

Dynamic Learning over Beliefs about Farming in the American West*

Katherine Hauck and Tiemen Woutersen

October 22, 2023

Abstract

Many settings involve dynamic learning and its consequences, including farmers learning about fertilizer in developing countries, students learning about the returns to college majors, and firms learning about the demand for their products. To reflect this dynamic learning, we extend a popular dynamic optimization model to allow agents to change their information sets over time and to update their beliefs. In particular, we allow for dynamic optimization, unobserved differences between agent types, and allow these types to have biased Bayesian beliefs. We apply this model to the setting of individual land acquisition in the late 19th century, when the U.S. government granted over 11% of the land area in the country through homesteading. Individuals who homesteaded were required to farm the land. Thus, these individuals learned about the value of farming. They used this updated belief about the value of farming to decide whether to (i) acquire the title of the farm by farming for five years, (ii) abandon the farm, or (iii) sell the farm. We digitize and match novel plot-level histories of land acquisition, land resale, and agricultural production to estimate the farmers' learning process. We use our model to construct counterfactuals which show the impact of individuals' beliefs on their decisions. Specifically, 62% of the farmers who abandoned their farms in the first five years would not have done so if they had begun with the beliefs they would have held after ten years of farming. In a second counterfactual, we demonstrate that if the U.S. government had not offered homesteading, only 31% of the homesteaders would have purchased the land instead and 69% would have opted not to farm. These results indicate that without the Homestead Act, the speed of western expansion would have been reduced by 38%.

Keywords: Dynamic optimization, Bayesian updating, Homesteading, Optimality of contracts

JEL classification: C11, C61, C80, N51

*We thank Zach Brown, Gautam Gowrisankaran, Ivana Komunjer, Ashley Langer, Juan Pantano, Tauhidur Rahman, John Rust, and Matthijs Wildenbeest for very helpful discussions. This work was funded by the Economic History Association, the National Science Foundation (#2117324), and the University of Arizona.

1 Introduction

How do individuals' changing beliefs about their own ability impact their decisions over time? We address this question in the setting of land acquisition and farming in the nineteenth century under the Homestead Act. The Homestead Act, beginning in 1863, granted more than 270 million acres of land in the western part of the U.S., opening up this area to low-cost, low-risk settlement to individual farmers (National Park Service). This amount represents more than 11% of the land in the U.S.. Widely regarded to be one of the most influential land laws in American history, the Homestead Act provided for nearly all the land west of Ohio to be available to individual farmers at a low cost, fundamentally reshaping the demographic and agricultural makeup of the western half of America. At the same time, this land was also available for direct purchase by those same individuals because the Homestead Act was superimposed on top of a pre-existing land law system.

Farmers learning new information over time is a key feature of our setting. We consider two counties in eastern Kansas, in which 24% homesteaders abandoned their farms. Further, the majority of this abandonment happened quickly, within the first two years of the homesteader starting the farm. An additional 31% of homesteaders sold their farm within about the first 6 years. These proportions are shown in Figure 1. This high rate of abandonment and selling for homesteaders - and how quickly it occurred after the homesteader started farming - indicate that farmers had a high amount of uncertainty about the value of farming in Kansas. When farmers learned new information through the process of actually farming, they updated their beliefs about the value of farming, and used those updated beliefs to make new decisions, causing many farmers to abandon or sell their homesteads. Therefore, learning new information over time is highly relevant to the decision-making process of farmers in our setting. Because of the importance of learning to farmers, we quantify how mistaken beliefs about farming affected land acquisition, and how those mistaken beliefs mattered for western expansion.

In order to capture how learning new information over time impacted farmers' decisions, and therefore western expansion, we build a dynamic discrete choice model in which farmers can make a decision every 6 months between (i) continuing to farm, (ii) selling their farm, or (iii) abandoning their farm. However, because learning over time is such a crucial component in farmers' decision making, we incorporate this learning into the model. Therefore, our methodological innovation in this paper is to expand the typical full-information dynamic

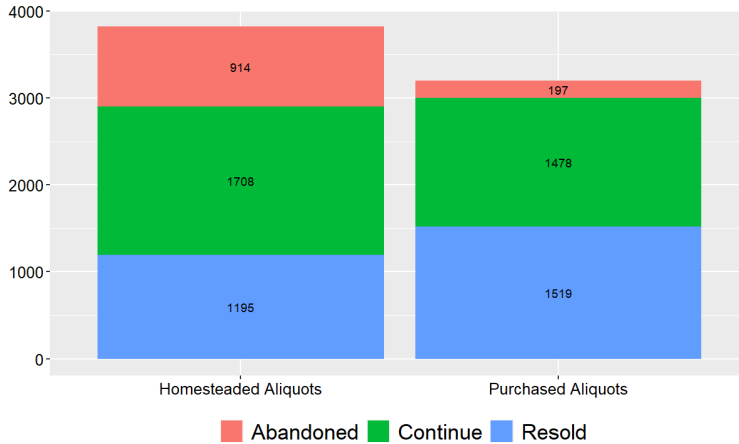


Figure 1: Number of homesteaded and purchased aliquots which were abandoned, resold, and farmed over 12 years, calculated from the regression sample

discrete choice model to capture the type of learning that farmers in our setting experience in three specific ways: (i) updating of Bayesian beliefs, (ii) allowing for farmers to be optimistic or pessimistic, and (iii) unobserved heterogeneity among farmers.

First, we allow farmers' decisions to depend on their changing beliefs about the value of farming in Kansas, which we capture with updating of Bayesian beliefs. Second, we allow the farmers' beliefs to be biased, meaning the farmers may be optimistic or pessimistic about the value of farming. Third, we allow for unobserved heterogeneity among farmers. In our context, this unobserved heterogeneity means that we allow some farmers to have a higher ability to farm than others. Neither the farmers nor the researcher know which farmers have a higher farming ability *ex ante*, but we allow their decisions and beliefs to depend on this heterogeneous unobserved ability. Individuals with different unobserved abilities are allowed to abandon their farm at different rates, leading to dynamic selection.

These three methodological expansions to the typical full-information dynamic discrete choice model - (i) updating of Bayesian beliefs, (ii) allowing for farmers to be optimistic or pessimistic, and (iii) unobserved heterogeneity among farmers - complicate solving the model. This complication arises because a full-information dynamic discrete choice model relies on the assumption of stationarity - i.e., a constant autocorrelation structure over time - to show that the value function is a contraction mapping. Including Bayesian beliefs, as we do, means that this type of stationarity no longer holds. Therefore, we make a methodological contribution to the dynamic discrete choice literature by combining results from game theory

to solve our model. The intuition for solving the model is that we show that the model is stationary in expectation, and that the agents use this expectation when making decisions. Specifically, we relax the stationarity requirement of a typical dynamic discrete choice model to only require stationarity in expectation.

Further, we show that, when the prior mean of the beliefs in two periods are close to each other, the fixed points in these periods may be close to each other as well. We call these fixed points, which arise when similar priors imply fixed points that are numerically close to each other, *sequential fixed points*, and we use them to solve for the unique fixed points of the value function. We propose a numerical optimization routine that exploits this feature. We show that the optimization routine converges at a geometric rate in computer time and that it works well in our application.

We find that farmers were initially highly pessimistic about the value of farming in Kansas. In particular, farmers initially undervalued farming by about \$2,500 in 1870 dollars, which amounts to about about \$50,000 today. However, learning more information over time allowed farmers to refine their beliefs about farming, and by the end of the first year, they were about half as pessimistic about the value of farming. Specifically, after one year, farmers undervalued farming by about \$1,300 in 1870 dollars. Further, while we allow for individuals to value farming differently, we find that in our setting, all farmers had approximately the same valuation. We attribute this to the idea that so much learning was occurring for all types of farmers that previous farming experience in places like Vermont was not particularly relevant.

The model we construct allows us to create several counterfactuals which are of both historical and policy relevance, in order to better understand how mistaken beliefs about farming impacted western expansion. Firstly, we construct a counterfactual of what would have happened to land acquisition if the U.S. government had not allowed homesteading and had only sold the land. Because the Homestead Act was such a far-reaching land policy and because of the invasive nature of western expansion, this counterfactual answers an important question in and of itself. Calculating this type of counterfactual requires some type of discrete choice model, like the one we create. This counterfactual is also important for contract diversity where the volume of a good is increased by having several contracts. We find that the speed of western expansion would have been reduced by 37.5% without the Homestead Act because 69% of the homesteaders would have chosen not to acquire land if

purchasing it had been the only option. This result yields a back-of-the-envelope calculation which indicates agricultural output in Kansas would have decreased by \$270,000,000 per year in 1870 dollars without the Homestead Act. We further compare our estimates to a standard full information model of the demand for land to demonstrate the value of incorporating the agents' beliefs into counterfactual analysis.

Then we construct a second set of counterfactuals to show what fraction of the homesteaders would still have abandoned their farm if they had begun with the information they would have learned after (i) five years and (ii) ten years. These counterfactuals show the effect of changes in the information set on decision-making and have particular policy relevance in understanding how much mistaken beliefs lead to sub-optimal decisions. In our context, abandonment represents a policy failure, and a government policy such as publicizing the annual rainfall in a homesteaded location for the previous ten years may have given the homesteaders more information about the value of farming, leading to lower abandonment. We find that if homesteaders had begun with the information they learned after ten years, 62% of the homesteaders who abandoned their farms would not have done so. Because a quarter of homesteaders abandoned their farms, this counterfactual represents an economically significant change in the number of completed homesteads, and therefore in the value of agricultural output. These counterfactuals would not be possible to estimate in a standard full information model like Rust (1987). However, these counterfactuals are relevant in our setting because of the amount of learning by farmers, so we construct our model to allow for them. They quantify the impact of the farmers' mistaken beliefs on policy failure.

We collect and digitize individual level data in Kansas for this paper. In particular, we use five main individual-level data sources: (i) individual purchase and homestead decisions; (ii) sale deeds of individual resale decisions; (iii) the Kansas agricultural censuses of 1870 and 1880; (iv) the identified U.S. Population Censuses of 1860, 1870, and 1880 for Kansas; and (v) plot-level historical land characteristics. Matching together data on land acquisition and land resale at the plot level shows how long the farmer held the land. The 1870 and 1880 Kansas agricultural censuses record production at the farm level. We match these production data to the plot-level land acquisition data to compare the output between purchasers and homesteaders. We further match the land acquisition and resale data to the identified U.S. Population Censuses of 1860, 1870, and 1880 for Kansas, and to historical land characteristics at the plot level. We use this matching to show that demographic and land characteristics

are stable across land decisions at the individual level.

Addressing the question of how individuals' changing beliefs impact their decisions over time is relevant in a wide range of empirical settings, such as farmers in developing countries learning about different farming methods including fertilizer, students learning about the returns to different majors in college, and firms learning about consumer demand for their products. The model we create can be applied in such settings.

1.1 Related Literature

This paper makes a contribution to several literatures. First, land acquisition, allocation, and property rights is a major focus of development literature. Field, Field, and Torero (2006) and Field (2007) explore the impacts of issuing property titles in Peru. Montero (2022) analyzes the effect of land reform in El Salvador on productivity. Further, there is a large literature on learning in agricultural methods. For example, Duflo, Kremer, and Robinson (2008) analyze why farmers in Kenya do not participate in fertilizer uptake at what appears to be the optimal amount. Our paper complements this work by analyzing a land allocation policy which specifically incentivized learning. The Homestead Act required farmers to actually farm. Our results indicate that, because people can have highly mistaken beliefs and because learning is a gradual process, government policy which involves people learning is more effective when it incentivizes people to continue the process long enough to learn enough to change their beliefs.

