in atomic congestion games and show that the PoA and PoS converge at the limit. We conjecture that the Price of ϵ-Anarchy and Price of ϵ-Stability converge in our setting as well.

3. Model and Non-Existence of Equilibria

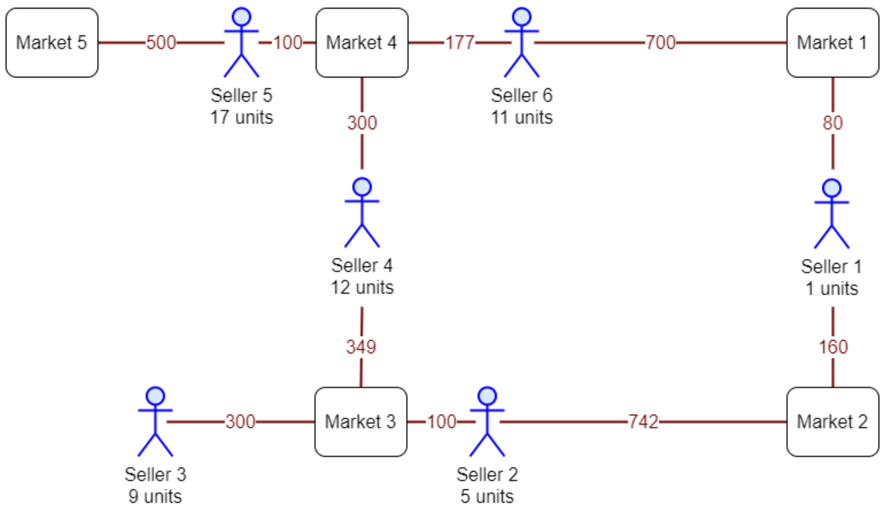
We first describe a general model that captures the seller dynamics described in the introduction. We show that there are instances of our model for which no equilibrium exists, and furthermore, no *approximate* equilibrium exists. We then show that under some mild assumptions, an equilibrium in the large always exists.

3.1. Model

Consider an economy for a perishable good consisting of n sellers and m markets. Sellers and markets are heterogeneous: seller i would like to sell q_i of the good and market j has size α_j. For each seller i and market j there is a cost of traveling from i to j, denoted c_{ij}. We assume that for all $i \in [n]$ and $j, j' \in [m], j \neq j'$, $c_{ij} \neq c_{ij'}$.[8] Each seller i chooses a market and sells their product there. Alternatively, the seller can choose to stay at home (not travel to any market). We denote the market that seller i chooses by $\boldsymbol{\sigma}_i$, where $\boldsymbol{\sigma}_i = \perp$ denotes that the seller stays at home. A *strategy profile* (sometimes simply referred to as *strategy*) is the set of all agents' choices: $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_n)$. For reasons detailed in the introduction, we only consider *pure* strategies—where each agent's strategy is deterministic. The price per unit volume of the good at market j in the strategy profile $\boldsymbol{\sigma}$ is $p_j(\boldsymbol{\sigma}) = \alpha_j v_j(\boldsymbol{\sigma})^{\beta}$, where $v_j(\boldsymbol{\sigma}) = \sum_{i:\boldsymbol{\sigma}_i = j} q_i$ and $\beta < 0$ is the price elasticity for the good (we assume the elasticity is constant across markets). The utility of seller i from strategy profile $\boldsymbol{\sigma}$, where $\boldsymbol{\sigma}_i = j$, is $u_i(\boldsymbol{\sigma}) = q_i p_j(\boldsymbol{\sigma}) - c_{ij}$. We denote such an economy by the tuple $(n, m, \beta, \boldsymbol{\alpha}, \mathbf{q}, \mathbf{C})$, where $\boldsymbol{\alpha}$ and $\mathbf{q}$ are vectors and $\mathbf{C}$ is a matrix. For each market $j \in [m]$, we denote the set of agents that are closer to market j than to any other market by $\textsc{Closest}(j)$. That is, $\textsc{Closest}(j) = \{i \in [n] : \forall j' \neq j, c_{ij} < c_{ij'}\}$. For strategy $\boldsymbol{\sigma}$, we define $(\boldsymbol{\sigma}_{-i}, \boldsymbol{\sigma}'_i)$ to be the strategy where all agents except i have the same strategy as they do in $\boldsymbol{\sigma}$ and i's strategy is $\boldsymbol{\sigma}'_i$.

A strategy is an *equilibrium* if no agent can profitably deviate (increase their utility by deviating unilaterally). Additionally, we define what it means for a strategy to be an *approximate* equilibrium:

[8] This assumption is made for simplicity and clarity in the proofs. It is straightforward to modify the results to remove this assumption.



For all $j \in \{1,\ldots,5\}$, $\alpha_j = 1000$ and $\beta = -1$.

Figure 2: Example where an equilibrium does not exist.

Definition 1. A pure strategy $\boldsymbol{\sigma}$ is an $\epsilon-$*equilibrium* if no agent can increase their utility by more than ϵ by unilaterally deviating. Formally, for any $i \in [n]$ and any strategy $\boldsymbol{\sigma}'_i$,

$$u_i(\boldsymbol{\sigma}_{-i}, \boldsymbol{\sigma}'_i) - u_i(\boldsymbol{\sigma}) \leq \epsilon.$$

3.2. Non-Existence of Approximate Equilibria

We show that for any $\epsilon \geq 0$, there exist instances of our model that do not admit an ϵ-equilibrium.

Theorem 1. *For any $\epsilon \geq 0$, there is an economy for which there does not exist an ϵ-equilibrium.*

Proof Outline. We describe an instance that does not admit an 0.1-equilibrium (Example 1 below), and show how to adapt it for any $\epsilon > 0$, to guarantee that there is no ϵ-equilibrium. We show that Example 1 does not admit a 0.1-equilibrium using exhaustive enumeration. To extend the result to arbitrary values of ϵ, we observe that the agents' utilities scale linearly with the market size and distances to the market. Hence, multiplying these values in Example 1 by any constant $\ell \geq 1$ results in an instance for which there is no $\ell\epsilon$-equilibrium. See Appendix B.1 for the complete proof. □

Example 1. There are six sellers and five markets, $\beta = -1$, and $\alpha_j = 1000$ for all markets $j \in \{1,\ldots,5\}$. The sellers' volumes and locations are given in Figure 2, where lines represent roads, and the distances of the sellers from the markets are noted on the road segments.

Example 1 is inspired by an example of Milchtaich (2009). Our example holds for functions of the form αv^{β}, whereas Milchtaich's function was only described by its value on a set of inputs (hence does not conform to our model). In addition, our example holds for distances that obey the triangle inequality, and requires fewer agents (six, as opposed to Milchtaich's eight).

It is straightforward to see that for any $n \geq 6$ and $m \geq 5$, one can construct an economy that does not admit an $\epsilon-$equilibrium: let 5 markets and 6 agents be as in Example 1, while the remaining markets and agents can be located arbitrarily, as long as the agents are sufficiently far from the markets of the example (so that any of the remaining $n-6$ agents could never attain a non-negative utility from going to Markets 1-5).

4. Equilibrium in the Large

The impossibility result of Theorem 1 relies on the fact that the economy in Example 1 is dispersed—each agent can only travel to a small subset of the markets without incurring negative utility, and each market serves only a small number of agents. Nevertheless, we show that under two mild assumptions on the market concentration, an $\epsilon-$equilibrium is guaranteed to exist when the number of agents is sufficiently large. We require the following definitions.

DEFINITION 2. An economy $E = (n, m, \beta, \boldsymbol{\alpha}, \mathbf{q}, \mathbf{C})$ is $\delta-$*concentrated* if for any market $j \in [m]$, at least a δ-fraction of the sellers are closer to market j than any other market. Formally, $\forall j \in [m], |\textsc{Closest}\,(j)\,| \geq \delta n$.

Let the NEAREST strategy be the strategy $\boldsymbol{\sigma}$ where $\forall i \in [n], \boldsymbol{\sigma}_i = j : i \in \textsc{Closest}\,(j)$.

DEFINITION 3. An economy $E = (n, m, \beta, \boldsymbol{\alpha}, \mathbf{q}, \mathbf{C})$ is *viable* if $\forall i \in [n], u_i(\textsc{Nearest}) > 0$.

At a high level, an economy is δ-concentrated if every market is located such that at least a δ-fraction of the agents are closer to it than to any other market; an economy is viable if agents derive a positive utility from participating in their closest markets. We will show that an ϵ-equilibrium always exists in any sufficiently large (with respect to the number of sellers) economy that is both δ-concentrated and viable. We note that we allow δ to be an arbitrarily small constant; we only require that for any market k, the number of agents in $\textsc{Closest}\,(k)$ increases with the total number of agents in the economy. Viability, however, is quite a strong requirement, and we note that while it does not hold for the markets that we study in Section 6, the total volume of the farmers that would not obtain a positive utility if all farmers travel to their nearest market is extremely small. We believe that these farmers would use an aggregator regardless; we remark upon this in more detail in Section 6. The proof that there always exists an ϵ-equilibrium is constructive: we describe an algorithm (denoted Algorithm 1) that always outputs such an ϵ-equilibrium.

4.1. Algorithm Description

We give an informal description of Algorithm 1 below; the detailed pseudocode for Algorithm 1 appears in Appendix A. We note that the constants used in the description below ($v^*, v^\dagger$ and μ) depend on the problem parameters; their more precise and formal definitions appear in the pseudocode.

1. **Initialization:** Initialize the agent locations using the NEAREST strategy.
2. **Iterative Adjustment:** The core of the algorithm is an iterative process that continues until an ϵ-equilibrium is reached. In each iteration:
   (a) If any agent has a negative utility, set their strategy to $\perp$.
   (b) Check whether one of the following conditions holds:
      i. All markets have a volume equal to or greater than a predefined threshold v^* and there is an agent i and a market k such that $u_i(\boldsymbol{\sigma}_{-i}, k) > u_i(\boldsymbol{\sigma}) + \epsilon$.
      ii. There is an agent i whose current strategy is $\perp$, and the volume of their closest market, denoted k, is below some threshold $v^\dagger$.
      iii. There is an agent i for whom the volume at their current market (σ_i) is more than μ times the volume at their closest market, k.
   (c) If any of the conditions hold, move the agent denoted i to the market denoted k in the condition. Otherwise, the current strategy is an ϵ-equilibrium.

4.2. Main Result

Our main theorem, given formally below, states that Algorithm 1 computes an ϵ-equilibrium in a sufficiently large viable δ-concentrated economy. The result is in fact, slightly stronger, as the strategy profile also satisfies individual rationality:

DEFINITION 4. A pure strategy $\boldsymbol{\sigma}$ is *individually rational* (IR) if the utility of each agent is non-negative: $\forall i \in [n], u_i(\boldsymbol{\sigma}) \geq 0$.

Note that a strategy being an ϵ-equilibrium does not automatically guarantee that it is individually rational.

THEOREM 2. *For any $m > 0, \beta \in (-1, 0), 0 < \underline{\alpha} \leq \overline{\alpha}, 0 < \underline{q} \leq \overline{q}$, let $\mathcal{E}_{m,\beta,\underline{\alpha},\overline{\alpha},\underline{q},\overline{q}}$ denote the set of economies $E = (n, m, \beta, \boldsymbol{\alpha}, \mathbf{q}, \mathbf{C})$ such that $\underline{\alpha} \leq \alpha_j \leq \overline{\alpha}$ for each $j \in [m]$ and $\underline{q} \leq q_i \leq \overline{q}$ for each $i \in [n]$.*

For any $m, \beta, \underline{\alpha}, \overline{\alpha}, \underline{q}, \overline{q}$ as above, and any $\epsilon, \delta > 0$, there exists some $n^* > 0$ such that for every $n \geq n^*$, every viable δ-concentrated economy $E \in \mathcal{E}_{m,\beta,\underline{\alpha},\overline{\alpha},\underline{q},\overline{q}}$ with n agents admits an individually rational $\epsilon-$equilibrium. Furthermore, Algorithm 1 computes such an ϵ-equilibrium.

Proof Outline. The proof consists of two main parts. In the first part, we show that if the strategy is not an ϵ-equilibrium, one of the conditions of Algorithm 1 (denoted i., ii. and iii. under 2(b) in the algorithm description above) must hold. That is, either all market volumes are sufficiently large, and there is some agent whose utility can increase by more than ϵ by deviating, or there is an agent that is not at their closest market, and the market volumes obey certain conditions.

This part of the proof shows that the algorithm does not encounter any errors during its execution. In the second part, we define a potential function, and show that at every iteration of the algorithm, the value of the potential function increases. We note that we do not claim that the deviation is profitable for the deviating agent, only that it causes the potential to increase. As there is a finite number of strategies, the algorithm must terminate. By the algorithm definition, it terminates only if the strategy profile is an ϵ-equilibrium. The complete proof appears in Appendix B.2. □

4.3. Price of ϵ-Stability

We now show that when the economy is δ-concentrated, viable and the number of agents is sufficiently large, requiring that the allocation is an ϵ-equilibrium comes with no loss of welfare.

Definition 5. The *price of ϵ-stability* is the ratio of the utility of the best ϵ-equilibrium relative to the optimal solution.

To show that the price of ϵ-stability is 1, we show that every optimal allocation is necessarily an ϵ-equilibrium. Formally, the result is the following.