Work in economic history on homesteading and land use on the American frontier is related to this development literature. In general, previous economic research on homesteading has been focused on the causal impacts of homesteading on long-term outcomes (Hansen and Libecap (2004), Smith (2020), Allen and Leonard (2021), and Mattheis and Raz (2021)). However, these papers do not estimate how settlers made their initial land acquisition decisions. Our paper is the first step in that direction by modeling how settlers made acquisition decisions in the presence of uncertainty. Because we use a discrete choice model, we are able to make a significant contribution to the literature on land acquisition historically in America by estimating how western expansion would have been different without the Homestead Act.

Further, this paper builds on industrial organization models of demand which include beliefs, including Akerberg (2003), Crawford and Shum (2005), Covert (2015), Darden

(2017), Steck (2018), and Hodgson (2019). However, none of these papers allow for a bias in beliefs, as ours does. Additionally, these papers assume the priors are population densities and are known to the researcher, i.e., the assumption is made that all agents initially think they are average. We do not assume the priors are population densities or known to the researcher and instead estimate them directly. This allows us to let the beliefs be biased. Allowing for biased beliefs is important in many empirical settings.

The rest of the paper is organized as follows. Section 2 lays out the historical background which motivates the model. Section 3 states the model. Section 4 discusses identification and estimation. Section 5 describes the data. Section 6 gives the estimation results and the counterfactuals. Section 7 concludes.

2 Historical Background

Four main facts about the setting motivate our modeling choices. First, Kansas in the late nineteenth century was an area with high uncertainty for farmers. Second, contemporaneous experts disagreed about whether farmers were optimistic or pessimistic about the value of farming. Third, many farmers changed their decision about how to acquire land. Fourth, unusual administrative features of the Homestead Act provided sequential decisions and required farmers to learn about farming. These four facts all indicate that learning new information was relevant to farmers' decision-making process, and we structure our model to be able to capture this learning. We first provide some general historical background, and then we discuss these facts and their impact on our model further below.

Individuals had two main ways to acquire land from the federal government in the late nineteenth century: (i) direct purchase and (ii) homesteading. Farmers could not ensure they acquired better quality land by using one method over the other. The same land was simultaneously available for both methods of land acquisition - by the same individuals - but purchasing land was significantly more expensive than homesteading the same land. Purchasing land cost \$1.25 per acre¹, and individuals received the title to the land immediately. Homesteading land, on the other hand, was significantly less expensive, but it meant that the individual did not receive the title immediately. Instead, homesteaders paid an initial \$10 dollar application fee, and then had three options. Firstly, homesteaders could continue

¹The standard farm size was 160 acres, so purchasing typically cost \$200.

to farm for five years and then receive the title for a \$4 fee. Secondly, after six months, homesteaders could pay to purchase the title more quickly (called commutation). Commutation cost the homestead application fee plus the direct purchase cost of \$1.25 per acre. Thirdly, homesteaders could abandon their farm at any point and not get the title.

We examine two counties in eastern Kansas from 1863 to 1890. Kansas was a heavily agricultural state which was highly impacted by the Homestead Act. The two counties we study were selected because they lie in the eastern part of Kansas, east of the hundredth meridian, which is the dividing line between land suitable for small farms and arid land suitable for ranching and large farms. While individuals were allowed to purchase very large tracts of land, the counties in this paper were selected because this practice was not common in them. Purchased aliquots are not statistically larger than homesteaded tracts in our sample. In our sample, the average farm size is 158 acres.

While there were other methods of acquiring land from the federal government at this time, purchasing and homesteading land were the two which were both most widely used in Kansas and available to nearly everyone, so we limit our analysis to these methods. In fact, between 1863 and 1890, the years in our analysis, more than 80% of the land acquired from the federal government by individuals was acquired through what the Bureau of Land Management classifies as either cash sales or homesteading.

The first fact about the historical setting which influences our model building is that Kansas was an area with high information uncertainty for farmers. Kansas was on the western frontier in the late nineteenth century. When the Homestead Act passed in 1862, Kansas had just recently become a state in 1861. In the decades after the Homestead Act, Kansas received the third most in-migration of any state. Many farmers were highly uncertain about their own farming abilities or preferences because one purpose of the Homestead Act was to act as a safety valve for dissatisfied urban workers (Gates (1968)). Farmers did not have a good understanding of land quality and thus were initially uncertain about the value of farming (Shannon (1966)). Likewise, the types of soils and methods of farming which settlers encountered west of the Mississippi River deviated so greatly from the soil composition they might have been previously familiar with (in the eastern half of the U.S. or in Europe) that historians believe previous farming experience may not have been particularly beneficial (Shannon (1966)).

These facts make Kansas an ideal setting to study the impact of uncertainty on learning

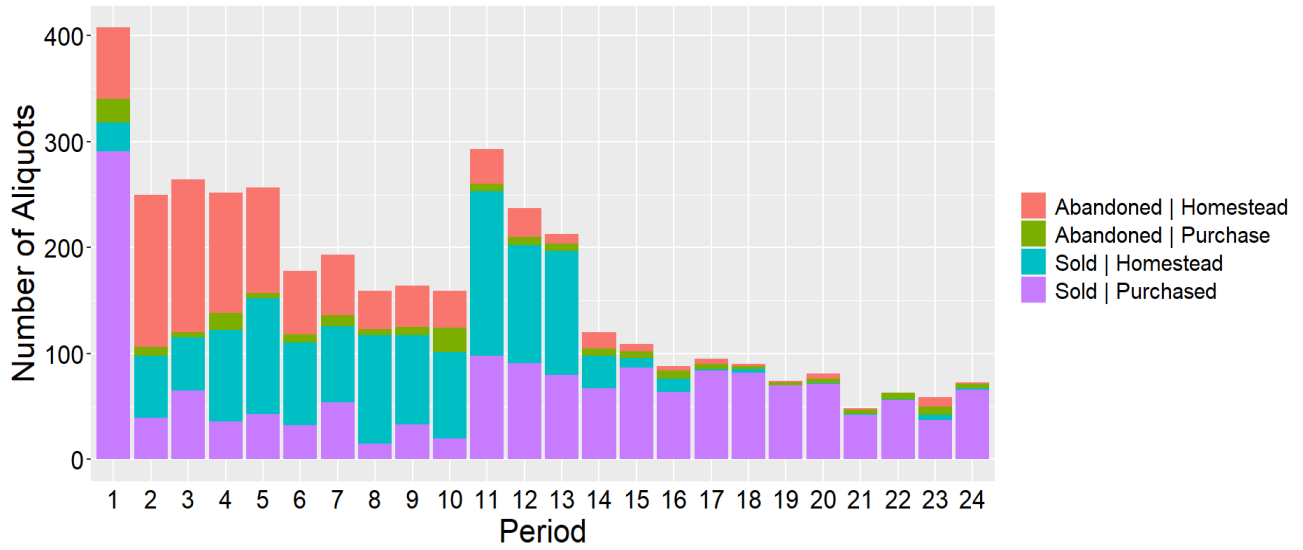


Figure 2: Number of aliquots resold and abandoned by period, calculated from the regression sample

about agricultural ability because the farmers in our sample had recently moved to eastern Kansas and were unfamiliar with farming practices suitable to the area. Therefore, farmers were uncertain about the value of farming, and when they made decisions about whether or not to continue farming in Kansas, it was based on their belief about the value of farming, which may have been highly incorrect. We allow for these mistaken beliefs about the value of farming in our model.

The second fact which influences our model building is that contemporaneous experts disagreed about the direction of the information uncertainty for farmers. Some historical sources indicate that farmers were too optimistic about their farming prospects, while others indicate that farmers were too pessimistic. For example, some experts cautioned against starting a farm without at least \$1,000 and stated that the vast majority of farmers were starting with much less capital than \$1,000, meaning that these farmers may have been too optimistic (Peck (1967)). On the other hand, contemporaneous newspapers wrote that “The record-making general drought of 1860 left an unfortunate aftermath by creating a haunting doubt in the minds of many people ... that subsequent drought periods might lengthen interminably into another [drought]” (Malin (1946)). This evidence indicates that farmers may have been too pessimistic about their farming prospects. We allow for both optimism and pessimism of farmers in our model in order to test these statements empirically.

The third fact which influences our model building is that farmers valued the option to change their decision highly. The high uncertainty about farming practices is exemplified by the fraction of farmers who chose to abandon or sell their farms, shown in Figures 1 and 2. Notably, one-quarter of homesteaders chose to abandon their farms, and another one-third of homesteaders chose to sell their farms.² Together, almost 60% of homesteaders changed their decision about how to acquire land.

Figure 2 shows the number of periods (6 month time intervals) which farmers had held their land before abandoning or selling it.³ Periods are calculated from when the farmer begins farming, so period 1 includes everyone who had acquired land in the last six months, regardless of the calendar date. Specifically, Figure 2 illustrates the high percentage of farmers who abandoned their farm within the first two years. Abandonment took place after the farmer had already either paid the initial application fee for the homestead or the purchase price. Further, the cost of starting a farm in the Midwest in this period is generally estimated at approximately \$1,000 in 1870 dollars (approximately \$21,000 today) (Peck (1967)). Therefore, even if farmers who abandoned their farms had invested less money than average into their farms, these farmers had already invested a significant amount of money and time when they abandoned. Thus, the fact that 24% of homesteaders in our data abandoned their farm indicates that these farmers learned new information which significantly changed their belief about farming, and they valued the option to change their decision highly.

If abandoning were due to simply random negative shocks like droughts, the probability of abandoning would be more evenly distributed across the amount of time farmers had been on their land. The probability would be evenly distributed in such a case because farmers in our data started in different years, so a potential drought would hit them in different years. Instead, we see that the probability of abandonment is highly concentrated in the first two years after starting the farm, regardless of the calendar year. The pattern we observe is relative to how long the farmer had been farming, not the calendar year. This concentration in the probability of changing their decision when new information was of the

²Homesteaders generally commuted their farm just before reselling it, so we group these choices together in our data.

³Land that was originally purchased from the federal government may have been abandoned, similar to how homesteads were abandoned. However, this information was never recorded. Therefore, we define abandoned land that was originally purchased to be purchased land that was resold for less than 1 dollar per acre. This cutoff ensures that the farmer was selling the land at a loss because the initial price per acre they paid to purchase land from the federal government was 1.25 dollars per acre.

most value indicates learning. Further, Table 8 in the online appendix shows that the rate of abandonment is similar for all quarters of the year, indicating abandonment is not driven by cyclical weather shocks.

Figures 1 and 2 also show the percentage of farmers who sold their farms. In the case of homesteaders who sold before period 10, in order to sell, they commuted their farms by paying \$1.25 per acre after already paying the initial homestead application fee of \$10. The decision to commute a homestead was strictly more expensive than purchasing the same land. However, we observe about one-third of homesteaders commuted their farms in our data. This high percentage of homesteaders who commuted their farm is consistent with the idea that these farmers learned new information which caused them to change their decision. Because we observe so many farmers change their decisions after starting to farm, our model allows for learning new information to change farmers' beliefs and therefore their decisions.

The fourth element of the historical setting which influences our model building is that the Homestead Act was structured sequentially to require individuals to learn more about farming. This requirement of the Homestead Act makes it an ideal setting to study how learning new information impacts decisions because the administrative features of the Act included both learning and sequential decisions. Compared to purchasing the same land, the Homestead Act provided a low-cost way to learn more about farming before committing to the title because homesteaders were required to actually farm the land. Therefore, they learned about their ability to farm. Further, they used this updated information to decide whether to (i) acquire the title of the farm by farming for five years, (ii) abandon the farm, or (iii) sell the farm.

It is notable that the three options available to homesteaders - continuing to farm, commuting and/or selling the farm, or abandoning the farm - were only available after already paying the initial homestead application fee and starting the farm. For example, homesteaders could not get the title in any manner for at least six months. This element of the Homestead Act meant that homesteaders were able - and, in fact, required - to learn more information on which to base their new decision. We include the sequential nature of the homesteaders' decision-making process in our model.

3 Model

3.1 Bayesian Model with Heterogeneous Types

A major component of our setting is that we observe farmers changing their decisions when they learn new information. In order to capture this learning over time and its impact on decision-making, we use a dynamic discrete choice model. However, a standard dynamic discrete choice model assumes that agents have full information, which would not adequately capture farmers' information uncertainty in our setting. Therefore, we expand a typical full information dynamic discrete choice model, such as Rust (1987), in three ways to explicitly model the learning process we see reflected in our data.

Firstly, we expand the typical dynamic discrete choice model to allow for farmers to have a belief about the value of farming in Kansas and to base their decisions on this belief, rather than on the true value. We allow this belief to change over time as farmers learn new information. Specifically, we model this changing belief using Bayesian updating. This belief is important to include in our model because of the high information uncertainty experienced by farmers in our setting. This modeling decision captures the fact that farmers learn over time and base their decisions on that new information.

Secondly, we expand the typical full information dynamic discrete choice model to allow the farmers' beliefs to be biased, meaning that farmers can be optimistic or pessimistic about how valuable farming will be. This optimism or pessimism is important to allow for in our model because some contemporaneous experts believed that farmers were too optimistic, while some reported that farmers were too pessimistic, and we allow for either option.

Finally, we expand the typical full information dynamic discrete choice model by allowing for unobserved heterogeneity among farmers. In our setting, this means that we allow some farmers to have a higher ability to farm than others, and that this higher ability is unobserved by both the researcher and the farmers themselves. We call this unobserved heterogeneity the farmer's type. This unobserved heterogeneity among farmers is important to allow for in our model because there may have been selection on unobservables into different land acquisition decisions. This modeling decision reflects the fact that different farmers may have different backgrounds which position them to be more successful at farming, and that some of this background is unobservable.

Expanding a dynamic discrete choice model in the above three ways complicates solving

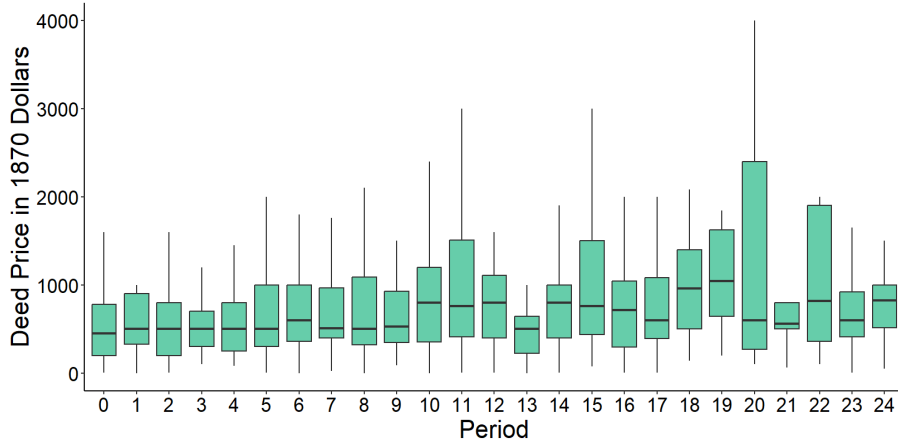


Figure 3: Resale price by period shows that stationarity holds, calculated from the regression sample

the model. Specifically, this complication arises because a typical full information dynamic discrete choice model relies on the assumption of stationarity - i.e., a constant autocorrelation structure over time - to show that the value function is a contraction mapping. Adding Bayesian beliefs which are updated over time to such a model, as we do, means that this type of stationarity no longer holds. Intuitively, farmers do not know what they may learn in the future, but they need to make a decision now, and this decision must incorporate the possible information from the future. Therefore, in order to show that the value function is a contraction mapping, we combine techniques from the game theory literature to calculate the distribution of random variables and outcomes to solve the optimization problem.

Specifically, in order to solve our model, we empirically implement the single-agent version of the Perfect Bayesian Equilibrium by Fudenberg and Tirole (1991). Watson (2017) states conditions on updating beliefs and these are satisfied in our model. Kamenica and Gentzkow (2011) introduce the concept of Bayesian Persuasion and make the assumption of Bayesian plausibility in the context of their game theory model. They state that a posterior is Bayesian plausible if the expected posterior probability equals the prior. A feature of our model is that this Bayesian plausibility is satisfied, meaning that we can use the results from Kamenica and Gentzkow (2011) about which priors agents use. These results mean that stationarity holds in expectation, and that it is this expectation which farmers use in making decisions.⁴

In order for stationarity in expectation to hold in our context, the price of developed

⁴See the appendix and the online appendix.

farmland cannot change in a predictable manner. In our application, this assumption holds because the price of developed farmland in Kansas in the late 1800's is stable. For stationarity in expectation to be violated, the price of farmland would have to change in a predictable manner, which it does not. We find that the correlation coefficient between resale price and periods held is low (0.0894). Further, Figure 3 shows that the resale price of developed farmland does not change significantly no matter how many periods it is held. While we typically think of land values as increasing over time, in our setting, they do not because Kansas was on the frontier, and there was a large amount of inexpensive land available. Land prices remaining stable over time is stronger than the necessary stationarity in expectation condition. Because stationarity holds in expectation in our model, the value function is a contraction mapping. This result is proved in the appendix. Therefore, our methodology allows for non-stationarity by using a conditional expectation version of a fixed point theorem, which means we can show the value function is a contraction mapping. Such non-stationarity can also be relevant outside a Bayesian framework and allows for future priors to be known in expectation.

A few previous papers have combined a dynamic discrete choice model with Bayesian beliefs, including Akerberg (2003), Crawford and Shum (2005), Covert (2015), Darden (2017), Steck (2018), and Hodgson (2019). However, our paper makes a methodological contribution to this literature in two ways. First, none of the previous models have allowed for the beliefs to be biased, which we do. In our context, it is necessary to allow the beliefs to be biased because contemporaneous experts disagreed about whether farmers were highly optimistic or highly pessimistic. Without a bias in beliefs, we could not capture this optimism or pessimism which impacted farmers' decisions.

Second, previous literature assumes that the prior beliefs are known to the researcher. Specifically, the researcher makes the assumption that the agents' prior belief is the population density, i.e., everyone starts by believing they are average. We do not assume the priors are known to the researcher, but instead estimate them directly in the model. We could not allow for the beliefs to be biased without estimating them directly.

Further, we estimate the informational content of farming experience which farmers use to update their beliefs, meaning we estimate both the beliefs and the information used to update those beliefs. The beliefs are the unobserved state variables, and we call the information farmers use to update their beliefs *information realizations*. These information realizations

differ from the structural disturbances by Pakes, Porter, Ho, and Ishii (2015), which were developed outside a model with Bayesian updating. Estimating these information realizations helps us to understand historical farming in the Midwest by clarifying the conditions under which farmers made decisions. Such estimation is also useful in other dynamic programming models with nested fixed points, as well as models of stationary games, to clarify decision conditions in those settings.

3.1.1 Model Specification

Farmers' decisions are based on their beliefs in our model, not on the true values, which reflects the real decision-making process. Individuals learn about their own ability to farm, and we use their choices in each period to back out their learning process. First, we outline the general form of the model, and then we detail how we incorporate (i) information realizations which the farmers use to update their beliefs, (ii) biased Bayesian beliefs, and (iii) which prior beliefs to use. Finally, we explain the model implementation in our setting.

In every period, farmer i has the option to either (i) continue to farm; (ii) sell the farm; or (iii) abandon the farm, based on the new information they gain in every period. If they choose to sell or abandon, they do not continue to the next period; they only continue to period $t + 1$ if they choose to continue to farm in period t . Let this decision and these choices be denoted by $d_{it} \in \{C, S, A\}$, where C indicates continuing, S indicates selling, and A indicates abandoning. We assume that farmers choose optimally, based on their beliefs, in order to maximize the discounted expected utility at time t . Each farmer belongs to an unobserved type, and we assume that all farmers of the same type are identical. Let farmer i be of type $m \in \{A, B\}$.

The farmer does not observe their true ability, μ_m , to farm in Kansas. This ability can also be thought of as a farmer's true value of farming in Kansas. Instead of observing their true ability, the farmer has a prior over their ability to farm. Call the mean of this prior after observing t periods of data $\bar{\pi}_{it}$, where farmer i is of type m . Let $title_{it}$ be a binary variable indicating whether or not the farmer has the title in period t , and let $periods_{it}$ indicates how many periods are left before acquiring the title if $title_{it} = 0$. Let $z_{it} = \{periods_{it}, \bar{\pi}_{it}, X_t\}$, where X_t are the information realizations the farmers learn each period. Then $\{title_{it}, z_{it}\}$

represent the state variables for all i and t . The utility of individual i in period t is given by

$$u(\text{title}_{it}, z_{it}, d_{it}) = \begin{cases} \bar{\pi}_{it} + \varepsilon_{C,it} & \text{if } d_{it} = C \text{ and individual } i \text{ is type } m \\ \text{Sell}_{it} + \varepsilon_{S,it} & \text{if } d_{it} = S \\ \varepsilon_{A,it} & \text{if } d_{it} = A, \end{cases}$$

where Sell_{it} is the utility of individual i from selling the farm in period t . The variable Sell_{it} and the error terms $\varepsilon_{C,it}$, $\varepsilon_{S,it}$ and $\varepsilon_{A,it}$ are all unobserved to the empirical researcher and independently distributed from each other for all i and t .