Theorem 3. *For any $m > 0, \beta \in (-1, 0), 0 < \underline{\alpha} \leq \overline{\alpha}, 0 < \underline{q} \leq \overline{q}$, let $\mathcal{E}_{m,\beta,\underline{\alpha},\overline{\alpha},\underline{q},\overline{q}}$ denote the set of economies $E = (n, m, \beta, \boldsymbol{\alpha}, \mathbf{q}, \mathbf{C})$ such that $\underline{\alpha} \leq \alpha_j \leq \overline{\alpha}$ for each $j \in [m]$ and $\underline{q} \leq q_i \leq \overline{q}$ for each $i \in [n]$.*

For any $m, \beta, \underline{\alpha}, \overline{\alpha}, \underline{q}, \overline{q}$ as above, and any $\epsilon, \delta > 0$, there exists some $n^* > 0$ such that for every $n \geq n^*$, the price of ϵ-stability of every viable δ-concentrated economy $E \in \mathcal{E}_{m,\beta,\underline{\alpha},\overline{\alpha},\underline{q},\overline{q}}$ with n agents is 1.

The proof appears in Appendix B.3. We note that this result does not guarantee that Algorithm 1 computes an optimal, or even an almost optimal allocation. We conjecture that the ϵ-equilibrium generated by Algorithm 1 extracts almost all of the possible welfare, and this is supported by simulations (Section 7 and Appendix D), in which we show that Algorithm 1 typically extracts almost all of the welfare, even in small economies. We remark upon this further in the conclusion.

5. Understanding and Comparing Seller Strategies

To understand the impact of ex-post equilibrium recommendations on sellers and markets, we first needed to identify the strategies currently used by sellers. Ashkrof et al. (2020) and Camerer et al. (1997) document strategies used by ridesharing drivers; we contribute by examining how farmers in India make similar decisions about where to sell their goods. To do so, we surveyed 345 onion farmers and 157 potato farmers in Maharashtra, India. Onion and potato farmers were chosen as they are important crops in India. Farmers were asked about important factors that influence the market they select, farming experience, farm size, sales channels, and sale frequency. Appendix C contains a table for each question in the survey and either the percent of farmers responding in a given way (for categorical response questions) or the mean and standard deviation of responses (for numerical response questions such as the quantity sold in a market in a single trip). We note a few interesting themes from the survey results.

Farmers take many factors into account when making their decision. Transportation costs are important to the farmers' decision, with 42% of all farmers (40.6% of onion farmers, 45.2% of potato farmers) saying that it plays a role in their decision. Other key factors are the previous day's price 38.8% (39.4%, 37.6%), market size 33.9% (31.9%, 38.2%), their relationship with the buyer 33.9% (33.9%, 33.8%), the amount that they produced 29.7% (28.7%, 31.8%), and even family tradition 12.9% (13.9%, 10.8%). In particular, path dependence is still strong, with 43.2% (43.8% 42%) saying that they consider a relationship with a buyer or tradition as an important factor.

A majority of farmers 54.4% (53.6%, 56.1%) sell their produce at most once a month, which underscores the importance of this decision, as this is typically their only source of income. Given the complexity and importance of this decision, farmers typically consult with their neighbors and friends: 86.1% of farmers (89%, 79.6%) consult at least one other farmer to determine which market to sell their produce at, and 69.7% (72.2%, 64.3%) consult with at least two other farmers. It is therefore unsurprising that neighboring farmers usually end up going to the same market; the vast majority 89% (91%, 84.7%) say that they generally sell their produce at the same market as farmers close to them.

We also note that despite the digitization of markets generally and the introduction of electronic marketplace options like e-Choupal, only 3.2% (2%, 5.7%) typically sell their produce in an e-marketplace. Furthermore, 22.5% (20%, 28%) sell their produce to (or through) an aggregator.

In addition to the survey, we consider strategies that were utilized in previous literature (e.g., Parker et al. 2016) as well as strategies discovered from conversations with both employees at RML and farmers. Through this process, we identify four baseline strategies, which we detail below.

5.1. Seller Strategies

In the proposed recommendation strategy (Rec), sellers follow a recommended strategy, where the recommendation profile is intended to elicit a pure (ex-post) equilibrium. We compare it to two non-adaptive baseline strategies (Nearest and Greedy) where farmers do not react to the previous market conditions and two adaptive (Best Response and Iterative Best Response) strategies where farmers take into account previous market decisions.

- Nearest: In the Nearest strategy, sellers travel to their nearest market. Despite its simplicity and lack of consideration for market sizes, this strategy is common, as indicated by our survey; 33.5% of the farmers surveyed typically use it. This strategy was also used as a benchmark in Parker et al. (2016) as its use appears to have been widespread in the Indian agricultural market in the years immediately after the introduction of RML.
- Greedy: In the Greedy strategy, sellers maximize their welfare, based on the market size (α) and the travel costs, assuming that all markets have an identical volume before accounting for the seller's contribution. This assumption is made so that only the market size (which is known at the time of the decision) will play a role in the decision, and not the the market volume, which is speculative. For simulation purposes, we took this volume to be 0; we note that i) our results do not change qualitatively when we use different base volumes and ii) when the base volume is very large, Greedy reduces to Nearest. We include this strategy as it has an intermediate level of sophistication. On the one hand, it takes into account both the market size (which in turn affects the expected price) and the travel costs; on the other hand, it is not strategic, as it does not take into account the other players' actions.
- Best Response (BR): In the BR strategy, sellers simultaneously best-respond once to some specific market outcome. Here, sellers know the values of α, the travel costs, and the market volumes (and hence prices) that resulted from an initial allocation of sellers to markets, and best-respond to this setting. In the simulation, sellers are best responding to the volumes resulting from the Greedy strategy. We obtain similar results if the sellers best respond to different volumes (for example, those generated by Nearest). This strategy takes into account most of the factors that the farmers surveyed reported that they take into considerations: the transportation cost, market size, amount produced and yesterday's price.
- Iterative Best Response (IBR): In the IBR strategy, sellers simultaneously best-respond repeatedly. This strategy captures the sellers' welfare at a steady state that is reached by repeatedly playing a best-response. This setting is also initialized using Greedy (again, we get similar results with different initializations). We stop the simulation game when either: i) we reach an equilibrium (i.e., each seller's best-response is their current choice), ii) we enter a short cycle (specifically, the same strategy profile had already been observed within the past 10 iterations), or iii) we have reached the maximal allowable number of iterations. If we reach a cycle, we chose a strategy profile uniformly at random from the profiles on the cycle. We chose a cycle length of 10 because we found that in virtually all settings, cycles tended to be at most 10 profiles long.

We note that in all four baseline strategies, similar sellers (in terms of location and production amount) will tend to travel to the same market, hence they are all consistent with the farmers reporting that they typically travel to the same market as their neighbors. In all of the simulation results presented in this paper, all sellers play the same strategy. Results are similar when sellers play different strategies. For brevity, we omit those results.

5.2. Key Metrics

We use six key metrics to compare REC to the other strategies. The first five use a direct comparison between REC and a baseline strategy $S \in \{\text{NEAREST}, \text{GREEDY}, \text{BR}, \text{IBR}\}$. The final metric is non-comparative/strategy-specific; nevertheless it is useful in order to obtain more insights about the advantages of the REC strategy.

1. **Welfare Ratio**: Let $\text{SW}(S)$ denote the total seller welfare (the sum of the sellers' utilities) when sellers use strategy S. The welfare ratio with respect of REC to strategy $S \in \{\text{NEAREST}, \text{GREEDY}, \text{BR}, \text{IBR}\}$ is defined as $\text{SWR}(S) = \frac{\text{SW(REC)}}{\text{SW}(S)}$.
2. **Percent Better under** REC: The percent of sellers whose utility is strictly better under REC than under strategy S.
3. **Conditional Improvement Percent:** The average percentage increase in utility of sellers under REC, conditioned on improvement. Formally, if the utility of seller i when sellers use strategy S is u_i^S, the seller's improvement is $\frac{u_i^{\text{REC}} - u_i^S}{u_i^{\text{REC}}}$. We use this metric both as a general one and to gauge the fairness of the REC strategy. For the latter, we consider individual improvement. We note that the unconditional improvement is measured by $\text{SWR}(S)$.
4. **Ratio of Price CV:** Let $P(S)$ denote the vector of market prices when sellers use strategy $\boldsymbol{\sigma}$. The coefficient of variation of market prices when sellers use strategy S is $\text{CV}(S) = \sigma(P(S))/P\bar{(}S)$, where $\sigma(P(S))$ and $P\bar{(}S)$ are the standard deviation and mean of $P(S)$, respectively. The Ratio of Price CV with respect to strategy $\boldsymbol{\sigma}$ is $\text{CVR}(S) = \frac{\text{CV(REC)}}{\text{CV}(S)}$. We note that markets that have no volume do not have a price and, therefore, are not included in the calculation of the coefficient of variation. The Ratio of Price CV is an important metric since volume moving from low price markets to high price markets translates directly into reductions in geographic price dispersion resulting in net (seller plus buyer) welfare gains (Jensen 2007). An alternative would be to measure buyer welfare, but that would rely on additional assumptions about the behavior of those participants.
5. **HHI Ratio:** To measure market concentration, we use the Herfindahl–Hirschman Index: $\text{HHI}(S) = \sum_{j=1}^{m} \frac{v_j(S)^2}{\left(\sum_{j=1}^{m} v_j(S)\right)^2}$, where $v_j(S)$ denotes the total volume at market j when sellers use strategy S. The Ratio of Market Volume Concentration with respect to strategy $\boldsymbol{\sigma}$ is $\text{HHIR}(S) = \frac{\text{HHI(REC)}}{\text{HHI}(S)}$.

6. **Percent Regretful:** The percentage of sellers that, upon the revelation of the realized volumes, conclude that they would have attained a higher utility if they had gone to a different market.

For SWR(S), CVR(S) and HHIR(S), we use the median across simulations, and not, for example, the mean, in order to minimize the impact of outliers on the metrics.

6. Numerical Analysis Using Data From India's Agriculture Markets

We conduct a numerical analysis to compare Rec with the baseline strategies, using price and subscription data from RML to generate realistic problems. The price data includes the price and volume transacted for 367 crops in 1,318 markets throughout 16 states. The data starts on January 1, 2008 and ends on January 31, 2011 for a total of 1,848,349 crop-market-days. We estimate the price elasticities using individual regressions for each crop of the form:

$$\ln(price_{mt}) = \alpha_m + \tau_t + \beta \times \ln(volume_{mt}) + \epsilon_{mt}, \tag{1}$$

where $\ln(price_{mt})$ ($\ln(volume_{mt})$) is the log of the price (volume) for market m on day t, α_m are market fixed effects that capture time-invariant differences across markets, τ_t are year fixed effects that capture trends in market prices and volumes over time, and ϵ_{mt} is the idiosyncratic shock. The parameters of interest are β and α_m, which will allow us to set prices after observing market volumes based on farmers' market selection decisions. Table 2 shows the estimated parameters used for the crops that we selected. It is important to note that our objective is to estimate reasonable parameters for the numerical analysis; we do not claim that the estimated α_m and β are perfectly accurate. We further discuss and expand on the data, parameter estimation process, and our crop selection criteria in Appendix E.

We perform 500 simulations and analyze the median value of the key metrics across simulation runs. For each simulation run, seller and market locations are set to the locations we observe in the RML data (hence the number of farmers and markets are constant across simulation runs), and the set of market sizes $\{\alpha_j\}$ is constant across simulation runs for the same crop, but these are randomly shuffled among the known market locations across simulation runs. Seller volume is modeled as a normal distribution, truncated at zero. We estimate transportation costs using the crow's flight distance between farmers and markets, 70 Rupees per liter of fuel and 6.3 liters per 100 kilometers (Global Fuel Economy Initiative 2010).

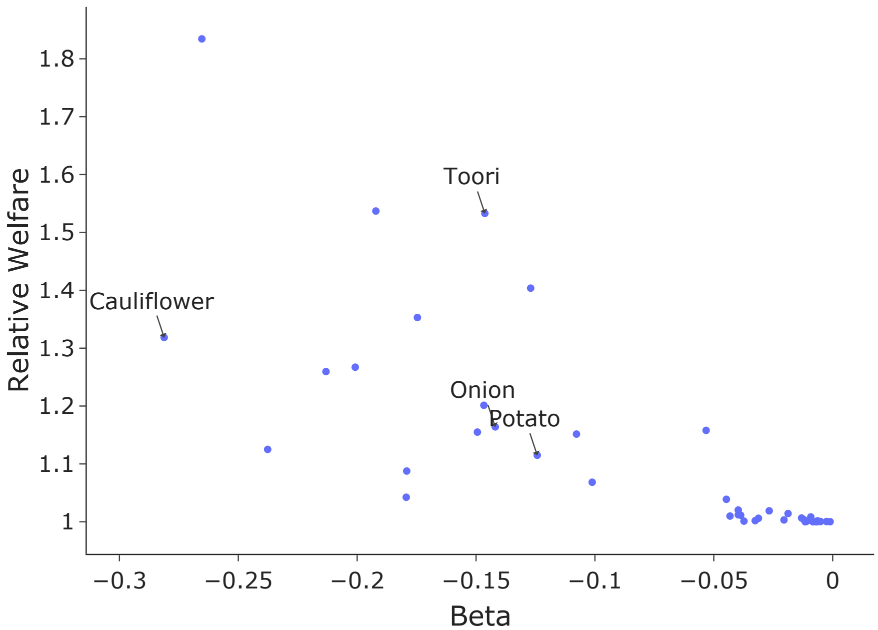


Figure 3: Welfare Ratio comparing Rec to Greedy for the crops in our dataset.