In a full information model like Rust (1987), the agent knows the parameter values. In our case, individuals have beliefs about the values according to their type and subsequently update these beliefs. This type is unobserved to the empirical researcher. The farmers update their own prior beliefs but do not know how many types exist, meaning they cannot convey type-specific information to each other. The prior mean $\bar{\pi}_{it}$ is based on the information set that the farmer has at time t . All farmers that are of type m use the same prior $\pi_{mt}(\mu)$. In our case, a sufficient statistic for the beliefs is the prior mean of μ_m . If farmer i is type m then their prior mean of μ_m is

$$\bar{\pi}_{mt} = \int \mu_m \pi_{mt}(\mu_m) d\mu_m. \quad (1)$$

Allowing the farmers to base their decisions on their priors implies that the value functions and Bellman equation will depend on these priors as well. This motivates the central innovation of our paper. To allow for the value functions and Bellman equation to depend on priors, we write $u(\text{title}_{it}, z_{it}, d_{it})$ as the addition of two parts: the part of the utility function that is observable to the farmer, and the part of the utility function that depends on the prior. The reason for writing the utility in these two parts is that the decisions of the farmer depend on the (updated) prior rather than the farm output. This gives

$$u(\text{title}_{it}, z_{it}, d_{it}) = u_{\text{farm obs}}(\text{title}_{it}, z_{it}, d_{it}) + E\{u_{\text{farm unobs}}(\text{title}_{it}, z_{it}, d_{it})\},$$

where the expectation is taken with respect to the prior, and

$$u_{\text{farm obs}}(\text{title}_{it}, z_{it}, d_{it}) = \begin{cases} \varepsilon_{C,it} & \text{if } d_{it} = C \\ \text{Sell}_{it} + \varepsilon_{S,it} & \text{if } d_{it} = S \\ \varepsilon_{A,it} & \text{if } d_{it} = A, \end{cases}$$

and

$$u_{\text{farm unobs}}(\text{title}_{it}, z_{it}, d_{it}) = \begin{cases} \mu_m & \text{if } d_{it} = C \text{ and individual } i \text{ is type } m \\ 0 & \text{if } d_{it} = S \\ 0 & \text{if } d_{it} = A. \end{cases}.$$

These utility functions imply the following value function,

$$V(\text{title}_{it}, z_{it}) = \sup_{\Lambda_{it}} \{u_{\text{farm obs}}(\text{title}_{it}, z_{it}, d_{it}) + E\{u_{\text{farm unobs}}(\text{title}_{it}, z_{it}, d_{it})\} + E[\sum_{s=t+1}^{\infty} \delta^{s-t} u(\text{title}_{is}, z_{is}, d_{is})]\},$$

where δ is the discount factor and Λ_{it} denotes the set of decision rules for each individual i at time t . These decision rules are functions of previous decisions, beliefs, and state variables. For those farmers that have the title, we can write the last equation in a recursive form

$$V(\text{title}_{it} = 1, z_{it}) = E \left[\sup_{d_{it} \in \{C, S, A\}} \{u_{\text{farm obs}}(\text{title}_{it} = 1, z_{it}, d_{it}) + u_{\text{farm unobs}}(\text{title}_{it} = 1, z_{it}, d_{it}) + \delta E[V(\text{title}_{i,t+1} = 1, z_{i,t+1})]\} \right], \quad (2)$$

where the state variables, $z_{i,t+1}$ and $\text{title}_{i,t+1}$, are functions of the state variables in the previous period, z_{it} and title_{it} , and the decision d_{it} .

All farmers face a non-stationary environment in the sense that they update their priors every period. As we argued above, the farmers do not yet know their future priors or posteriors. As we will show in Lemma 1, Bayesian updating implies that individuals will predict the future value function using their current priors. This allows for non-stationarity at the individual level for all farmers. This result allows us to solve the model. The Bayesian

updating yields that the predicted prior mean in future periods is the current prior mean, i.e., $E(\bar{\pi}_{i,t+s}) = \bar{\pi}_{it}$ for all $s \geq 0$. Further, we assume that the expected utility from selling the farm remains constant, i.e., $E(\text{Sell}_{i,t+s}) = \text{Sell}_{it}$ for all $s \geq 0$. In our application, this assumption holds because the price of developed farmland in Kansas in the late 1800's is stable (see Figure 3).

This property of Bayesian updating and the assumption on the expected utility of selling the land yields that $E\{V(\text{title}_{i,t+1} = 1, z_{i,t+1})\} = V(\text{title}_{it} = 1, z_{it})$, meaning the expected value in the next period equals the value in this period, for farmers with the title. This equality allows us to write the value as a function of the state variables and itself,

$$V(\text{title}_{it} = 1, z_{it}) = E \left[\sup_{d_{it} \in \{C, S, A\}} \{u_{\text{farm obs}}(\text{title}_{it} = 1, z_{it}, d_{it}) + u_{\text{farm unobs}}(\text{title}_{it} = 1, z_{it}, d_{it}) + \delta V(\text{title}_{it} = 1, z_{it})\} \right]. \quad (3)$$

For homesteaders who do not have the title yet, we do not have the equality $E\{V(\text{title}_{i,t+1} = 1, z_{i,t+1})\} = V(\text{title}_{it} = 1, z_{it})$. Instead, we use the priors to predict the future value functions and then use backward induction from the period that they acquire the title.

Now we turn to the specifics of our model, including the information realizations, biased Bayesian beliefs, and which priors to use.

3.1.2 Information Realizations

A farmer receives a signal each period of the success of their farm and their own ability and uses this information to make a continuation decision. All farmers learn the same information: X_t is not farmer specific. For example, in period 1, farmers learn about how easy it is to buy farm equipment on credit. This allows us to better understand how farmers made decisions under uncertainty. The farmers observe the information realizations X_t , $t = 1, 2, \dots$, and use these to update their beliefs. The empirical researchers do not observe X_t , $t = 1, 2, \dots$, but can measure the prior belief in every period. This implies that the changes in the prior from period to period can identify the information realizations X_t for every period. Estimating X_t allows the empirical researcher to understand the magnitude of the factors influencing decision-makers. We estimate X_t , $t = 1, 2, \dots$, directly as parameters in the model, up to the normalizations of zero mean and variance $\sigma_{data}^2 = 1$.

3.1.3 Allowing for Biased Beliefs

Some individuals are better suited to farming than others and therefore receive more benefits from farming. Let $\mu_m, m \in \{A, B\}$ be the type-specific farming ability. Individuals learn about μ_m each period by Bayesian updating. If they learn their farming ability is low, they may decide to abandon their farm. Different types may abandon at different rates, leading to dynamic selection. Updating their beliefs about their own farming ability allows individuals to make more informed decisions over time.

We assume beliefs are normally distributed. Individuals may exhibit optimism or pessimism, meaning their normally distributed posterior means are allowed to be lower or higher than their true abilities. This bias shifts the distribution of farmers' beliefs away from the distribution of true farming ability.

Before starting to farm, let the belief of type m be normally distributed, centered around $\bar{\mu}_m + \eta_m$ (where η_m is a type-specific bias term), and have variation σ_{prior}^2 . Farmers observe information X_t every period. Therefore, after observing t periods of data, the mean of the prior is

$$\bar{\pi}_{it} = \bar{\mu}_m + \frac{\eta_m + \frac{\sigma_{prior}^2}{\sigma_{data}^2} \cdot \sum_{s=1}^t X_s}{1 + t \cdot \frac{\sigma_{prior}^2}{\sigma_{data}^2}},$$

and the variance of the prior after t periods is

$$\sigma_{\pi,t}^2 = \frac{1}{\frac{1}{\sigma_{prior}^2} + \frac{t}{\sigma_{data}^2}}.$$

Note that the effect of the bias η_m decreases over time. The belief in each period combines the prior with the new information in a Bayesian framework such that as t increases, the farmer puts more weight on the information and less weight on the prior.

3.1.4 Which Priors to Use

A central issue in a model with beliefs is which beliefs agents use when making decisions. Agents optimize with respect to their beliefs, so which beliefs to use is important. We prove that farmers use their current beliefs in making decisions. Kamenica and Genzktow (2011)

call this result Bayesian plausibility. We show that Bayesian plausibility holds in our model, so we can use their results to solve our model.

The farmers have beliefs about their ability to farm, and they know they will learn more about that ability in the future. When solving their dynamic optimization problem at time t , the farmers use the prior belief that they have at time t . The intuitive reason for this is that the farmers do not have future information to update their priors.

Consider farmers at $t = 0$. In later periods, the farmers will observe the information realizations X_t for $t = 1, 2, \dots$. The farmers know that these realizations, given their ability μ , are normally distributed around zero with variance σ_{data}^2 . The key to determining which prior belief to use is the prior predictive distribution. The prior predictive distribution is the distribution of X_t for $t = 1, 2, \dots$ without conditioning on μ . That is, we average the conditional distribution of $X_t|\mu$ using the prior on μ as a weight. Thus, this average is given by $p(x_t) = \int p(x_t|\mu)\pi_{t=0}(\mu)d\mu$, where $p(x_t|\mu)$ is the conditional distribution and $\pi_{t=0}(\mu)$ is the prior at $t = 0$.

When a farmer observes X_1 , then this farmer can update their prior. Let this posterior be denoted by $\pi(\mu|X_1)$. Farmers do this updating in our model. Alternatively, the farmers could update their prior on all possible values of X_1 and then use the prior predictive distribution as a weight, i.e., $\pi_{t=1}(\mu) = \int \pi(\mu|x_1)p(x_1)dx_1$, where $\pi_{t=1}(\mu)$ is the prior at $t = 1$, without conditioning on the outcome X_1 . The following lemma states that this would not give the farmers additional information: $\pi_{t=1}(\mu)$ is the same as the original prior, and, further, this holds for all future periods. Therefore, farmers use the current prior when optimizing, even though they know they will learn more in the future.

Lemma 1: Priors

Let $\pi_{t=0}(\mu)$ denote the prior on μ at $t = 0$ and let $\pi_{t=0}(\mu)$ be the density of a normally distributed random variable with mean $\bar{\mu} + \eta$ and variance σ_{prior}^2 . The random variable X_t is observed at time t and let $X_t|\mu$ be i.i.d. over time and be normally distributed with mean μ and variance σ_{data}^2 . Let $\pi_{t=s}(\mu)$, $s = 0, 1, 2, \dots$, denote the expected prior at time s . Then $\pi_{t=s}(\mu) = \pi_{t=0}(\mu)$ for all s .

Proof: See appendix

After observing X_1 , the farmers update their prior and can use this prior $\pi(\mu|X_1)$ to predict future outcomes and expected priors. Lemma 1 also applies to such a conditional

prior, and these expected priors are the same as $\pi(\mu|X_1)$. Further, the same holds for the priors that condition on X_1, \dots, X_Q for any $Q \geq 1$.

Kamenica and Gentzkow (2011) introduce the concept of Bayesian persuasion and make the assumption of Bayesian plausibility. They state that a distribution of posteriors is Bayes plausible if the expected posterior probability equals the prior. The last lemma states that Bayesian plausibility holds in our model.

3.1.5 Model Implementation

Now we turn to the implementation of our model. Individuals in our model move through the following periods, which are each 6 months long. In period $t = 0$, individual i decides to either homestead or buy land. For most of the paper, we take the decision to acquire land as fixed, making our model robust to misspecifying this decision. We use 20 periods to estimate the model, but it can be estimated with infinite periods. In this setting, we assume that we have high and low ability farmers, but this model can be implemented with more than two types. For the high ability farmer, type A , we have μ_A , and for the low ability farmer, type B , we have μ_B .

We first consider the decision process of the homesteaders. The homesteaders acquire the title after five years and then have the same value functions and fixed points as the purchasers. In particular, let $V_{home,mt}$ denote the value function of the homesteader of type m in period t . From period 10 onward we have $V_{home,mt} = V_{mt}$. First, we outline the backward induction process for periods 1 through 9, and then we discuss these fixed points further below.

Let $Sell_{notitle}$ be the utility a homesteader receives from selling their farm before they would normally acquire the title. For a homesteader to receive the title before period 10, they have to pay \$200. Therefore, we can consider $Sell_{notitle}$ to be the utility from selling their farm after they have had to pay an additional \$200 for the title. Let κ denote the parameter that converts dollars into utility. The estimation of $Sell_{notitle}$ and κ are discussed in Section 4. As above, $d_{it} \in \{C, S, A\}$ for $t \geq 1$ and, in addition, let $d_{i,0} \in \{purchase, home\}$ for $t = 0$. We further define

$$Sell_{it} = \begin{cases} Sell_{notitle} + \kappa \cdot 200 & \text{if } d_{i,t \geq 10} = S \text{ and } d_{i,0} = home \\ Sell_{notitle} & \text{if } d_{i,t < 10} = S \text{ and } d_{i,0} = home. \end{cases}$$

Homesteads Use Backward Induction in Periods 1 through 9. Before acquiring the title, the homesteaders use backward induction from their estimates of the fixed points. Thus, for these earlier periods we have

$$V_{home,mt} = E \left[\sup_{d_{it} \in \{C,S,A\}} [1(d_{it} = C) \cdot (\varepsilon_{C,it} + \bar{\pi}_{mt} + \delta V_{home,m,t+1}) + 1(d_{it} = S) \cdot (\varepsilon_{S,it} + Sell_{it}) + 1(d_{it} = A) \cdot \varepsilon_{A,it}] \right], \text{ where } t = 1, \dots, 9.$$

In our application, we assume that the random shocks $\varepsilon_{C,it}$, $\varepsilon_{S,it}$, and $\varepsilon_{A,it}$ given $\bar{\pi}_{mt}$ and X_t are realizations from a type I extreme value distribution⁵ and are independent of each other, given $\bar{\pi}_{mt}$ and X_t . This yields a specific form for the value functions that we state in the online appendix. Further, this assumption on the random shocks and the value functions implies the following probabilities of continuing for each type for homesteaders,

$$\Pr_{m,home}(\text{continue at } t) = \frac{\exp(\bar{\pi}_{mt} + \delta V_{home,m,t+1})}{1 + \exp(\bar{\pi}_{mt} + \delta V_{home,m,t+1}) + \exp(Sell_{notitle} + 1(t \geq 10)\kappa \cdot 200)}$$

for $t = 1, 2, \dots$

The homesteader has the title from period 10 onward, so the homesteader no longer has to pay the \$200 fee to the government for selling the land. The indicator $1(t \geq 10)$, the utility of money κ , and the amount \$200 account for this in the last expression. The related probabilities for selling and abandoning are in the online appendix.

Homesteaders Use Fixed Points in Periods 10 through 20. Consider farmer i at time t , who homesteaded their farm, and let this farmer be of type m . If this farmer continues to farm, then the farmer expects to receive the relevant part of $u_{\text{farm obs}}(\text{title}_{it}, z_{it}, d_{it})$, i.e., $\varepsilon_{C,it}$; the prior mean of μ_m , i.e., $\bar{\pi}_{mt}$; and the discounted expected value function, i.e., $\delta E\{V(\text{title}_{i,t+1}, z_{i,t+1})\}$. Thus, the reward for continuing to farm is $\varepsilon_{C,it} + \bar{\pi}_{mt} + \delta E\{V(\text{title}_{i,t+1}, z_{i,t+1})\}$. It is important to note that the farmer's decision does not depend on the true value of μ_m but on their beliefs, i.e., the prior mean of μ_m . This modeling decision makes the model more realistic and, further, computationally faster, because the prior mean has only to be calculated once for every fixed point calculation. If the farmer sells or abandons the

⁵The cumulative distribution function of this extreme value distribution is $F(\varepsilon) = \exp(-\exp(-\varepsilon))$.

farm, they receive $Sell_{it} + \varepsilon_{S,it}$, or $\varepsilon_{A,it}$, respectively.

Once the homesteader has the title, their value function is a sequence of fixed points. Specifically, earlier, we showed that $E\{V(title_{i,t+1} = 1, z_{i,t+1})\} = V(title_{it} = 1, z_{it})$, meaning the expected value in the next period equals the value in this period, for farmers with the title. Further, let $V_{mt}(V)$ denote the value function of a farmer that has the title to the farm and is of type m . Note that the prior mean and information realizations are the same for all i of the same type. The error terms $\varepsilon_{C,it}$, $\varepsilon_{S,it}$, and $\varepsilon_{A,it}$ are farmer specific, but we can calculate the value function $V_{mt}(V)$ by taking the expectation of these individual error terms. Using equation (3), we can write $V_{mt}(V)$ as a function of these state variables and the value V ,

$$V = E \left[\sup_{d_{it} \in \{C,S,A\}} [1(d_{it} = C) \cdot (\varepsilon_{C,it} + \bar{\pi}_{mt} + \delta V) + 1(d_{it} = S) \cdot (\varepsilon_{S,it} + Sell_{it}) + 1(d_{it} = A) \cdot \varepsilon_{A,it}] \right]. \quad (4)$$

This is a contraction mapping, as shown in the appendix, because we show that stationarity holds in expectation in our model. Let $V_{mt,\text{fixed}}$ denote a fixed point for type m in period t . The lemma below gives the conditions for the uniqueness of this fixed point. The prior mean changes in every period, so the fixed point $V_{mt,\text{fixed}}$ needs to be calculated for every period. Note that when the prior means are very similar in adjacent periods, the fixed points $V_{mt,\text{fixed}}$ and $V_{m,t+1,\text{fixed}}$ may be close to each other as well. We call such a set of fixed points that are pairwise close to each other *sequential fixed points*, and we now propose an algorithm that works well in such a case and is consistent in a much larger class.

Suppose the empirical researcher estimates the fixed point for type m in period t , i.e., $V_{mt,\text{fixed}}$. The prior means of period t and period $t-1$ may be close together, so that $V_{mt,\text{fixed}}$ may be a good starting value for calculating $V_{m,t-1,\text{fixed}}$. In particular, consider the sequence

$$\begin{aligned} V_{m,t-1,R=1} &= V_{m,t-1}(V_{mt,\text{fixed}}), \\ V_{m,t-1,R=2} &= V_{m,t-1}(V_{m,t-1,R=1}), \\ &\dots \\ V_{m,t-1,R=r} &= V_{m,t-1}(V_{m,t-1,R=r-1}) \text{ for } r > 1. \end{aligned}$$

After the researcher finds $V_{m,t-1,\text{fixed}}$, they then use the same method to find $V_{m,t-2,\text{fixed}}$, etc.. The following lemma gives the conditions under which this algorithm converges at a geometric rate in r .

Lemma 2: Contraction Mapping and Sequential Fixed Point Algorithm

Let

$$V_{mt}(V) = E \left[\sup_{d_{it} \in \{C,S,A\}} [1(d_{it} = C) \cdot (\varepsilon_{C,it} + \bar{\pi}_{mt} + \delta V) + 1(d_{it} = S) \cdot (\varepsilon_{S,it} + Sell_{it}) + 1(d_{it} = A) \cdot \varepsilon_{A,it}] \right]. \quad (5)$$

Let (i) the conditional expectations in the last expression exist for all m and t ; (ii) $0 \leq \delta < 1$; and (iii) $V \in [0, V_{high}]$ where $0 < V_{high} < \infty$. Then $V_{mt}(V)$ is a contraction mapping with a unique fixed point $V_{mt,\text{fixed}}$ for all m and t . Further, the sequential fixed point algorithm converges at a geometric rate in r for all m and t , i.e.,

$$|V_{mt,R=r} - V_{mt,\text{fixed}}| \leq \delta^r |V_{mt,R=1} - V_{mt,\text{fixed}}| \text{ where } r > 1.$$

Proof: See appendix

By using the fixed point of an adjacent period as a starting value, the sequential fixed point algorithm is fast in computer time in our application. The geometric rate of convergence in Lemma 2 is the same as for the case of full information. However, the Bayesian updating implies a fixed point for every period and type, meaning that the fast convergence is relatively important.

An important assumption of Banach's (1922) fixed point theorem is the boundedness of V , i.e., $V_{high} < \infty$. This assumption is satisfied in our application because farms sold, on average, for \$1,480 when homesteaders sold, and \$1,475 when purchasers sold. Further, since the government was giving undeveloped land away through homesteading, the boundedness of the value of farming seems reasonable. Thus, the main conditions of the Banach (1922) fixed point are satisfied, and we discuss in the appendix that the other ones are satisfied as well. Therefore, we have a unique solution for $V_{mt,\text{fixed}}$ for every period.

3.1.6 Dynamic Selection

Farmers have prior beliefs about their ability to farm and update these priors over time. These (updated) beliefs may affect their decision to sell or abandon the farm. In our application, we find that the productivity of farmers is not homogeneous and that the high productivity farmer is more likely to continue to farm. In our model, we allow for different rates of abandoning, selling, and continuing by ability, i.e., type-specific survival functions. In industrial organization, such selection is called competitive selection (see for example Bowden (1992)), and in labor economics, it is called dynamic selection (see for example Van den Berg (2001) and Ridder and Woutersen (2003)). Let $H_{m,home,t}$ denote the type-specific survival function for homesteaders, i.e., the probability of continuing through at least time t for homesteaders of the type. That is,

$$H_{m,home,t} = \prod_{s=1}^t \Pr(d_{ms} = C), \quad m \in \{A, B\}.$$

We allow for dynamic selection by first calculating the fraction of the homesteaders who are high ability at $t = 0$. Let $f_{A,home,t=0}$ denote this fraction,

$$f_{A,home,t=0} = \frac{\Pr(A|HS \text{ at } t = 0)}{\Pr(A|HS \text{ at } t = 0) + \Pr(B|HS \text{ at } t = 0)}.$$

Next, we can use the survival functions that condition on type to calculate this fraction for any t ,

$$f_{A,home,t} = \frac{\Pr(A|HS \text{ at } t = 0) \times H_{A,home,t}}{\Pr(A|HS \text{ at } t = 0) \times H_{A,home,t} + \Pr(B|HS \text{ at } t = 0) \times H_{B,home,t}}, \quad t = 1, 2, \dots$$

The fraction $f_{A,home,t}$ shows the dynamic selection process over time because $f_{A,home,t}$ is the fraction of homesteaders that are still farming at time t and are type A . We would expect high ability farmers (type A) to be more likely to continue relative to the low ability farmers (type B), so $f_{A,home,t}$ should increase towards 1 as t increases when $\Pr(A|HS \text{ at } t = 0) > 0$, $\Pr(B|HS \text{ at } t = 0) > 0$. We use $f_{A,home,t}$ to calculate the probability of continuing in each period t , which demonstrates how farmers make different decisions under different information sets. Such dynamic selection based on unobservables also occurs in duration models.

Perhaps the easiest way to deal with the dynamic selection is to calculate the survival function that conditions on homesteading but not on type. Let $H_{home,t}$ denote this survival function. This survival function is a weighted average of the survival functions $H_{A,home,t}$ and $H_{B,home,t}$ that do condition on type,

$$H_{home,t} = H_{A,home,t} \times \Pr(A|HS \text{ at } t = 0) + H_{B,home,t} \times \Pr(B|HS \text{ at } t = 0).$$

Note that we can use the data to estimate this survival function because we observe which farmers continue to farm and for how long. Further, we can estimate the probabilities of selling or abandoning at time t , conditional on farming up to that point. Therefore, the probability of selling the farm in period t , conditional on being type m , equals the probability of continuing to farm in every previous period (including $t - 1$) multiplied by the probability of selling the farm in period t , conditional on being type m . That is,

$$\Pr(d_{mt} = S, d_{mk} = C \text{ for } k = 1, \dots, t - 1) = \Pr(d_{mt} = S) \prod_{s=1}^{t-1} \Pr(d_{ms} = C).$$

We can calculate this sequence for every type and then take the weighted average to get the unconditional probability. This unconditional probability gives us the survival function for homesteaders. We also calculate the survival function and probabilities for the purchasers, as we condition on observables wherever possible.