	No. of Farmers	No. of Markets	β	α Min	α Median	α Max	Farmer Volume Mean	Farmer Volume Std. Dev.
Cauliflower	2762	127	-0.281	37.206	2364.421	11377.228	180.222	621.365
Onion	2912	187	-0.142	827.109	3121.573	7261.557	432.578	1169.078
Potato	22556	126	-0.124	139.546	764.513	2588.425	254.395	909.403
Toori	10205	31	-0.146	223.110	432.455	694.533	5.325	8.395

Table 2: Scenario parameters estimated from RML data.

We analyze a total of 48 crops. Figure 3 shows the median Welfare Ratio for Rec compared to Greedy for all crops in our estimation sample. As expected, there is a clear relationship between β and the Welfare Ratio. Rather than showing full results for all 48 crops, we focus on the four highlighted crops: (1) Potato and (2) Onion because they are important crops in this context and farmers' responses to the survey about these crops were used as motivation for the comparison farmer selling strategies, (3) Cauliflower because it has the lowest (most extreme) value of β, and (4) Toori (a type of gourd) because it has highly concentrated farmers.

Table 3 summarizes our results: compared to all other strategies, Rec leads to higher welfare, lower price variability, and generally lower market concentration (with the exception of HHIR(Nearest) which achieves more dispersed volumes for all crops). An overwhelming majority of farmers regret their market choice for any strategy other than Rec. In hindsight, these farmers would have preferred to go to another market or possibly not go to any market. Under Rec, by design, no farmer regrets their decision.

		Nearest	Greedy	BR	IBR
Cauliflower	Welfare Ratio	2.744	1.318	2.085	1.318
	Percent Better Under Rec	89.500	99.976	99.877	99.975
	Improvement Percent	64.682	41.735	63.077	42.146
	Ratio of Price CV	0.131	0.129	0.133	0.132
	HHI Ratio	2.154	0.173	0.173	0.173
	Percent Regretful	99.777	99.976	99.977	99.977
Onion	Welfare Ratio	1.614	1.164	1.476	1.164
	Percent Better Under Rec	90.639	99.876	99.567	99.812
	Improvement Percent	37.772	14.828	31.837	18.112
	Ratio of Price CV	0.025	0.018	0.022	0.019
	HHI Ratio	3.608	0.139	0.139	0.139
	Percent Regretful	99.996	99.929	99.941	99.957
Potato	Welfare Ratio	2.677	1.115	1.271	1.115
	Percent Better Under Rec	96.547	99.501	99.562	99.521
	Improvement Percent	62.778	14.270	24.882	19.212
	Ratio of Price CV	0.081	0.147	0.150	0.156
	HHI Ratio	2.872	0.237	0.237	0.237
	Percent Regretful	99.976	99.904	99.915	99.927
Toori	Welfare Ratio	1.137	1.533	1.944	1.602
	Percent Better Under Rec	71.321	86.344	95.246	90.848
	Improvement Percent	18.325	15.148	97.687	61.222
	Ratio of Price CV	0.283	0.374	0.251	0.246
	HHI Ratio	1.063	0.106	0.108	0.163
	Percent Regretful	94.056	96.807	97.042	97.540

Table 3: Summary of key metrics.

For the four crops we focus on, there are no simulations in which any of the baseline strategies achieve a higher total seller welfare than Rec. We note that we did observe some instances where BR or IBR achieve higher welfare than Rec in other crops. These occur almost exclusively in one crop—a type of pomegranate—where farmers are quite small. In reality, these farmers likely sell their produce to an aggregator and not to the open market.

The welfare gains are also shared among a large portion of the farmers. At the low end, 71% of Toori farmers are better off under Rec than Nearest. This is actually the second lowest value we see across all crops, with all other crops seeing at least 80% of farmers sharing in the welfare gains. This inclusive growth is not just seen for crops in which there are large welfare gains from using Rec. In fact, the total welfare gain from following Rec as opposed to Greedy for Banana farmers is only 0.1%, but over 99% of farmers have a higher welfare.

We examined to what extent the *viability* assumption of Section 4 holds in reality. We found that some crops were viable (i.e., all farmers had a positive utility under NEAREST), but most were not. In all of the crops, however, the volume of farmers that obtained a negative utility was very small. For example, the total volume of all Toori farmers with negative utility was just 1.3% of the total farmer volume. For Cauliflower, it was 0.16%, and for Potato and Onion, it is less than 0.001%.

Finally, we provide a quantification of farmer welfare to Rupees. Rather than calculating the ratio of farmer welfare, we calculate the change in farmer welfare. The median change in total farmer welfare across simulation runs varies greatly across the crops with Toori farmers earning about 438 additional Rupees and Potato, Cauliflower and Onion farmers earning around 29,752, 36,143, and 156,367 additional Rupees each, respectively. We note that this is simply an estimate of farmer welfare and does not include other welfare gains associated with reductions in geographic price dispersion.

7. Numerical Analysis using Synthetic Data

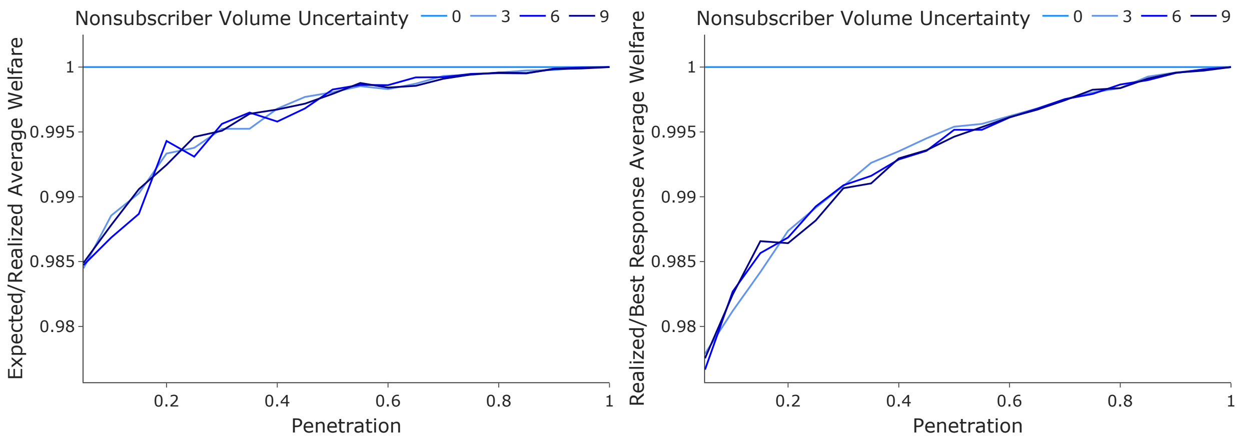
We now compare REC with the baseline strategies, using synthetic data. For each simulation run, seller and market locations are selected at random from the unit square. All markets have the same elasticity β, and have uniformly distributed market sizes (α_j). Seller volumes are also uniformly distributed and each seller's travel cost to a market is their Euclidean distance multiplied by a common cost multiplier c. The baseline parameters and distributions used in the analysis are summarized in Table 4. In Subsection 7.1, we examine the robustness of the service to incomplete market penetration: where not all sellers subscribe to the service. In Subsection 7.2, we analyze the system's fairness, with respect to whether it advantages large sellers over small ones. For ease of exhibition, the analysis of is done with respect to the GREEDY strategy; we obtain similar results for the other strategies. For each configuration, we perform 500 simulations and analyze the median value of the key metrics across simulation runs. More details about the baseline results and additional parameter sensitivity analyses are available in Appendix D.

7.1. Sensitivity to Penetration and Volume Uncertainty

Up to this point we have assumed that the service has 100% penetration. As such, the service provider knows how much volume each market will receive and the price each seller will receive at the market. In reality, no service can have complete market penetration. Without full penetration, the service cannot perfectly predict market volumes and prices meaning sellers will not receive the prices they were expecting. This presents two issues: 1) sellers may have an expected welfare that is different from the one they receive, and 2) sellers may observe prices at the end of the day

Parameter	Value/Distribution	Heterogeneity
Price elasticity β	-0.1	§D.2: -0.05 to -0.30
Market size α_j	Uniform 10 to 50	-
Cost multiplier	20	-
Farmer location	Uniform on unit square	-
Market location	Uniform on unit square	-
Number of markets m	15	-
Number of sellers n	1000	-
Seller Volume v_i	Uniform 0.5 to 1.5	§D.2: max from 0.6 to 6 in increments of 0.2
Service penetration	100%	§7.1: 5% to 100% in increments of 5%
Non-subscribed seller volume uncertainty	0	§7.1: 0, 3, 6, 9

Table 4: Scenario Parameters for Synthetic data.



(a) Realized vs. Expected Welfare **(b)** Best Response vs. Realized Welfare

Figure 4: Average Welfare as a function of service penetration and volume uncertainty

and regret their market choice. We evaluate the impact of partial penetration on both of these issues by varying the number of sellers who have subscribed to the service, and the precision with which the service can estimate seller volume for non-subscribed sellers. To do this, we assume that non-subscribed sellers choose their nearest market and have some (unknown, but estimable to the service) volume. If there is no volume uncertainty, the service knows the seller's size and strategy, hence the resulting prices will be precisely as predicted. When there is volume uncertainty then seller sizes are only observed after they arrive at the market. As such, prices may be higher or lower and sellers may be disappointed by the prices they receive or regret not visiting a different market which realized higher prices than expected. We implement the volume uncertainty by allocating unsubscribed sellers to their nearest market and calculating the expected volume from unsubscribed farmers as the sum of volume from unsubscribed sellers allocated to that market. We draw the market's realized volume from unsubscribed sellers from a normal distribution that is truncated at zero (to preserve non-negative volume constraints) and at two times the expected volume from unsubscribed farmers (to ensure the service's expectation is correct on average). The service knows the expected volume from unsubscribed farmers and the standard deviation is a multiple of the market's expected volume from unsubscribed farmers ranging from zero (no uncertainty) to nine times the expected size (high uncertainty).

Figure 4a shows that sellers receive nearly all of their expected welfare. There is a small dropoff for lower levels of penetration, but the total welfare loss is less than 1% even at 25% penetration. Similarly, Figure 4b shows that sellers could only gain a little welfare by best responding indicating that they will have little regret at the end of the day.

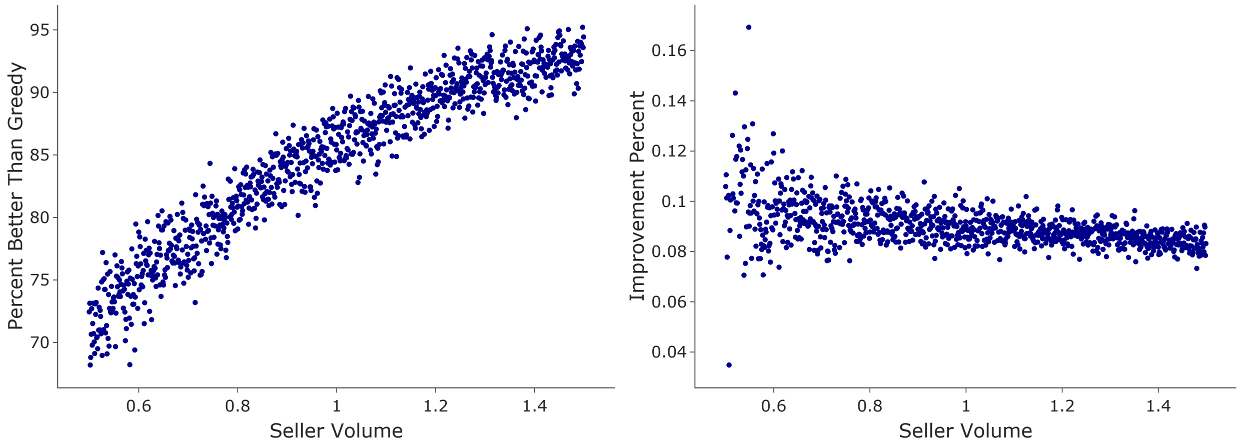
We note that while we model the uncertainty in terms of the seller's size, this actually captures a wider range of uncertainty, including seller behavior. For example, some sellers may not use Nearest, but instead Greedy or a randomized strategy or another strategy we have not considered. That would lead to over or under estimates of the volume at the markets and accompanying mis-estimations of the price. As such, these results demonstrate the general robustness of Rec.

7.2. Fairness

We investigate whether the Rec strategy disproportionately benefits large sellers (sellers with higher volumes) compared to small ones. This is indeed the case on average when comparing individual sellers' welfare under Greedy and Rec, as Figure 5a shows. It appears to be largely a benefit of scale as both small and large sellers pay the same travel cost per unit distance to the market; Figure 5b shows that sellers whose welfare is higher under Rec have fairly consistent improvement percentages. Nevertheless, to promote equitable growth opportunities, strategies may be needed to subsidize small sellers or some of the benefits to large sellers could be used to facilitate shared transportation costs for small sellers through, e.g., the creation of local farming cooperatives. The high travel costs can lead to a situation in which aggregators provide a Pareto-improving service. For example, consider ten small farmers that jointly pay nine times the travel cost to an aggregator who gives the farmers exactly the price they would have received in the market. Each small farmer now pays less than their original travel cost but has the same revenue, leading to a net welfare improvement.