We have now described the Bayesian model for homesteaders. Recall that in period 0, individuals may choose to homestead or purchase land. Conditional on type, the model for purchasers is the same as for homesteaders with one exception: purchasers gain the title in period 0, meaning that $Sell_{it} = Sell_{notitle} + \kappa \cdot 200$ if $d_{it} = S$ and $d_{i,0} = purchase$.

3.2 Full Information Model

In our Bayesian model, individuals do not have perfect information. Instead, they learn information about their farming ability through the process of farming. We compare the estimation results of the Bayesian model to the results of a full information model given in equation (1). In the full information model, the farmers know their ability, μ , and this parameter has the same value for all farmers. Thus, for the full information model we get the value functions and choice probabilities by replacing $\bar{\pi}_{mt}$ by μ for all m and t .

4 Identification and Estimation

In this section, we describe the variation in our data that allows us to identify the structural parameters of the Bayesian model. We estimate the following parameters: κ (converts dollars into utility), μ_m (the type-specific ability), η_m (the type-specific bias), $H_{A,home,t=0}$ (the proportion of homesteaders who are type A), $H_{A,purchase,t=0}$ (the proportion of purchasers who are type A), σ_{prior} (the standard deviation of the prior), δ (the discount rate), γ (the disutility from being a farmer), ρ (the flexibility of being a purchaser versus homesteader), and X_t (the information realizations in each period). The parameters γ and ρ are used in the counterfactuals, which we discuss in Section 6.3.1. $Sell_{notitle}$ denotes the utility from selling the farm before acquiring the title through homesteading and equals the net proceeds of selling times the parameter κ . The value functions are functions of the above parameters.

The parameter κ , and thereby $Sell_{notitle}$, is identified by the ratio of the probability of selling over abandoning. Intuitively, this is because if a farmer wants to exit, they can either sell or abandon, and the ratio between those two choices identifies the value of money. In particular, the probability that a farmer of type m sells in period t is

$$\Pr_{m,home,t}(S) = \frac{\exp\{Sell_{notitle} + 1(t \geq 10)\kappa \cdot 200\}}{1 + \exp\{\bar{\pi}_{mt} + \delta V_{mt}\} + \exp\{Sell_{notitle} + 1(t \geq 10)\kappa \cdot 200\}}.$$

Averaging over types using the probability that a farmer in period t is type m , $\Pr_{home,t}(\text{type} = m)$, as the weight yields

$$\begin{aligned} \Pr_{home,t}(S) &= \sum_{m \in \{A,B\}} \Pr_{m,home,t}(S) \times \Pr_{home,t}(\text{type} = m) \\ &= \sum_{m \in \{A,B\}} \frac{\exp\{Sell_{notitle} + 1(t \geq 10)\kappa \cdot 200\} \times \Pr_{home,t}(\text{type} = m)}{1 + \exp\{\bar{\pi}_{mt} + \delta V_{mt}\} + \exp\{Sell_{notitle} + 1(t \geq 10)\kappa \cdot 200\}}. \end{aligned}$$

Note that $Sell_{notitle}$ and κ do not depend on the type. Similarly, the probability of abandoning the farm is

$$\Pr_{home,t}(A) = \sum_{m \in \{A,B\}} \frac{\Pr_{home,t}(\text{type} = m)}{1 + \exp\{\bar{\pi}_{mt} + \delta V_{mt}\} + \exp\{Sell_{notitle} + 1(t \geq 10)\kappa \cdot 200\}}.$$

Taking the ratio of $\Pr_{home,t}(S)$ and $\Pr_{home,t}(A)$ for any t identifies κ , and thereby identifies

$Sell_{notitle}$. Further, this identification argument only relies on the probabilities that are conditional on the farmer being a homesteader. Being a homesteader is observed, and the same argument applies when conditioning on being a purchaser. Thus, κ is identified using just the homesteaders or just the purchasers. After identifying κ, γ and ρ are identified by the same arguments as in McFadden (1978).

We use results from the duration literature to allow us to identify the unobserved mixture of heterogeneity among farmers. The model is a discrete hazard model with constant hazard, in the duration terminology by Hausman and Woutersen (2014), who show nonparametric identification of the mixing distribution. Assuming that the effect of the prior converges to zero over time means that the model becomes stationary. In particular, the probabilities of continuing, selling, and abandoning no longer depend on time after conditioning on type. Thus, the model is a discrete hazard model with constant hazard. Therefore, the assumption that the effect of the prior converges to zero means that the unobserved mixture of abilities, $\mu_m, m \in \{A, B\}$, and their probabilities are identified. Alternatively, the Kansas 1880 census has detailed information on the individual farmers' revenue and cost. The revenue minus cost of homesteaders and purchasers, and thereby the average values of μ_m , are identified by the Kansas 1880 census. We use the method with the Kansas census as a robustness check.

Continuation decisions in the later periods are primarily based on the true ability, while continuation decisions in the early periods are primarily based on the initial prior belief about ability. As $\mu_m, m \in \{A, B\}$, and the probability of type m are identified, the fractions of farmers who continue, abandon, or sell in the early periods identify the bias parameter $\eta_m, m \in \{A, B\}$. How quickly the farmers update their beliefs, revealed by the difference in the probabilities of choosing to continue, abandon, or sell over time, identifies the standard deviation of the prior, σ_{prior} . The changes over time in the prior means identify the information, X_t , that caused these changes for all periods. In particular, when farmers receive a negative realization, the fraction selling and abandoning will jump up from the trend, and when farmers receive a positive realization, the fraction selling and abandoning will jump down from the trend.

We are able to separately identify the discount factor δ because the value function increases in the years before the homesteader gets the title. Separately identifying δ and μ_m is possible because of the institutional features of our setting and our rich dataset: the homesteaders have to either wait to get the title or pay for it earlier, so we can identify the discount

factor from this tradeoff. The homesteaders discount from $t = 10$ (when they acquire the title). The value functions in $t = 10$ and in all other periods are unique by Banach's (1922) fixed point theorem, as discussed before. Finally, $H_{A,home,t}$ and $H_{A,purchase,t}$ are functions of these fixed points and the other identified parameters. The estimation is further discussed in the online appendix.

5 Data Description

This paper uses five individual-level data sources: (i) individual purchase and homestead decisions; (ii) sale deeds of individual resale decisions; (iii) the Kansas agricultural censuses of 1870 and 1880; (iv) the identified U.S. Population Censuses of 1860, 1870, and 1880 for Kansas; and (v) plot-level historical land characteristics. Using individual farm level data, as opposed to the county level data generally used in previous homesteading literature, is necessary to estimate the decision-making process based on learning new information because information is learned at the individual level. Matching together individual data on land acquisition and land resale at the plot level shows how long the farmer held the land. The 1870 and 1880 Kansas agricultural censuses record production at the farm level. We match these production data to the plot-level land acquisition data to provide a robustness check in which we use individual production data to account for the difference in farming ability, instead of estimating the difference as part of the model. We also use individual-level data from the U.S. Population censuses of 1860, 1870, and 1880 and the National Commodity Crop Productivity Index from the Soil Survey Geographic (SSURGO) database to demonstrate the farmers and land quality do not differ along observable characteristics. Matching between the data sources is described in the online appendix.

5.1 Initial Land Acquisition from the Federal Government

The Bureau of Land Management (BLM) tract books record all original acquisitions of land from the federal government, including the homesteads and purchased land used in this analysis. The level of observation is the farm. The tract books record both successful and unsuccessful homestead claims and purchases, including the original date filed, the date the patent was applied for, the date the patent was acquired, the name of the owner, the size of the acreage, the price per acre, the Act it was filed under, and the state, county, township-

range, section, and aliquot (PLSS). The typical farm size is 160 square acres. The land in the tract books contain only new, unbroken land, not land already used for farming that changed hands.

We digitize 9,279 unique homestead and cash purchase observations from two counties in Kansas: Bourbon County and Woodson County. These two counties have comparable latitudes, but Woodson County is slightly farther west than Bourbon County. This difference means that it was acquired by individuals from the federal government later than land in Bourbon County, reducing the likelihood that our results are driven by time-varying factors such as periods of drought or business cycles. The choice of counties allows us to keep the aridity, temperature, and weather conditions similar.

While high ability farmers may be more likely to choose high quality land, the fact that our setting is geographically small mitigates this issue. Land within a small area is of similar quality and has similar market access (see Section 5.4). Additionally, historians concur that farmers in the nineteenth century did not have a good understanding of land quality. Shannon (1966) concludes that the majority of a settler's information about whether the soil was of good quality for growing crops was based on the soil's color. Both because the land quality in these counties does not vary greatly and because farmers did not have good initial information about soil quality, better farmers selecting better land poses little threat to our identification.

5.2 Land Resale

About one third of the counties in Kansas have available all resale deeds of land from the Registers of Deeds. These deeds begin with the original resale from the first recognized owner who acquired it from the federal government and stretch into the twentieth century. Each deed is a handwritten legal document which lists the name(s) of the seller, the name(s) of the buyer, the number of acres sold, the exact aliquot (PLSS), the price per acre, and the exact date (day/month/year) of the sale. To our best knowledge, these deeds have never been used for any type of quantitative research.

We digitize the sale deeds for Bourbon and Woodson counties from 1852 to 1896, which total 18,473 observations. Many of these deeds represent sales after the initial resale of land, and this paper uses only those deeds which are initial resales by the owner who acquired it from the federal government via either purchasing or homesteading (i.e., the first resale).

The time period we use covers the vast majority of the early land investment activity in these two counties. Since we are only concerned with the initial resale period, the period from 1852 to 1896 captures the majority of the initial resale. By definition, tract book aliquots acquired later in the period are less likely to match with resale deeds than tract book aliquots acquired early in the period, since the resale deeds extend only until 1896. However, 84% of the tract book aliquots in these counties were acquired before 1874, meaning that the deed data extends 22 years after their acquisition from the federal government. For this reason, we use only the matched dataset with tract books acquired before 1874. The match rate in 1873 is 68% and the match rate in 1861 is 75%, so we do not obtain a statistically lower match rate (at the 5% level) from the resale deeds to the tract book observations acquired late in the period than from the resale deeds to the tract book observations acquired early in the period.

We observe the exact date (day/month/year) that the owner acquired the land from the federal government, the exact date that they obtained the title, and the exact date that the farm was resold. This allows us to calculate the number of days the land was held by the original owner before they resold it in order to determine how quickly farmers update their information. Figure 2 illustrates this. Purchased land is much more likely to be resold within the first six months than at any other time in our data. Homesteaded land is frequently abandoned for the first several years, with the abandonment rate dropping off after the homesteaders obtain the title in period 10. Homesteads are regularly sold before period 10, meaning that homesteaders often paid \$200 additional to obtain the title more quickly. There is an uptick in the number of homesteads sold at period 10 when they receive the title naturally; however, few homesteads are resold after about seven years (period 14).

Matching between the tract books and the deeds yields a dataset of 7,011 observations. Using these 7,011 observations, we construct a matrix of each observation in each period, i.e., a 7,011 by 24 matrix.⁶ Each cell of the matrix indicates whether the farmer continued, sold, or abandoned in that period, or whether they have exited the game (if they sold or abandoned previously).

⁶We use periods 21 through 24 for the final fixed points.

5.3 Individual Agricultural Censuses

In order to provide supporting evidence for our model, we use data from the Kansas agricultural censuses of 1870 and 1880. We use the production data to conduct a robustness check for the difference in farming ability directly, without imposing the structure of the model. These censuses are individual-level censuses, separate from the published federal agricultural censuses, which were conducted at the county level. The federal censuses only publish farm output at the county level, but it is important for dynamic optimization to use data at the individual level because that is the level at which optimization decisions occur. Each observation in the Kansas agricultural census includes the name of the farmer, whether they are the farm owner or a tenant, the acreage of the farm, the farm value in dollars, the types and amounts of the crops and livestock grown on the farm (e.g., 100 bushels of wheat and 8 milk cows), the value of the livestock in dollars, the date of the census, and the township in which the farm is located. We use the location, census date, and name of the farmer to match the Kansas agricultural census data to the BLM tract books.

5.4 Individual Demographic and Soil Quality Data

To show that both the demographic characteristics of purchasers and homesteaders are similar and that the land they farm is similar, we use individual demographic data from the U.S. Full Count population censuses of 1860, 1870, and 1880 and farm-level soil quality data from the National Commodity Crop Productivity Index from the Soil Survey Geographic (SSURGO) database. We match the BLM tract book observations to both of these datasets at the individual level using the location (PLSS) for the soil quality data and the location, name, and year for the U.S. population censuses.

The demographic data from the U.S. population censuses allows us to demonstrate that there are no observable differences in the distribution of demographic characteristics based on land acquisition decisions at the individual level.⁷ We use the age, gender, personal property, occupation class, and number of children variables from the U.S. population censuses.

The soil quality data classify eight soil categories based on their effectiveness for farming. This measure is called the non-irrigated land capability class, where higher numbers indicate “greater limitations and narrower choices for practical use” (United States Depart-

⁷This result is shown in Figures 14 through 19 in the online appendix.

ment of Agriculture (2013)). This is a more accurate measure than farmers had at the time. Nearly 80% of the land in our sample is classified as above average soil quality for the U.S.. Therefore, the idea that better farmers can select better land does not pose a significant threat to identification in our model because the vast majority of land in our sample was high quality. Further, soil quality does not differ in a meaningful way on land acquired by homesteaders and by purchasers: on average, homesteaded land has a soil classification of 3.46 and purchased land has a soil classification of 3.55 in our sample, both of which are above average quality. Further, as we find below that purchasers have higher farming ability than homesteaders, we observe that better farmers actually selected slightly worse quality land on average (3.55 vs. 3.46), meaning that we do not have a threat to identification from better farmers selecting land for its higher soil quality.

6 Results and Counterfactuals

In this section, we describe the results from the Bayesian and full information models, and the counterfactual estimates we construct from both of them. The full information model is comparable to a typical Rust (1987) specification in which the farmers know their true ability to farm. The results of our Bayesian model demonstrate its advantages over a full information model: it allows us to estimate the biases in the beliefs about farming ability for each unobserved type, the standard deviation of the prior belief, and the information realizations. These parameters are not a part of a full information model, but they are important because they empirically demonstrate the value of learning in our context.

Using our Bayesian model, we estimate two counterfactuals. Firstly, we estimate what fraction of homesteaders would have bought land if they could only purchase it. This counterfactual is a policy counterfactual which quantifies the importance of the Homestead Act in western expansion. Secondly, we estimate what fraction of homesteaders would have abandoned their farms in the first five years if they had begun with different information sets. This counterfactual demonstrates the importance of beliefs on decision-making in our setting. The first counterfactual is possible to calculate under a full information specification, which we do. The second counterfactual can only be calculated using our Bayesian model. These counterfactuals further demonstrate the advantage of our Bayesian model over a full information model because they incorporate the value of learning, which is critical in our

setting.

6.1 Empirical Results from the Full Information Model

The farmers, homesteaders and purchasers, build a farmhouse and develop the land. If they subsequently sell the property, then the average proceeds are \$1480 for the homesteader and \$1475 for the land buyer for an 160 acre farm. The parameter that converts dollars into utility is κ , and the estimate for κ is $5.7 \cdot 10^{-4}$ (s.e. $4.9 \cdot 10^{-5}$).

The estimated ability to farm is substantially different for homesteaders and buyers. When estimating a full information model, i.e., no prior beliefs, the estimates for their farming ability μ are 0.363 (s.e. $1.6 \cdot 10^{-2}$) for the homesteaders and 0.925 (s.e. $6.5 \cdot 10^{-2}$) for purchasers. A positive μ is interpretable as farmers receiving positive benefits from farming in the current period in the full information model. In particular, μ is the difference between the expected utility from farming and the expected utility from the outside option (abandoning). This utility is substantially larger for purchasers.

To facilitate a more detailed understanding of these results, we convert the parameter estimates into 1870 dollars using $\hat{\kappa}$, the parameter which converts dollars into utility. The farming ability for the homesteaders is equivalent to \$636.84, and \$1,622.81 for buyers, per period over the outside option. However, the full information model does not capture the differences in farming ability among farmers, nor the bias in beliefs. A negative bias in beliefs may lead to discontinuing farming despite a positive μ .

We estimate $Sell_{notitle}$, the utility a homesteader receives from selling their farm before they would normally acquire the title, at 0.732 (s.e. $6.3 \cdot 10^{-2}$), which is comparable to \$1,284.95 in 1870 dollars. Recall that the utility a farmer gets from selling the farm if they have the title is $Sell_{notitle} + \kappa \cdot 200$, so the dollar value (in 1870 dollars) from selling the farm if they already have the title is estimated at \$1,485. Since the average resale value of a farm in our data is \$1,478, our results are consistent.

The full information model fits the data with a pseudo- R^2 of 0.596. In both the model and in the data, the vast majority of farmers continue to farm each period.⁸ In the full information model, by definition, the probabilities of continuing, selling, and abandoning do not change in any period for purchasers and do not change after period 10 for homesteaders. However, in the data, the probabilities of each choice are not the same every period; specifically, the

⁸See Figures 8 and 9 in the online appendix.

probability of selling the farm if the farmer purchased the land originally is much higher in period 1 than in any other period, and for homesteaders, the probability of selling the farm is the highest in periods 11, 12, and 13. The full information model is unable to capture these differences in the probabilities of each choice by period. Further, the full information model cannot capture the significant decrease in the probabilities of abandoning after period 10 or selling after period 13 for homesteaders because the probabilities of each choice do not change after the homesteader receives the title. These changes in the choice probabilities are better captured by our Bayesian model.

The above abilities to farm were estimated without assuming a common κ . Below, our model with unobserved types strictly generalizes the full information model with homesteaders and purchasers.

6.2 Empirical Results from the Bayesian Model

Above, we detailed the results from the full information model. However, the full information model does not take into account the fact that individuals acquiring land may have limited knowledge about farming and learn more information over time. Because this learning was a crucial feature of our setting, we turn to the results from our Bayesian model. Table 2 in the appendix reports the parameter estimates from the Bayesian model and the full information model.

The Bayesian model fits the data significantly better than the full information model (pseudo- R^2 of 0.821 vs. pseudo- R^2 of 0.596). This difference is largely driven by the fact that the Bayesian model matches the data better than the full information model for the selling and abandonment decisions for the early periods. The Bayesian model is able to capture the fact that the choice probabilities change in every period, even after the farmer has the title, because their beliefs about their own farming ability change in every period based on the new information they receive.⁹

Using the Bayesian model, we estimate the farming ability for the high ability farmers, $\bar{\mu}_A$, at 1.249 (s.e. $8.9 \cdot 10^{-2}$). This can also be understood as the value of farming in the current period, so the low ability type values farming in the current period slightly less, at 1.237 (s.e. $8.1 \cdot 10^{-2}$). Using $\hat{\kappa}$, we convert these parameters into 1870 dollars: the farming ability for high ability farmers is equivalent to \$2,191.23 and \$2,170.18 for the low ability

⁹See Figures 10 and 11 in the online appendix for the fit of the Bayesian model to the data.

farmers per period over the outside option. This result is consistent with external data sources: the 1870 and 1880 Kansas Agricultural censuses values average farm production plus farm value for homesteaders at \$1,201.09 and for purchasers at \$2,263.46, conditional on the farm having existed for between 8 and 12 years.¹⁰

The different values of farming in the current period for each type culminate in different fixed points for the value of farming. We estimate $V_{A,final}$ to be 7.622 (s.e. 0.367) and $V_{B,final}$ to be lower, at 7.542 (s.e. 0.342). These final fixed points are equivalent to \$13,371.93 for the high ability farmers and \$13,231.58 for the low ability farmers in 1870 dollars. These numbers are the expected discounted utilities of farming over the outside option expressed in dollars. Additionally, we are able to estimate the discount factor, and find $\delta = 0.759$ (s.e. $5.2 \cdot 10^{-4}$). This discount rate is consistent with high-risk projects.

Farmers with different abilities may also exhibit differently biased beliefs. We estimate the bias parameter for the high ability farmers as -1.442 (s.e. 0.147), and -1.502 (s.e. 0.114) for the low ability farmers. Since both $\eta_m, m \in \{A, B\}$, are negative, farmers of both types are pessimistic about their abilities. In 1870 dollars, these biases are equivalent to -\$2,529.82 for the high ability farmers and -\$2,635.09 for the low ability farmers, meaning that initially, farmers anticipated their farms being about \$2,500 less productive than was accurate. This result that farmers in Kansas in the late 1800's were pessimistic about their outcomes is consistent with historical accounts: a newspaper in Junction City, Kansas in 1870 published that farmers "prophesy a failure of crops whenever it is dry for a 'straight' week" (Martin (1870)).

We find that high ability farmers are more likely to continue rather than sell or abandon relative to the low type, which is partly due to a higher farming ability, and partly due to less pessimism. Lower ability farmers have a lower ability and a slightly less accurate belief about that ability. However, both types of farmers are able to refine their beliefs over time using Bayesian updating. This causes their beliefs to be less biased: $\eta_m, m \in \{A, B\}$, represents their initial bias, and the whole bias term is decreasing in t . For example, by period 10, the bias is -0.602 and -0.623 for types A and B respectively (-\$1,056.14 and -\$1,092.98 in 1870 dollars, respectively), meaning that the bias has decreased by more than half after about five years. Over time, farmers' beliefs about their ability move closer to their true ability as they learn more information. How the bias for type B changes over the twenty periods

¹⁰These census values have been adjusted to reflect values for six months, to match the six-month periods in our model.

is shown in Figure 7.¹¹ In the first three periods, the biases decrease rapidly and then level off to remain fairly stable for about ten more periods. The biases then decrease again from periods 13 to 20.

The fact that the bias has not disappeared completely even by period 20 highlights the importance of the information realizations X_t . When individuals receive negative information realizations, their already pessimistic bias reduces more slowly than if they had received positive realizations in that period. In the first two periods, individuals receive positive realizations and the bias reduces by approximately 50% in these early periods: by period 2, the bias is -0.709 for type A and -0.753 for type B (see Table 7). These biases are equivalent to -\$1,243.86 and -\$1,321.05 in 1870 dollars, respectively, meaning that by period 2, farmers underestimate their productivity by about \$1,300, compared to underestimating it by about \$2,500 initially. Because of this large initial reduction in the bias, estimating the information realizations is critical in many settings. The large magnitude of the bias $\eta_m, m \in \{A, B\}$, may lead to farmers abandoning at a higher rate than is optimal in the early periods.