8. Discussion

New business models are leveraging information technology to make markets more efficient in developing economies. Harnessing the benefits of these new models may require more sophisticated strategies for ensuring that market participants can more easily coordinate in an incentive-aligned way. In this paper, we propose a framework for recommending markets to sellers, such that the resulting allocation is an ex-post equilibrium. The recommended equilibria typically lead to higher



(a) Percent Better Under Rec than Greedy

(b) Improvement Percent

Figure 5: Small sellers are less likely to be strictly better in Rec than Greedy as seen by the increasing slope in (a), but conditional on improving are similarly or slightly better off than comparable large sellers as evidenced by the decreasing slope in (b).

farmer welfare as well as lower geographic price dispersion and market concentration when compared to market selection strategies identified by surveying onion and potato farmers in Maharashtra, India. The numerical analysis using data from India's agricultural markets demonstrates that the framework is feasible and effective for realistic scenarios. We also show that improvements are retained, and even amplified, as we alter some of the key parameters. From this, we infer that the algorithm will work even with imprecisely estimated parameters or for crops with a broad range of price elasticity and farmer volume ranges.

Our work has managerial and policy implications. We show that adding a market recommendation service on top of price information services can benefit both individual farmers and the economy in general. Importantly, farmers will not regret following the recommendations. Policy makers should note that offering market information can be helpful, but providing coordination assistance can lead to even larger welfare gains. Because the recommendations also cause a reduction in geographic price dispersion—which translates into net welfare gains—and market concentration, the welfare improvements are not limited to only subscribers or even only sellers. Hence, policy makers may wish to encourage widespread use of market recommendation services in an effort to reduce poverty around the country. Policy makers may also wish to encourage aggregation services for crops where the cost of transportation is high relative to an individual farmers' volumes. We note that this solution could be market-based (for-profit aggregators) or not (non-profits, farmer-run cooperatives, or governmental drop off stations). An important aspect of the framework is that the recommendations do not need to replace the information services currently provided, but can be used to augment them. In fact, supplying the farmers with market information can offer transparency and help convince them that following the recommendation is indeed in their self interest.

We prove that the Price of ϵ-Stability is asymptotically one, but we have not been able to prove bounds on the Price of ϵ-Anarchy (which would be defined analogously to the Price of ϵ-Stability). Proving a bound on the Price of ϵ-Anarchy seems much more difficult, and many techniques that already exist do not appear to work. For instance, the smoothness framework of Roughgarden (2015), which has been successfully applied in many related settings (e.g., Cominetti et al. 2019, Feldman et al. 2016) does not seem to apply to our setting due to the travel costs, although we cannot rule out some clever application that relies on the triangle inequality. We conjecture that the price of ϵ-Anarchy is asymptotically one as well; in all of our simulations, any equilibrium computed captures almost all of the welfare, and the ratio increases as the market size grows (see Appendix D.3).

8.1. Notes on Implementation

Despite the theoretical impossibility results, we found that Algorithm 1 always found an exact equilibrium (i.e., setting $\epsilon = 0$.) We also implemented a simpler code: a best-response algorithm. We ran over one billion simulations, and this algorithm was always able to efficiently compute an (exact) equilibrium. We also experimented with several variations on best-response, when more than one seller could improve their utility by deviating. For example, we compared the equilibria achieved by prioritizing sellers with a larger quantity of goods versus those with a smaller quantity. We did not find any significant differences between the equilibria in any of the key metrics. In an effort to facilitate deployment of the framework, the numerical analysis code is available at: `https://osf.io/7at4n/?view_only=eb63433d94454162a989821c6847fc50`. It will be hosted on a public GitHub after publication.

Despite the ability to efficiently compute an exact equilibrium, there are still some challenges to implementation. First is obtaining an accurate estimate of the behavior of sellers that are not subscribed to the services as well as subscribed sellers who do not adhere to the recommendation for any reason. Specifically, one would need to estimate the base volume at each market on each day; i.e., the volume that is generated by unsubscribed and non-compliant farmers. We have shown that the recommendation is robust to relatively imprecise estimates of base volumes. Achieving this level of precision does not seem insurmountable by combining historical price and volume data with a survey of subscribed farmers.

Second, the values of α (market size) and, to a lesser extent, β, may change over time. The recommender must have a reasonable estimate of these parameters for a given market on a given day in order to make useful recommendations. We did not find substantial variation in these metrics from day to day in the data. However, we do observe variability on longer time scales, implying that it may be sufficient to use the previous day's value as an estimate for the following day's, although more rigorous study would be required to verify this. Related to this point, in this paper we only considered a one-shot game in which farmers are only concerned with their welfare for the day. In reality, some crops can be held for several days before going to the market. In such situations, farmers who did not go to the market today may go to the market tomorrow. Maximizing farmer welfare over a longer time-frame involves recommending a (market, day) pair to each farmer, and the algorithm can easily be modified to incorporate this change. However, making market-day recommendations increases the dimensionality of the problem—there would be 'number of days' times 'number of markets' options for each of the farmers, as opposed to just the number of markets. In order to extend the recommendations this way, it is important to understand the stability of the values of α and β over time.

Third, we assume that farmers do not strategically coordinate in any way. While our survey shows that farmers do discuss their market decisions with other farmers, we do not find strong evidence of strategic coordination. Nevertheless, farmers can coordinate in at least two ways. First, they can jointly determine their markets in an effort to maximize their welfare. To the extent that this is not global coordination, farmers will not be able to achieve the welfare gains observed in the framework. Second, farmers can attempt to reduce their selling costs by either aggregating volumes through a local cooperative or selling directly to an aggregator who takes the produce to a market. These aggregation options are particularly interesting, as these intermediaries are sometimes considered predatory and only take money from the farmers by using their superior market knowledge. Instead, they can be a useful selling channel that limits negative welfare situations for farmers. However, they do present a challenge for the current implementation of our algorithm. While aggregators can report their volume and be immediately included in the algorithm, there is currently no mechanism for recommending that farmers aggregate volume in some way such that market participation rates reach acceptable levels. Implementing this would increase the complexity of the problem as it would require each farmer to make the market/aggregation decision before the aggregator makes the market decision.

References

Abebe, Gumataw K, Jos Bijman, Annie Royer. 2016. Are middlemen facilitators or barriers to improve smallholders' welfare in rural economies? Empirical evidence from Ethiopia. *Journal of Rural Studies* **43** 203–213.

Anshelevich, E., A. Dasgupta, J. Kleinberg, É. Tardos, T. Wexler, T. Roughgarden. 2008. The price of stability for network design with fair cost allocation. *SIAM Journal on Computing* **38**(4) 1602–1623.

Ashkrof, Peyman, Goncalo Homem de Almeida Correia, Oded Cats, Bart van Arem. 2020. Understanding ride-sourcing drivers' behaviour and preferences: Insights from focus groups analysis. *Research in Transportation Business & Management* **37** 100516.

Aumann, Robert J. 1974. Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics* **1**(1) 67–96.

Azevedo, Eduardo M, Eric Budish. 2019. Strategy-proofness in the large. *The Review of Economic Studies* **86**(1) 81–116.

Camerer, Colin, Linda Babcock, George Loewenstein, Richard Thaler. 1997. Labor Supply of New York City Cabdrivers: One Day at a Time. *The Quarterly Journal of Economics* **112**(2) 407–441.

Candogan, Ozan, Asuman Ozdaglar, Pablo A. Parrilo. 2013. Dynamics in near-potential games. *Games and Economic Behavior* **82** 66–90.

Caragiannis, Ioannis, Angelo Fanelli, Nick Gravin, Alexander Skopalik. 2015. Approximate pure nash equilibria in weighted congestion games: Existence, efficient computation, and structure. *ACM Trans. Econ. Comput.* **3**(1).

Cason, Timothy N, Tridib Sharma. 2007. Recommended play and correlated equilibria: an experimental study. *Economic Theory* **33**(1) 11–27.

Chakrabarty, Deeparnab, Aranyak Mehta, Viswanath Nagarajan. 2005. Fairness and optimality in congestion games. John Riedl, Michael J. Kearns, Michael K. Reiter, eds., *Proceedings 6th ACM Conference on Electronic Commerce (EC-2005), Vancouver, BC, Canada, June 5-8, 2005*. ACM, 52–57.

Cominetti, Roberto, José R Correa, Nicolás E Stier-Moses. 2009. The impact of oligopolistic competition in networks. *Operations Research* **57**(6) 1421–1437.

Cominetti, Roberto, Matteo Quattropani, Marco Scarsini. 2022. The buck-passing game. *Mathematics of Operations Research* **47**(3) 1731–1756.

Cominetti, Roberto, Marco Scarsini, Marc Schröder, Nicolás Stier-Moses. 2023. Approximation and convergence of large atomic congestion games. *Mathematics of Operations Research* **48**(2) 784–811.

Cominetti, Roberto, Marco Scarsini, Marc Schröder, Nicolas E Stier-Moses. 2019. Price of anarchy in stochastic atomic congestion games with affine costs. *Proceedings of the 2019 ACM Conference on Economics and Computation*. 579–580.

Feldman, Michal, Nicole Immorlica, Brendan Lucier, Tim Roughgarden, Vasilis Syrgkanis. 2016. The price of anarchy in large games. *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*. 963–976.

Fotakis, Dimitris, Spyros Kontogiannis, Elias Koutsoupias, Marios Mavronicolas, Paul Spirakis. 2009. The structure and complexity of Nash equilibria for a selfish routing game. *Theoretical Computer Science* **410**(36) 3305 – 3326.

Fotakis, Dimitris, Spyros Kontogiannis, Paul Spirakis. 2005. Selfish unsplittable flows. *Theoretical Computer Science* **348**(2) 226 – 239.

Fournier, Gaëtan, Marco Scarsini. 2019. Location games on networks: Existence and efficiency of equilibria. *Mathematics of Operations Research* **44**(1) 212–235.

Frayer, Lauren. 2021. India's farmer protests: Why are they so angry? https://www.npr.org/sections/goatsandsoda/2021/03/02/971293844/indias-farmer-protests-why-are-they-so-angry. Last Accessed: 2024-07-30.

Gairing, Martin, Burkhard Monien, Karsten Tiemann. 2011. Routing (un-) splittable flow in games with player-specific affine latency functions. *ACM Trans. Algorithms* **7**(3) 31:1–31:31.

Global Fuel Economy Initiative. 2010. India's developing automotive fuel economy policy. https://www.globalfueleconomy.org/transport/gfei/autotool/case_studies/apacific/india/cs_ap_india.asp. Last Accessed: 2024-07-30.

Hansknecht, C., M. Klimm, Alexander Skopalik. 2014. Approximate pure nash equilibria in weighted congestion games. *Leibniz International Proceedings in Informatics, LIPIcs* **28** 242–257. doi:10.4230/LIPIcs.APPROX-RANDOM.2014.242.

Hsee, Christopher K., Ying Zeng, Xilin Li, Alex Imas. 2021. Bounded rationality in strategic decisions: Undershooting in a resource pool-choice dilemma. *Management Science* **67**(10) 6553–6567.

Jensen, Robert. 2007. The digital provide: Information (technology), market performance, and welfare in the south Indian fisheries sector. *The Quarterly Journal of Economics* **122**(3) 879–924.

Koutsoupias, Elias, Christos Papadimitriou. 1999. Worst-case equilibria. Christoph Meinel, Sophie Tison, eds., *STACS 99*. Springer Berlin Heidelberg, Berlin, Heidelberg, 404–413.

Milchtaich, Igal. 1996. Congestion games with player-specific payoff functions. *Games and Economic Behavior* **13**(1) 111 – 124.

Milchtaich, Igal. 2009. Weighted congestion games with separable preferences. *Games and Economic Behavior* **67**(2) 750 – 757.

Monderer, Dov, Lloyd S. Shapley. 1996. Potential games. *Games and Economic Behavior* **14**(1) 124 – 143.

Moreno, Diego, John Wooders. 1998. An experimental study of communication and coordination in noncooperative games. *Games and Economic Behavior* **24**(1-2) 47–76.

Paccagnan, Dario, Martin Gairing. 2024. In congestion games, taxes achieve optimal approximation. *Operations Research* **72**(3) 966–982.

Papadimitriou, Christos. 2001. Algorithms, games, and the internet. *Proceedings of the thirty-third annual ACM symposium on Theory of computing*. 749–753.

Papadimitriou, Christos H, Tim Roughgarden. 2008. Computing correlated equilibria in multi-player games. *Journal of the ACM (JACM)* **55**(3) 1–29.

Parker, Chris, Kamalini Ramdas, Nicos Savva. 2016. Is IT enough? evidence from a natural experiment in india's agriculture markets. *Management Science* **62**(9) 2481–2503.

Rabin, Matthew. 2000. Risk aversion and expected-utility theory: A calibration theorem. *Econometrica* **68**(5) 1281–1292.

Rosenthal, Robert W. 1973. A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory* **2**(1) 65–67.

Roughgarden, Tim. 2015. Intrinsic robustness of the price of anarchy. *Journal of the ACM (JACM)* **62**(5) 1–42.

Scarsini, Marco, Marc Schröder, Tristan Tomala. 2018. Dynamic atomic congestion games with seasonal flows. *Operations Research* **66**(2) 327–339.

Schulz, Andreas S, Nicolás E Stier Moses. 2003. On the performance of user equilibria in traffic networks. *SODA*, vol. 3. 86–87.

Wu, Manxi, Saurabh Amin, Asuman E. Ozdaglar. 2021. Value of information in bayesian routing games. *Operations Research* **69**(1) 148–163.

Appendix A: Pseudocode for the Algorithm

Algorithm 1: Algorithm for computing an ϵ-equilibrium

Input: An economy $E = (n, m, \beta, \boldsymbol{\alpha}, \mathbf{q}, \mathbf{C})$, $\epsilon > 0$

Output: A strategy profile $\boldsymbol{\sigma}$

$\dot{\alpha} \leftarrow \frac{\max(\boldsymbol{\alpha})}{\min(\boldsymbol{\alpha})}$, $\overline{\alpha} = \max(\boldsymbol{\alpha})$, $\overline{q} \leftarrow \max(\mathbf{q})$
$v^* \leftarrow \max\left\{ \left(\frac{\epsilon}{-\beta\overline{q}\overline{\alpha}}\right)^{\frac{1}{\beta}}, 2\overline{q} \right\}$
foreach $i \in [n]$ **do**
  $\sigma_i \leftarrow k : i \in \textsc{Closest}(k)$ `// Initialize strategy` $\boldsymbol{\sigma}$
end
while do
  if $\exists i \in [n]$ *such that* $u_i(\boldsymbol{\sigma}) < 0$ **then**
    $\boldsymbol{\sigma} \leftarrow (\boldsymbol{\sigma}_{-i}, \perp)$
  end
  else if $\forall j \in m,\ v_j(\boldsymbol{\sigma}) \geq v^*$ **then**
    if $\exists i \in [n], k \in [m]$ *such that* $u_i(\boldsymbol{\sigma}_{-i}, k) \geq u_i(\boldsymbol{\sigma}) + \epsilon$ **then**
      $\boldsymbol{\sigma} \leftarrow (\boldsymbol{\sigma}_{-i}, k)$
    end
    else
      return $\boldsymbol{\sigma}$ `//` $\boldsymbol{\sigma}$ `is an` ϵ`-equilibrium`
    end
  end
  else if $\exists i \in [n], k \in [m]$ *such that* $v_k < v^*(2\dot{\alpha})^{-\frac{m}{\beta}}$, $\sigma_i = \perp$ *and* $i \in \textsc{Closest}(k)$ **then**
    $\boldsymbol{\sigma} \leftarrow (\boldsymbol{\sigma}_{-i}, k)$
  end
  else
    let i, j, k be such that $i \in \textsc{Closest}(k), \sigma_i = j, v_k < v_j (2\dot{\alpha})^{\frac{1}{\beta}}$
    $\boldsymbol{\sigma} \leftarrow (\boldsymbol{\sigma}_{-i}, k)$
  end
end

Appendix B: Proofs

B.1. Proof of Proposition 1

THEOREM 1. *For any $\epsilon \geq 0$, there is an economy for which there does not exist an ϵ-equilibrium.*

Proof. As $\alpha_j = 1,000$ for all j and all the volumes are at least 1, no seller will go to any market more than $1,000$ distance units, as this would result in a negative utility. Hence, each seller can only go to one of its adjacent markets, or remain at home. There are only two possible destinations (not including $\perp$) for all sellers except Seller 3, who must go to Market 3 (or stay at home). The distances guarantee that no agent will choose to stay at home, therefore we only need to consider $2^5 = 32$ possible strategies. It is easy to verify that none of these strategies is in an equilibrium, as in each of the strategies, moving to a different market would improve the utility of at least one seller. Furthermore, it is easy to verify that each of these improvements is at least 0.1.

We give a high level intuition behind the example. The sellers' volumes and distances guarantee that the following hold:

1. Seller 1 prefers Market 1 to Market 2 unless Seller 6 is at Market 1.
2. Seller 2 prefers Market 2 to Market 3 unless Seller 1 is at Market 2.
3. Seller 4 prefers Market 3 to Market 4 unless Seller 2 is at Market 3.
4. Seller 5 prefers Market 4 to Market 5 unless Seller 4 is at Market 4.
5. Seller 6 prefers Market 1 to Market 4 unless Seller 5 is at Market 4.

There is a small caveat that Seller 4 prefers Market 3 if both Sellers 5 and 6 are at Market 4, however this never occurs.

As a result of these preferences, there can be no equilibrium. Assume that each seller goes to their preferred market (i.e., Sellers 1 and 6 go to Market 1 and so on). Seller 1 will deviate to Market 2, causing Seller 2 to deviate to Market 3, causing Seller 4 to deviate to Market 4, causing Seller 5 to deviate to Market 5. This in turn causes Seller 6 to deviate to Market 4, which causes Sellers 1, 2, 4 and 5 to return to Markets 1,2,3 and 4 respectively. This is the strategy from which we started. □

B.2. Proof of Theorem 2

THEOREM 2. *For any* $m > 0, \beta \in (-1, 0), 0 < \underline{\alpha} \le \overline{\alpha}, 0 < \underline{q} \le \overline{q}$*, let* $\mathcal{E}_{m,\beta,\underline{\alpha},\overline{\alpha},\underline{q},\overline{q}}$ *denote the set of economies* $E = (n, m, \beta, \boldsymbol{\alpha}, \mathbf{q}, \mathbf{C})$ *such that* $\underline{\alpha} \le \alpha_j \le \overline{\alpha}$ *for each* $j \in [m]$ *and* $\underline{q} \le q_i \le \overline{q}$ *for each* $i \in [n]$.

For any $m, \beta, \underline{\alpha}, \overline{\alpha}, \underline{q}, \overline{q}$ *as above, and any* $\epsilon, \delta > 0$*, there exists some* $n^* > 0$ *such that for every* $n \ge n^*$*, every viable* δ*-concentrated economy* $E \in \mathcal{E}_{m,\beta,\underline{\alpha},\overline{\alpha},\underline{q},\overline{q}}$ *with* n *agents admits an individually rational* $\epsilon-$*equilibrium. Furthermore, Algorithm 1 computes such an* ϵ*-equilibrium.*

Proof. The proof follows from Lemmas 2 and 3 below. Lemma 2 shows that one of the four 'if-else' clauses of Algorithm 1 (those of lines 7,10,18 and 21) always holds. This ensures that there are no runtime errors (i.e., the algorithm does not 'crash'). In Lemma 3, we define a potential function, and show that at every iteration of the 'while' loop, the value of the potential function increases. As there is a finite number of strategies, the algorithm must terminate. By the algorithm definition, it terminates at a strategy in which there are no users that can increase their utility by at least ϵ: an ϵ-equilibrium. □

We first prove a technical lemma.

LEMMA 1. *For* $\beta \in (-1, 0), \alpha_k > 0, v_k \ge q_i > 0$*, the following hold:*

- $(\beta+1) q_i \alpha_k v_k^{\beta} \left(1 + \frac{\beta q_i}{2 v_k}\right) \le \alpha_k \left((v_k + q_i)^{\beta+1} - v_k^{\beta+1}\right) \le (\beta+1) q_i \alpha_k v_k^{\beta}$,
- $\beta q_i \alpha_k v_k^{\beta} \le \alpha_k v_k \left((v_k + q_i)^{\beta} - v_k^{\beta}\right)$.

Proof. We rewrite the expressions in the lemma statement as follows.

$$\alpha_k v_k \left((v_k + q_i)^{\beta} - v_k^{\beta}\right) = \alpha_k v_k^{\beta+1} \left(\left(1 + \frac{q_i}{v_k}\right)^{\beta} - 1\right),$$

$$\alpha_k \left((v_k + q_i)^{\beta+1} - v_k^{\beta+1}\right) = \alpha_k v_k^{\beta+1} \left(\left(1 + \frac{q_i}{v_k}\right)^{\beta+1} - 1\right)$$

The Maclaurin series for $f(x) = (1+x)^b$ is as follows: $1 + bx + \frac{b(b+1)}{2} x^2 + \cdots$. From this we can deduce that, when $x \in (0, 1)$:

$$1 + (\beta+1)x + \frac{\beta(\beta+1)}{2} x^2 \le (1+x)^{\beta+1} \le 1 + (\beta+1)x,$$

and

$$1+\beta x \leq (1+x)^{\beta},$$

as β is in $[-1,0)$. The derivations follow, setting $x=\frac{q_i}{v_k}$. For clarity, we give the full derivation of the first inequality of the first bullet point as an example:

$$\begin{aligned}
\alpha_k\left((v_k+q_i)^{\beta+1}-v_k^{\beta+1}\right) &= \alpha_k v_k^{\beta+1}\left(\left(1+\frac{q_i}{v_k}\right)^{\beta+1}-1\right) \\
&\geq \alpha_k v_k^{\beta+1}\left((\beta+1)\frac{q_i}{v_k}+\frac{\beta(\beta+1)}{2}\left(\frac{q_i}{v_k}\right)^2\right) \\
&= (\beta+1)q_i\alpha_k v_k^{\beta}\left(1+\frac{\beta q_i}{2v_k}\right).
\end{aligned}$$

□

LEMMA 2. *Let $m,\beta,\underline{\alpha},\overline{\alpha},\underline{q},\overline{q},\delta$ and $\mathcal{E}_{m,\beta,\underline{\alpha},\overline{\alpha},\underline{q},\overline{q}}$ be as in Theorem 2, and let $\dot{\alpha}=\frac{\overline{\alpha}}{\underline{\alpha}}$ be the maximal ratio of any two values of α_j in any economy in $\mathcal{E}_{m,\beta,\underline{\alpha},\overline{\alpha},\underline{q},\overline{q}}$. Let $v^*\in\mathbb{R}_+$ be a real number, and set $n^*=\left(\frac{\beta+1}{2}(2\dot{\alpha})^{-m}\right)^{\frac{1}{\beta}}\frac{mv^*}{\underline{q}\delta}$. For any δ-concentrated economy $E\in\mathcal{E}_{m,\beta,\underline{\alpha},\overline{\alpha},\underline{q},\overline{q}}$ with $n\geq n^*$ agents and in any strategy $\boldsymbol{\sigma}$ in E, one of the following must hold:*

1. *for all $j\in[m]$, $v_j\geq v^*$;*
2. *there exist $i\in[n],k\in[m]$ such that (i) $v_k<v^*(2\dot{\alpha})^{-\frac{m}{\beta}}$, (ii) $\sigma_i=\perp$, and (iii) $i\in$ CLOSEST(k);*
3. *there exist $i\in[n],j,k\in[m]$ such that (i) $v_k<v_j(2\dot{\alpha})^{\frac{1}{\beta}}$, (ii) $\sigma_i=j$, and (iii) $i\in$ CLOSEST(k).*

Proof. The proof is by contradiction; we will show that it can not be that none of the conditions holds. For clarity, we use j and k to denote markets that some agent i might move *from* and *to* in the execution of the algorithm, respectively.

Assume that there exists at least one market j such that $v_j<v^*$ (otherwise Condition 1 holds). For any set of markets $X\subseteq[m]$, define CLOSEST$(X)=\bigcup_{j\in X}$ CLOSEST(j). Iteratively define the sets $X_1,X_2,\ldots$ as follows. Denote the index of the market with the smallest volume by 1, and set $X_1=\{1\}$. Let X_z be the set of markets that agents in CLOSEST(X_{z-1}) are allocated in $\boldsymbol{\sigma}$, excluding any market in $X_1,\ldots,X_{z-1}$, and excluding $\perp$. Formally, $X_z=\{\sigma_i : i\in \text{CLOSEST}(X_{z-1})\}\setminus\left(\bigcup_{j=1}^{z-1}X_j\cup\perp\right)$. Let m' denote the number of non-empty sets X_z constructed this way. Clearly, $m'\leq m$.

Let $v_z^{\max}$ denote the largest volume of any market in X_z. For every $z \in \{2,\ldots,m'\}$, it must hold that $v_{z-1}^{\max} \geq v_z^{\max}(2\dot{\alpha})^{\frac{1}{\beta}}$, otherwise there exist i,j,k that satisfy Condition 3. We therefore have the following inequalities.

$$v^* > v_1^{\max} \geq v_{m'}^{\max}(2\dot{\alpha})^{\frac{m'}{\beta}} \geq v_{m'}^{\max}(2\dot{\alpha})^{\frac{m}{\beta}}.$$

For every $k \in \bigcup_{z\in\{1,\ldots m'\}} X_z$, it holds that $v_k < v^*(2\dot{\alpha})^{-\frac{m}{\beta}}$. As Condition 2 does not hold, there is no market $k \in \bigcup_{z\in\{1,\ldots m'\}} X_z$ for which there is an agent $i \in \textsc{Closest}(k)$ such that $\sigma_i = \perp$. Denote the number of markets in the union of all of the sets X_z by m'': $m'' = \left|\bigcup_{z\in\{1,\ldots m'\}} X_z\right|$. There must be strictly less than $\frac{m''}{\underline{q}}v^*(2\dot{\alpha})^{-\frac{m}{\beta}} = m''\delta n^*$ agents in total at markets $k \in \bigcup_{z\in\{1,\ldots m'\}} X_z$, and hence in $\textsc{Closest}\left(\bigcup_{z\in\{1,\ldots m'\}} X_z\right)$. But by the definition of δ-concentration, and the fact that there is no market $k \in \bigcup_{z\in\{1,\ldots m'\}} X_z$ for which there is an agent $i \in \textsc{Closest}(k)$ such that $\sigma_i = \perp$, there must be at least $m''\delta n$ agents in $\textsc{Closest}\left(\bigcup_{z\in\{1,\ldots m'\}} X_z\right)$, a contradiction. □

Lemma 3. *Let $m,\beta,\underline{\alpha},\overline{\alpha},\underline{q},\overline{q},\delta,\epsilon$ and $\mathcal{E}_{m,\beta,\underline{\alpha},\overline{\alpha},\underline{q},\overline{q}}$ be as in Theorem 2, and let $\dot{\alpha} = \frac{\overline{\alpha}}{\underline{\alpha}}$ be the maximal ratio of any two values of α_j in any economy in $\mathcal{E}_{m,\beta,\underline{\alpha},\overline{\alpha},\underline{q},\overline{q}}$. Set $v^* = \max\left\{\left(\frac{\epsilon}{-\beta\overline{q}\overline{\alpha}}\right)^{\frac{1}{\beta}}, 2\overline{q}\right\}$, $n^* = \left(\frac{\beta+1}{2}(2\dot{\alpha})^{-m}\right)^{\frac{1}{\beta}}\frac{mv^*}{\underline{q}\delta}$. For any viable δ-concentrated economy $E \in \mathcal{E}_{m,\beta,\underline{\alpha},\overline{\alpha},\underline{q},\overline{q}}$ with $n \geq n^*$ agents, let $\boldsymbol{\sigma}_1,\ldots,\boldsymbol{\sigma}_S$ be the strategies generated during the execution of Algorithm 1 on input E,δ,ϵ. Define the potential function $\Phi(\boldsymbol{\sigma}) = \sum_{i\in[n]} u_i(\boldsymbol{\sigma})$. Then for all $s \in [S-1]$, $\Phi(\boldsymbol{\sigma}_s) \leq \Phi(\boldsymbol{\sigma}_{s+1})$.*

We note that the potential function Φ was first defined by Rosenthal (1973). We first prove a simple result that we will need for the proof of Lemma 3.

Lemma 4. *let $\boldsymbol{\sigma}_1,\ldots,\boldsymbol{\sigma}_S$ be the strategies generated during the execution of Algorithm 1 on input E,δ,ϵ, where E,δ,ϵ are as in Lemma 3. For any strategy $\boldsymbol{\sigma} \in \{\boldsymbol{\sigma}_1,\ldots,\boldsymbol{\sigma}_S\}$, it holds that for every $j \in m$, $v_j(\boldsymbol{\sigma}) > \overline{q}$.*

Proof. We prove a stronger result: that for every $\boldsymbol{\sigma} \in \{\boldsymbol{\sigma}_1,\ldots,\boldsymbol{\sigma}_S\}, j \in m$, it holds that (i) $v_j(\boldsymbol{\sigma}) > \overline{q}$ and (ii) for every i such that $\sigma_i = j$, the utility of agent i would be strictly positive at market j when $v_j \leq v^* - \overline{q}$ (formally, for every i such that $\sigma_i = j$, and strategy $\boldsymbol{\sigma}'$ such that $\sigma'_\ell = j$ and $v_j(\boldsymbol{\sigma}') \leq v^* - \overline{q}$, it holds that $u_i(\boldsymbol{\sigma}') > 0$). The proof is by induction. For the base case, this clearly holds for $\boldsymbol{\sigma}_1$, as the economy is viable, the volumes of all markets are greater than v^* and all utilities are non-negative. For the inductive hypothesis, assume this holds for $\boldsymbol{\sigma}_{s-1}$. From the algorithm description, we can see that no agent leaves any market j such that $v_j(\boldsymbol{\sigma}_{s-1}) \leq v^*$ if (i) and (ii) hold for $\boldsymbol{\sigma}_{s-1}$. Furthermore, for any agent i that leaves any market j to a market k, it must hold that $v_k + q_i \geq v^* - \overline{q}$, and $u_i(\boldsymbol{\sigma}_s) > 0$. Therefore (i) and (ii) must also hold for $\boldsymbol{\sigma}_s$. □

We are now ready to prove Lemma 3.

Proof of Lemma 3. Consider any iteration $s \in [S-1]$ of the 'while' loop in Algorithm 1 (i.e., any iteration that is not the last iteration). For clarity, denote the strategy at the start of the loop by $\boldsymbol{\sigma}$, and strategy at the start of the following loop $\boldsymbol{\sigma'}$ (i.e., $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma'}$ correspond to $\boldsymbol{\sigma}_s$ and $\boldsymbol{\sigma}_{s+1}$ in the lemma statement). Denote by i the agent whose market allocation is different in $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma'}$; set $j = \sigma_i$ and $k = \sigma'_i$. Denote the revenue of agent i for strategy $\boldsymbol{\sigma}$ by $r_i(\boldsymbol{\sigma})$, let $r(\boldsymbol{\sigma}) := \sum_{i \in [n]} r_i(\boldsymbol{\sigma})$ and let $r_{-i}(\boldsymbol{\sigma}) := r(\boldsymbol{\sigma}) - r_i(\boldsymbol{\sigma})$.

Denote the volumes of markets j and k in $\boldsymbol{\sigma}$, *excluding* the volume of agent i, by v_j and v_k, respectively. Denote $\Delta c_i(\boldsymbol{\sigma}, \boldsymbol{\sigma'}) = c_{ij} - c_{ik}$; this is the change in the transportation cost of agent i when deviating from $\boldsymbol{\sigma}$ to $\boldsymbol{\sigma'}$. In addition, denote $\Delta r_i(\boldsymbol{\sigma}, \boldsymbol{\sigma'}) = r_i(\boldsymbol{\sigma'}) - r_i(\boldsymbol{\sigma})$, $\Delta r(\boldsymbol{\sigma}, \boldsymbol{\sigma'}) = r(\boldsymbol{\sigma'} - r(\boldsymbol{\sigma})$ and $\Delta r_{-i}(\boldsymbol{\sigma}, \boldsymbol{\sigma'}) = r_{-i}(\boldsymbol{\sigma'}) - r_{-i}(\boldsymbol{\sigma})$. The change in agent i's utility for deviating from $\boldsymbol{\sigma}$ to $\boldsymbol{\sigma'}$ is:

$$\begin{aligned}\Delta u_i(\boldsymbol{\sigma}, \boldsymbol{\sigma'}) &= \Delta r_i(\boldsymbol{\sigma}, \boldsymbol{\sigma'}) + \Delta c_i(\boldsymbol{\sigma}, \boldsymbol{\sigma'}) \\ &= q_i\left(\alpha_k(v_k + q_i)^\beta - \alpha_j(v_j + q_i)^\beta\right) - c_{ik} + c_{ij}.\end{aligned}$$

The change in the potential following agent i's deviation can be represented in two equivalent ways:

$$\begin{aligned}\Delta\Phi(\boldsymbol{\sigma}, \boldsymbol{\sigma'}) &= \Delta u_i(\boldsymbol{\sigma}, \boldsymbol{\sigma'}) + \Delta r_{-i}(\boldsymbol{\sigma}, \boldsymbol{\sigma'}) \\ &= \Delta r(\boldsymbol{\sigma}, \boldsymbol{\sigma'}) + \Delta c_i(\boldsymbol{\sigma}, \boldsymbol{\sigma'}).\end{aligned}$$

Finally, note that

$$\Delta r_{-i}(\boldsymbol{\sigma}, \boldsymbol{\sigma'}) = \alpha_j(v_j)\left(v_j^\beta - (v_j + q_i)^\beta\right) + \alpha_k v_k\left((v_k + q_i)^\beta - v_k^\beta\right),$$

as only agents at markets j and k experience a change in utility. We now consider the four 'if-else' clauses of Algorithm 1.

Case 1 (Lines 7-9 of Algorithm 1):

There exists some agent i for whom $u_i(\boldsymbol{\sigma}) < 0$, and we set $\boldsymbol{\sigma}' = (\boldsymbol{\sigma}_{-i}, \perp)$. The change in potential is as follows:

$$\begin{aligned}\Delta\Phi(\boldsymbol{\sigma},\boldsymbol{\sigma}') &= \Delta u_i(\boldsymbol{\sigma},\boldsymbol{\sigma}') + \Delta r_{-i}(\boldsymbol{\sigma},\boldsymbol{\sigma}') \\ &> \alpha_j v_j \left(v_j^\beta - (v_j + q_i)^\beta\right) \\ &> 0,\end{aligned}$$

where the first inequality is because $u_i(\boldsymbol{\sigma}) < 0$ and the second is because the revenue for the agents in market j increases when the volume at the market goes down.

Case 2 (Lines 10-13 of Algorithm 1):

In this case, for all $j \in m$, $v_j(\boldsymbol{\sigma}) \geq v^*$, and $\exists i \in [n], k \in [m]$ such that $u_i(\boldsymbol{\sigma}_{-i}, k) \geq u_i(\boldsymbol{\sigma}) + \epsilon$. Note that this case also holds for $\sigma_i = \perp$. We first show that $\alpha_k v_k \left((v_k + q_i)^\beta - v_k^\beta\right) \geq -\epsilon$:

$$\alpha_k v_k \left((v_k + q_i)^\beta - v_k^\beta\right) \geq \beta \alpha_k v_k^\beta q_i \tag{2}$$

$$\geq \beta \overline{\alpha q} (v^*)^\beta \tag{3}$$

$$= -\epsilon \tag{4}$$

Ineq. (2) is due to Lemma 1; Ineq. (3) is because β is negative; Eq. (4) is due to the definition of v^*. We are now ready to bound the change in the potential function.

$$\begin{aligned}\Delta\Phi(\boldsymbol{\sigma},\boldsymbol{\sigma}') &= \Delta u_i(\boldsymbol{\sigma},\boldsymbol{\sigma}') + \Delta r_{-i}(\boldsymbol{\sigma},\boldsymbol{\sigma}') \\ &= \Delta u_i(\boldsymbol{\sigma},\boldsymbol{\sigma}') + \alpha_j v_j \left(v_j^\beta - (v_j + q_i)^\beta\right) + \alpha_k v_k \left((v_k + q_i)^\beta - v_k^\beta\right) \\ &\geq \Delta u_i(\boldsymbol{\sigma},\boldsymbol{\sigma}') + \alpha_k v_k \left((v_k + q_i)^\beta - v_k^\beta\right) \qquad (5) \\ &\geq \epsilon - \epsilon = 0\end{aligned}$$

where Ineq. (5) is because $\alpha_j v_j \left(v_j^\beta - (v_j + q_i)^\beta\right) > 0$, as the revenue for the agents in market j increases when the volume at the market goes down.

Case 3 (Lines 18-20 of Algorithm 1):

From the assumption that the economy is viable, we know that $\forall i \in [n], u_i(\textsc{Nearest}) > 0$. Let k denote the market such that $i \in \textsc{Closest}(k)$. As the economy is δ-concentrated, the volume of market k under the strategy Nearest is at least

$$\delta n \underline{q} \geq \delta n^* \underline{q} = \left(\frac{\beta+1}{2}(2\dot{\alpha})^{-m}\right)^{\frac{1}{\beta}} m v^*.$$

Therefore,

$$\begin{aligned} u_i(\textsc{Nearest}) &\leq q_i \alpha_k \left(\left(\frac{\beta+1}{2}(2\dot{\alpha})^{-m}\right)^{\frac{1}{\beta}} m v^*\right)^{\beta} - c_{ik} \\ &\leq \frac{\beta+1}{2} q_i \alpha_k (2\dot{\alpha})^{-m} {v^*}^{\beta} - c_{ik}. \end{aligned}$$

So, as $u_i(\textsc{Nearest}) \geq 0$, we have that

$$c_{ik} \leq \frac{\beta+1}{2} q_i \alpha_k (2\dot{\alpha})^{-m} {v^*}^{\beta}. \tag{6}$$

We bound the change in potential:

$$\begin{aligned} \Delta\Phi(\boldsymbol{\sigma}, \boldsymbol{\sigma}') &= \Delta r(\boldsymbol{\sigma}, \boldsymbol{\sigma}') + \Delta c_i(\boldsymbol{\sigma}, \boldsymbol{\sigma}') \\ &= \alpha_k \left((v_k + q_i)^{\beta+1} - v_k^{\beta+1}\right) - c_{ik} \\ &\geq (\beta+1) q_i \alpha_k v_k^{\beta} \left(1 + \frac{\beta q_i}{2 v_k}\right) - c_{ik} && (7) \\ &\geq (\beta+1) q_i \alpha_k \frac{v_k^{\beta}}{2} - c_{ik} && (8) \\ &\geq \frac{(\beta+1) q_i \alpha_k}{2} \left(v^* (2\dot{\alpha})^{-\frac{m}{\beta}}\right)^{\beta} - c_{ik} && (9) \\ &= \frac{(\beta+1) q_i \alpha_k}{2} (2\dot{\alpha})^{-m} {v^*}^{\beta} - c_{ik} \\ &\geq 0. && (10) \end{aligned}$$

Ineq. (7) is due to Lemma 1; Ineq. (8) is because, from Lemma 4, $v_k \geq q_i$; Ineq. (9) is because of the assumption of Case 3; Ineq. (10) is due to Ineq. (6).

Case 4 (Lines 21-24 of Algorithm 1):

Clearly, as $\sigma_i' = k$ and $i \in \textsc{Closest}(k)$, it must hold that $\Delta c_i(\boldsymbol{\sigma}, \boldsymbol{\sigma}') \geq 0$. We now bound the change in potential resulting from chagning the strategy from $\boldsymbol{\sigma}$ to $\boldsymbol{\sigma}'$.

$$\begin{aligned}
\Delta\Phi(\boldsymbol{\sigma}, \boldsymbol{\sigma}') &= \Delta r(\boldsymbol{\sigma}, \boldsymbol{\sigma}') + \Delta c_i(\boldsymbol{\sigma}, \boldsymbol{\sigma}') \\
&\geq \alpha_k\left((v_k + q_i)^{\beta+1} - v_k^{\beta+1}\right) - \alpha_j\left((v_j + q_i)^{\beta+1} - v_j^{\beta+1}\right) && (11)\\
&\geq q_i(\beta+1)\left(\alpha_k v_k^{\beta}\left(1 + \frac{\beta q_i}{2v_k}\right) - \alpha_j v_j^{\beta}\right) && (12)\\
&\geq \underline{q}(\beta+1)\alpha_k\left(\frac{v_k^{\beta}}{2} - \dot{\alpha} v_j^{\beta}\right) && (13)\\
&\geq 0. && (14)
\end{aligned}$$

Ineq. (11) is because $\Delta c_i(\boldsymbol{\sigma}, \boldsymbol{\sigma}') \geq 0$; Ineq. (12) is due to Lemma 1; Ineq. (13) is because, from Lemma 4, $v_k > q_i$; Ineq. (14) is because $v_k < v_j(2\dot{\alpha})^{\frac{1}{\beta}}$, as this implies that $\frac{v_k^{\beta}}{2\dot{\alpha}} > v_j^{\beta}$ (recall that $\beta \in [-1, 0)$). □

B.3. Proof of Theorem 3

Theorem 3. *For any $m > 0, \beta \in (-1, 0), 0 < \underline{\alpha} \leq \overline{\alpha}, 0 < \underline{q} \leq \overline{q}$, let $\mathcal{E}_{m,\beta,\underline{\alpha},\overline{\alpha},\underline{q},\overline{q}}$ denote the set of economies $E = (n, m, \beta, \boldsymbol{\alpha}, \mathbf{q}, \mathbf{C})$ such that $\underline{\alpha} \leq \alpha_j \leq \overline{\alpha}$ for each $j \in [m]$ and $\underline{q} \leq q_i \leq \overline{q}$ for each $i \in [n]$.*

For any $m, \beta, \underline{\alpha}, \overline{\alpha}, \underline{q}, \overline{q}$ as above, and any $\epsilon, \delta > 0$, there exists some $n^* > 0$ such that for every $n \geq n^*$, the price of ϵ-stability of every viable δ-concentrated economy $E \in \mathcal{E}_{m,\beta,\underline{\alpha},\overline{\alpha},\underline{q},\overline{q}}$ with n agents is 1.

Proof. To prove the theorem, we show that any optimal solution is an ϵ-equilibrium. Let E be an economy as in the theorem statement, and let $\boldsymbol{\sigma}^*$ be a welfare-maximizing strategy. From Lemma 2, one of the following must hold for $\boldsymbol{\sigma}^*$:

1. for all $j \in [m]$, $v_j \geq v^*$;
2. there exist $i \in [n], k \in [m]$ such that (i) $v_k < v^*(2\dot{\alpha})^{-\frac{m}{\beta}}$, (ii) $\sigma_i = \perp$, and (iii) $i \in \textsc{Closest}(k)$;
3. there exist $i \in [n], j, k \in [m]$ such that (i) $v_k < v_j(2\dot{\alpha})^{\frac{1}{\beta}}$, (ii) $\sigma_i = j$, and (iii) $i \in \textsc{Closest}(k)$.

Therefore, one of the five conditions of the while loop of Algorithm 1 must hold. The proof of Lemma 3 shows that if the conditions of (i) Lines 7-9, (ii) Lines 10-13, (iii) Lines 18-20, or (iv) Lines 21-23 hold, there exist a deviation that increases the total welfare, a contradiction to the optimality of $\boldsymbol{\sigma}^*$. Therefore, the condition of Line 15 must hold; i.e., $\boldsymbol{\sigma}^*$ in an $\epsilon-$equilibrium. □

Appendix C: Farmer Survey

The survey was completed by a market research firm based in India. The survey was conducted in the farmers' native language when possible or Hindi. In total, 345 onion farmers and 157 potato farmers responded. In addition to the questions below, farmers were also asked 'How do you chose which market to sell your produce at?' and given the option to provide an open text response. Different ways of saying the price at the market and the distance to the market were the dominate explanations provided. The main questions, with their response distributions are given in Tables 5-14 below.

	Farming Experience					
	Less than 1 year	1-5 years	6-10 years	11-15 years	16-20 years	20+ years
Onion	1.2%	18.0%	30.0%	18.8%	15.4%	17.1%
Potato	0.0%	29.3%	34.4%	10.0%	12.1%	13.4%

Table 5: Response to the question: For how many years have you been farming?

	Farm Size				
	Less than 1 acre	1-2 acres	2-5 acres	5-10 acres	10+ acres
Onion	2.9%	28.4%	45.8%	17.4%	5.5%
Potato	3.8%	26.1%	45.2%	19.1%	5.7%

Table 6: Response to the question: What is the size of the farm you own in acres?

	Sales Method			
	Direct Market	Aggregator	E-marketplaces	Other
Onion	78.0%	20.0%	2.0%	0.0%
Potato	66.2%	28.0%	5.7%	0.0%

Table 7: Response to the question: How do you typically sell your produce?

	Quantity (Kg)	
	Mean	Standard Deviation
Onion	2038	6191
Potato	1610	2967

Table 8: Response to the question: How many kilos of onions or potatoes do you take to the market in one trip?

	Selling Frequency					
	Daily	Once a week	Once every two weeks	Once a month	Quarterly	All at once
Onion	6.4%	20.3%	19.7%	15.7%	11.3%	26.7%
Potato	8.9%	21.0%	14.0%	14.0%	14.0%	28.0%

Table 9: Response to the question: During a typical selling season, how frequently do you sell your produce in the market?

	Factors Considered						
	Prev. day's price	Family tradition	Amount produced	Distance to market	Market size	Transp. cost	Relation with middleman
Onion	39.4%	13.9%	28.7%	28.4%	31.9%	40.6%	33.9%
Potato	37.6%	10.8%	31.8%	33.8%	38.2%	45.2%	33.8%

Table 10: Response to the question: What factors do you typically consider when deciding where to sell your produce?

	Number of Farmers Consulted					
	0	1	2	3	4	5+
Onion	11.0%	16.8%	25.2%	20.9%	11.6%	14.5%
Potato	20.4%	15.3%	21.0%	21.0%	7.6%	14.6%

Table 11: Response to the question: How many other farmers do you talk to when deciding where to sell your produce?

	Common Market Choice	
	Yes	No
Onion	91.0%	9.0%
Potato	84.7%	15.3%

Table 12: Response to the question: Do the farmers close to you generally go to the same market as you?

	Distance to Nearest Market (Kms)	
	Mean	Standard Deviation
Onion	20.159	29.367
Potato	14.720	27.101

Table 13: Response to the question: How far are you from the nearest market? (in Kms)

	Distance to Usual Market (Kms)	
	Mean	Standard Deviation
Onion	42.172	83.317
Potato	35.433	60.254

Table 14: Response to the question: What is the distance between your location and the market you usually sell your produce to? (in Kms)

Appendix D: Numerical Analysis using Synthetic Data

In Section 7, we used synthetic data to explore how Rec performs on several key metrics compared to other strategies when parameters change. In Subsection D.1 we expand on the baseline results using the parameters in Table 4. In Subsection D.2 we study how the key metrics change as a function of the price elasticity and seller volume. In Subsection D.3, we assess how much of the optimal welfare Rec captures.

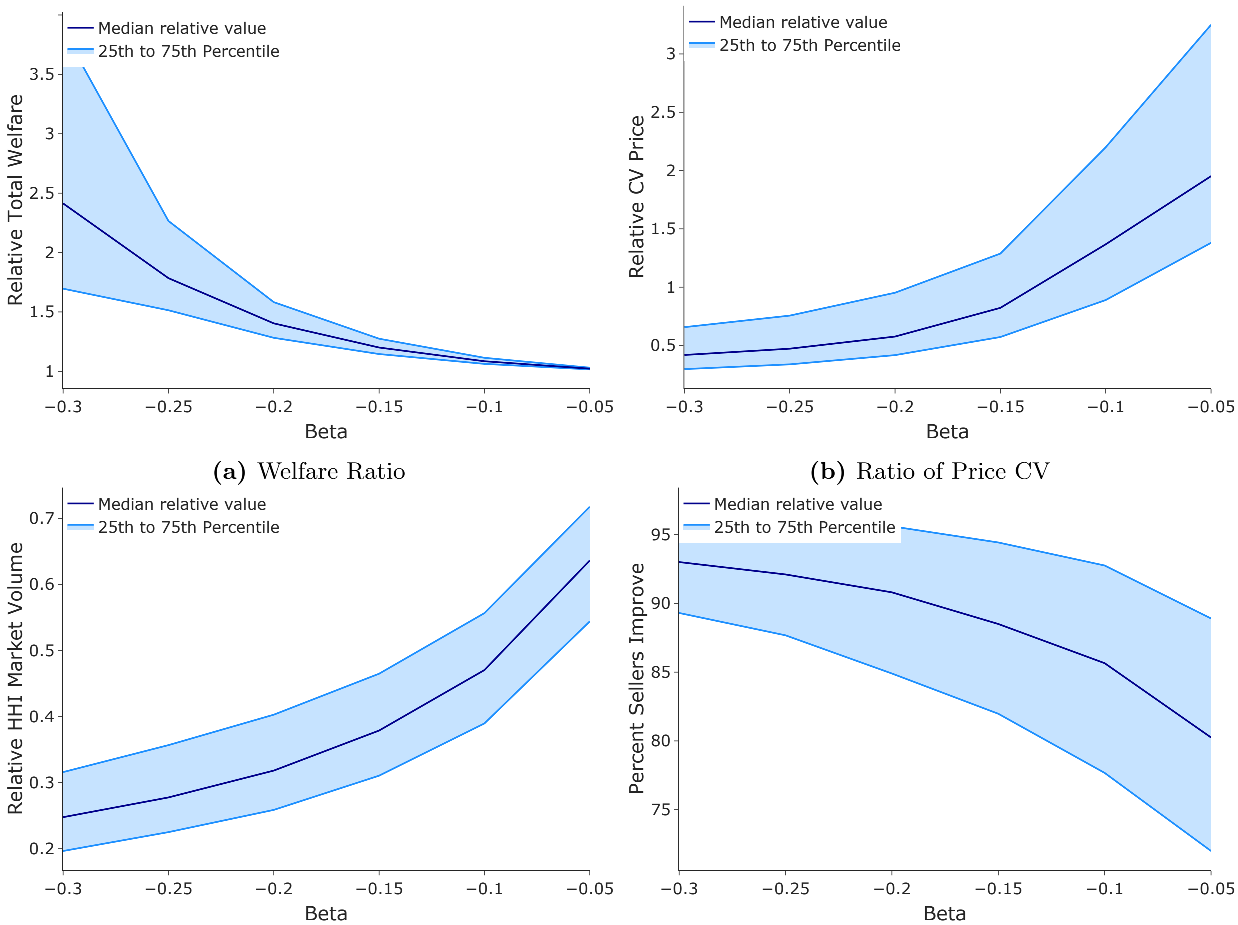
D.1. Baseline Results

The results of our baseline comparisons are summarized in Table 15. When the sellers use the Rec strategy, their total welfare is higher than when they use any other strategy. On the individual level, at least two thirds of the sellers have higher utilities under Rec. We can see that there is some trade-off between geographic price dispersion and market concentration (except for BR, which is worse than Rec with respect to both metrics). Importantly, we note that HHIR($\boldsymbol{\sigma}$) is always lower when CVR($\boldsymbol{\sigma}$) is higher. This is because the strategies for which this happens (Greedy and IBR) cause sellers to be concentrated in the largest markets, which in turn leads to a smaller variation in prices among the non-empty markets. Market concentration is lowest in Nearest, which is not surprising as the sellers and markets are positioned uniformly at random in our synthetic data. Comparing with Subsection 6, we see that when the sellers and markets are distributed more realistically, this advantage typically disappears. Finally, no strategy except Rec ever organically reaches an equilibrium. This implies that there are always sellers that regret their decision. In fact, we can see that at least three quarters of the sellers are regretful when using any baseline strategy. We also simulated scenarios in which not all sellers used the same strategy with a portion playing Nearest, Greedy, and BR. The results, available upon request, are similar to those in Table 15.

	Nearest	Greedy	BR	IBR
Welfare Ratio	1.263	1.085	1.446	1.216
Percent Better Under Rec	67.890	84.685	91.390	94.722
Improvement Percent	22.911	9.038	31.248	20.111
Ratio of Price CV	0.368	1.367	0.578	2.011
HHI Ratio	2.097	0.470	0.452	0.362
Percent Regretful	75.471	94.050	98.201	99.664

Median values of total seller welfare, coefficient of variation of market prices, HHI of market volumes, improvement percent, and percent of sellers better off using the Rec strategy for different seller strategies. Rec outperforms S when $\mathrm{SWR}(S) > 1$, $\mathrm{CVR}(S) < 1$, and $\mathrm{HHIR}(S) < 1$.

Table 15: Summary of key metrics.



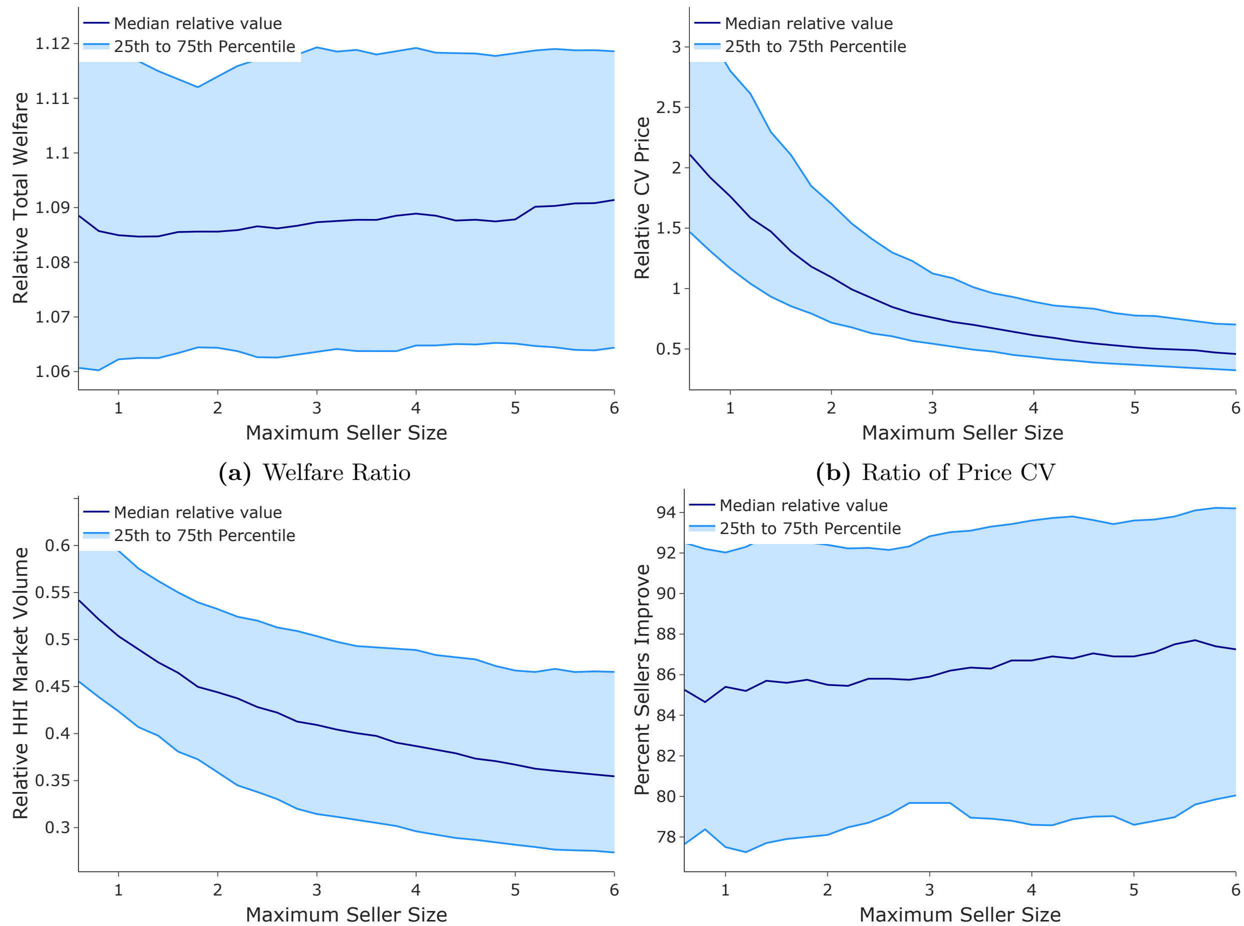
(a) Welfare Ratio

(b) Ratio of Price CV

(c) HHI Ratio

(d) Percent Better under Rec

Figure 6: Sensitivity of key metrics to price elasticity.



(a) Welfare Ratio **(b)** Ratio of Price CV

(c) HHI Ratio **(d)** Percent Better under Rec

Figure 7: Sensitivity of key metrics to seller size.

D.2. Sensitivity to Price Elasticity and Seller Volumes

In order to study the sensitivity of the key metrics to price elasticity, we vary the value of β from -0.05 to -0.3 in increments of 0.05. Figure 6 shows how four of the metrics are affected by β; for all four metrics, the advantages of Rec over Greedy increase as the price elasticity increases. Rec is better than Greedy for all values of β for both total welfare and market concentration and outperforms Greedy with respect to price CV for $\beta \leq -0.15$. As β decreases further, we see that the price variance when sellers use Rec is almost always better than for Greedy. The impact of a high elasticity can be substantial; for $\beta = -0.3$, sellers' welfare is about 2.5 times greater, almost all of the sellers are better off, and both price dispersion and market concentration are substantially better under Rec.

We also vary the distribution from which seller volumes are drawn, keeping the value of β as in Table 4. We change the top of the uniform distribution of seller volumes to 0.6 to 6 in increments of 0.2, keeping the bottom of the distribution 0.5 throughout. Again, we compare the Rec and Greedy strategies. The results are displayed in Figure 7. Welfare is largely unchanged as seller sizes increase. However, geographic price dispersion can be either lower (for small values of seller size) or higher (for large values of seller size).

D.3. Efficiency of the Recommendation Strategy

The example of Figure 1 shows that different equilibria can have different welfare. Furthermore, there exist cases where the optimal welfare is not at an equilibrium. Instead of recommending an equilibrium to the sellers, one might consider computing an allocation that maximizes total seller welfare and recommending those markets instead. There are two difficulties with this approach: (i) computing the optimal seller welfare is NP-hard for a large range of inputs (Chakrabarty et al. 2005), and (ii) the resulting allocation might not be an equilibrium, causing sellers to doubt whether the recommender is acting in their best interest. Nevertheless, we would like to determine how much of the total possible welfare is typically extracted by our strategy. In order to determine the optimal seller welfare, we implement a branch-and-bound optimization algorithm for instances with three to five markets and ten to twenty sellers.[9] We evaluate a total of 500 scenarios for each combination of number of markets and number of sellers for a total of 16,500 scenarios. In each scenario, we compare the optimal to Rec to determine how much of the seller welfare it obtains.

The Rec strategy captures 99.84% of the maximum total welfare on average and at least 98.85% of the maximum total welfare in 99% of the scenarios. Figure 8 shows the average relative welfare captured by Rec as the number of sellers and markets changes. While increasing the number of markets does reduce the relative welfare captured by Rec, increasing the number of farmers increases the relative welfare captured. Furthermore, we find that the optimal solution is an equilibrium in only 53.93% of the scenarios; in nearly half of our simulations some sellers would regret their decision to follow the welfare-maximizing recommendation despite the overall welfare being higher.

Appendix E: Estimating the Parameters

Here we provide more details about the process used to estimate the parameters and choose the crops for the scenarios of Section 6. To estimate the parameters in Model 1, we focus on crops and crop-markets that have sufficient data. As we noted in Section 6, the price data includes the price and volume transacted for 367 crops in 1,318 markets throughout 16 states. The data starts

[9] The branch and bound algorithm starts with all sellers staying home. Branches are then created as sellers are allocated to each of the markets. We note that the market price decreases when another seller is allocated to a market which impacts the utility of sellers who have already been allocated to this market. Branches are pruned when either 1) the utility of any allocated sellers is negative or 2) the sum of utilities for allocated sellers plus the sum of best potential (greedy) utilities is less than the best total utility observed so far.

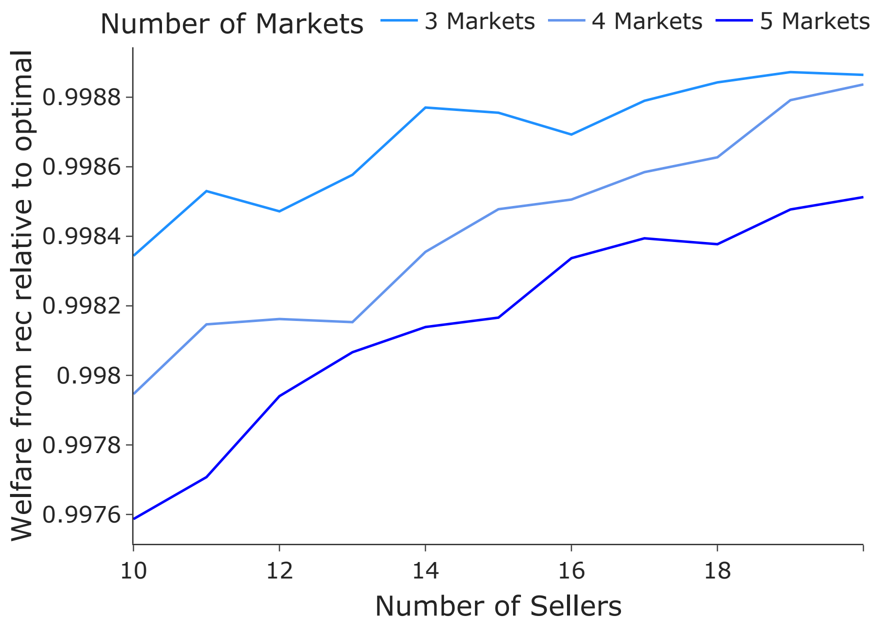
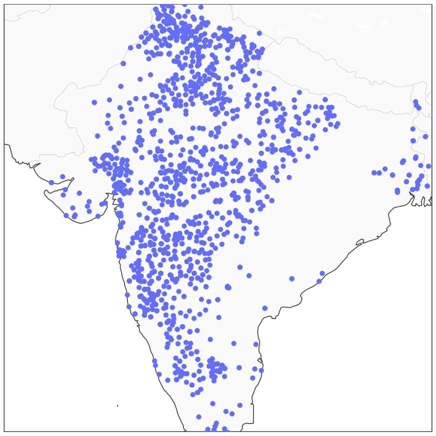


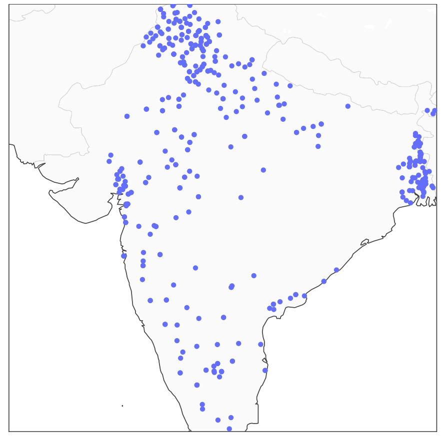
Figure 8: Relative Utility with REC vs. Welfare Maximizing

on January 1, 2008 and ends on January 31, 2011 for a total of 1,848,349 crop-market-days. The subscription data covers the same time period and informs us which two crops the farmer chose to receive price information on, and for each of those two crops, which three markets. We first remove any crop-markets with less than 30 observations to allow for reasonable estimation of α_m and then removing any crops with less than eight markets or less than 1,200 subscribers. This process results in 1,364,942 market-days across 65 crops in 1,099 markets throughout 15 states. After completing the data filtering steps above, we drop seventeen crops for which the estimated β is positive. This leaves us with 1,056,581 crop-market-day observations across 48 crops in 1,0003 markets throughout 15 states. For these crops, we observe 520,785 subscriptions, where a subscription is one farmer subscribed to one crop. A farmer can subscribe to two crops for one set price, but we count them as two subscriptions since they are for different crops.

Estimating farmer volume is more difficult, as we do not observe any information about the farmers' cultivated land area. Instead, we estimate a distribution of farmer sizes using the number of subscribers for a market and the volume at the markets. In particular, we first calculate the total volume transacted for each crop-market over the course of our data. We then divide by the number of (fractional) years between the first price observation and the last price observation that we have for the crop-market to obtain an estimate of the volume per year for each crop-market. We then convert to the volume per farmer per year for each crop-market by dividing by the number of farmers who subscribe to that crop-market. However, we can not assume that the market volume is fully accounted for by subscribed farmers because other farmers who do not have an RML



(a) Farmer locations



(b) Market locations

Figure 9: Farmer and market locations throughout India.

subscription also sell at the markets. To account for this, we assume that RML has a 5% market share for all crops and, hence, divide the volume per year by the number of subscribers multiplied by 20. Furthermore, we assume that farmers make five trips to the market each year, so we divide by five to obtain the volume per farmer per trip to the market. These two assumptions are based on conversations with farmers and employees at RML. In the numerical analysis, we obtain each farmer's volume by drawing from a truncated normal distribution, where the parameters are the mean and standard deviation of each markets' estimated volume per farmer per trip (calculated independently for each crop).

We note that we do not intend for the estimated parameters to be perfectly correct. A service provider would want to more clearly understand any bias in this estimation and control for any observable market-time varying characteristics. They may also want to implement an instrumental variable regression to remove any bias in the estimated coefficients. We do not run these regressions as we are using the estimates only to show a proof a concept.