Table 7 in the online appendix shows the estimates for the information realizations X_t . These realizations are the information that farmers use to update their beliefs. For example, farmers might learn in the first six months that they have difficulty purchasing farming equipment because they have limited credit in Kansas, which would create a negative information realization in the first period. We find that in the first two periods, farmers receive highly positive information realizations. In periods 3 through 13, they receive negative information realizations, and then in the last periods, the information realizations are positive again.

We estimate the standard deviation of the prior, σ_{prior} , as 0.443 (s.e. $2.3 \cdot 10^{-2}$). The standard deviation of the prior is relative to learning while farming, and 0.443 is comparable to five years of farming experience.¹² The context of this result is that more than half the U.S. population worked in agriculture at the time, meaning that most people had experience farming (Gates (1968)).

In our Bayesian model, the parameter $\mu_m, m \in \{A, B\}$ captures what we call farming ability, which may also reflect things like taste for farming or wealth. Therefore, to strengthen our results, we demonstrate that the distribution of the six land acquisition probabilities is consistent across the distribution of observable demographic characteristics from the U.S. population census. Specifically, we show that the probabilities of continuing, selling, or

¹¹The bias of type A is somewhat smaller and reported in Figure 13 in the online appendix.

¹²We calculate the years of farming experience to which σ_{prior} is equivalent by $\frac{1}{\sigma_{prior}^2} = 5.195$.

abandoning, conditional on having initially homesteaded or initially purchased, are constant across age brackets of the owner, gender of the owner, year the farm was started, occupational choice of the owner, number of children of the owner, and personal property of the owner. For example, we show that the probabilities for each land acquisition choice are the same for young, old, and middle-aged farmers: older farmers are not more likely to abandon their farms than younger farmers. The probabilities of different land acquisition decisions are the same at all levels of observable demographic characteristics, which indicates that our results for μ_m are not driven by demographic characteristics.¹³

Figure 4 illustrates how farmers made different decisions (continue, sell, or abandon) based on their different beliefs about their farming ability. These beliefs depend on the initial bias and the information realizations. When farmers think their expected farm output is high (i.e., their posterior belief is higher), the probability to continue is relatively large. When farmers think their expected farm output is low (i.e., their posterior belief is lower), the probability of selling or abandoning is relatively high. In other words, the model formalizes that individuals develop beliefs that affect their decisions. Notably, all the individuals in Figure 4 initially acquired land, so they all thought they were likely to be able to create a successful farm. However, learning more information about their farming ability in later periods allowed them to refine their decision, leading many farmers to abandon or sell their land. Figure 4 demonstrates how learning about their own ability impacts individuals' decisions under uncertainty. In particular, it shows that when their posterior mean is higher, they are less likely to sell or abandon their farm.

By estimating this model with multiple types, we are able to distinguish how high and low ability farmers self-select into purchasing land versus homesteading it. We find that 99.0% (s.e. 29.0%) of the purchasers are high ability, versus 1.0% (s.e. 0.973%) of the homesteaders. Of the high ability farmers, 98.8% (s.e. 15.7%) of them select into purchasing land and 1.2% (s.e. 15.7%) select into homesteading, while of the low type farmers, 0.8% (s.e. 13.3%) of them choose to purchase and 99.2% (s.e. 13.3%) of them homestead. This result, reported in Table 3 in the appendix, has empirical implications in the homesteading literature: while much of the previous literature has assumed that the decision to purchase versus homestead land is exogenous after various county-level controls, this result indicates that the farmers involved may be different and select into land acquisition decisions based on that unmea-

¹³This result is shown in Figures 14 through 19 of the online appendix.

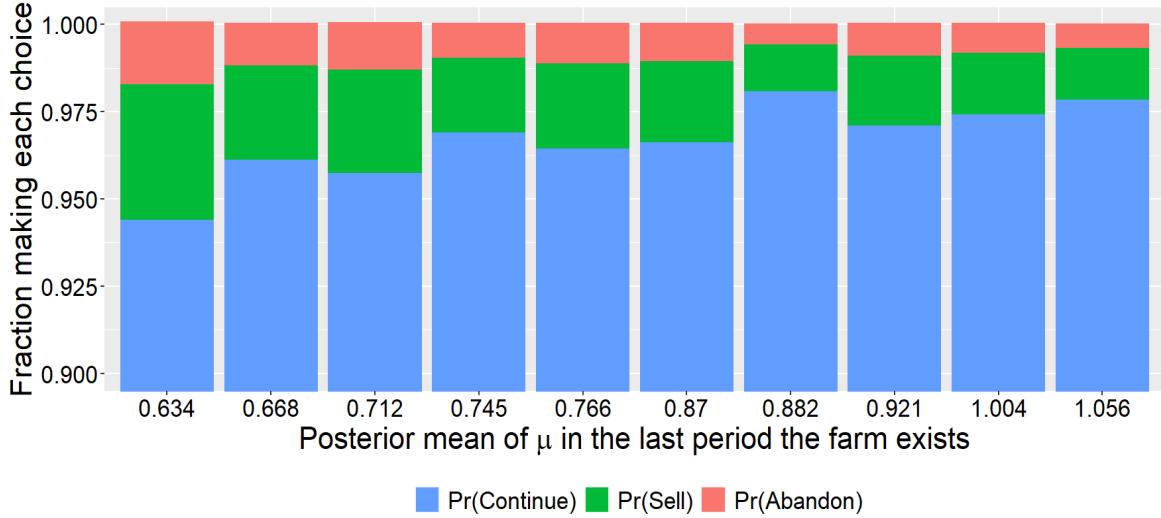


Figure 4: Farmers make different decisions based on their beliefs: when they believe their ability is higher, more of them continue to farm, and when they believe their ability is lower, more of them abandon and sell their farm

sured difference. For example, if the researcher is interested in how the purchase-homestead decision affects future farm productivity, even with individual-level controls, the results may be conflating the effect of farming ability with the effect of the purchase-homestead decision. We model the purchase-homestead decision with ability taken as given, and our model suggests that high and low ability farmers separate into purchasing and homesteading land. Determining the amount of self-selection into unobserved types is relevant for a large number of empirical settings.

As the farmers move through time, this selection by types continues. The ratio of high ability farmers to low ability farmers is increasing somewhat in t (0.45659 in period 1 to 0.45734 in period 20) because low ability farmers are more likely to exit through selling or abandoning their farm. This same selection process is observable in the ratio of the survival functions by type, i.e., $\frac{H_{A,home,t}}{H_{B,home,t}}$ and $\frac{H_{A,purchase,t}}{H_{B,purchase,t}}$. For example, $\frac{H_{A,home,t}}{H_{B,home,t}}$ increases from 1.013 in period 1 to 1.091 in period 20, meaning more high ability homesteaders are continuing to farm than low ability homesteaders. Figure 5 illustrates how the information realizations relate to the probability of continuing for each type. Because of dynamic selection, the ratio of high ability farmers to low ability farmers increases over time. In part, selection by type is due to the information realizations they receive each period, which are estimated by X_t . We use the X_t to better understand how farmers make decisions in a Bayesian setting.

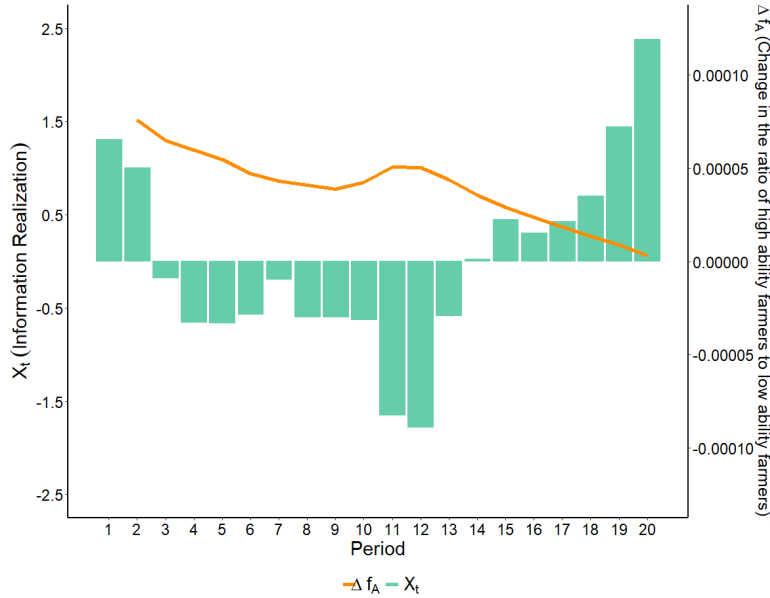


Figure 5: Low ability farmers are more susceptible to negative information realizations than high ability farmers

In particular, let Δf_A denote the change in the ratio of high ability farmers to low ability farmers. Figure 5 shows the relationship between the information realizations X_t and Δf_A . The ratio f_A changes most in the first few periods, when learning new information is of the most value to farmers deciding whether to sell, abandon, or continue. By the last periods, there is little change in the ratio of high ability farmers to low ability farmers because the information has already been learned. However, in periods 11 and 12, farmers receive highly negative realizations ($X_{11} = -1.646$ (s.e. 0.208) and $X_{12} = -1.775$ (s.e. 0.300)). In these periods, Δf_A increases: more high ability farmers are continuing and more low ability farmers are selling or abandoning when they receive negative realizations. This result indicates that low ability farmers are more susceptible to these negative realizations than are high ability farmers. This finding is consistent with Schultz’s (1975) concept of responding to disequilibria. In Schultz’s (1975) framework, low ability farmers are less insulated from negative realizations than high ability farmers, and how efficiently farmers are able to respond has value.

We use the data from the Kansas agricultural census of 1880 to provide supporting evidence for the difference in farming ability. Table 1 shows a back-of-the-envelope calculation for the average farm-level profit for purchasers and homesteaders at the individual level. We

calculate revenue as the number of bushels of each type of crop produced times the price per bushel of each crop plus the number of each type of livestock born times the price of each type of livestock in the census year. We calculate their cost as the cost of fertilizer and wages in the census year, plus the cost of two draft animals (to pull the plow), plus the value of machinery and the cost of fences, both multiplied by 0.2 to account for depreciation and capital cost, plus the cost of enough corn to feed the two draft animals, as given below,

$$Revenue_i = NumberOfBushels_{i,crop} \times Price_{crop} + NumberLivestock_{i,type} \times Price_{type}$$

$$Cost_i = FertilizerCost_i + Wages_i + 2 \cdot DraftAnimal + \\ + 0.2 \times (MachineryValue_i + FenceCost_i) + FeedCost.$$

Table 1: Annual Farm Profit by Initial Land Acquisition Decision

	Homesteaded	Purchased
Revenue	\$693.72 (540.84)	\$768.71 (626.07)
Cost	145.30 (62.15)	149.55 (69.70)
Profit	537.72 (518.62)	603.84 (591.96)
	$N = 284$	$N = 208$

The large standard deviations in Table 1 reflect the fact that many farmers abandoned because they earned negative profits. The average profit for purchasers is \$66.12 dollars larger (in 1880 dollars) than the average profit for homesteaders. This result is consistent with our results from the Bayesian model: purchasers are more likely to be high ability farmers.

We use the data from the Kansas agricultural census to externally verify the results from our Bayesian model: without relying on the modeling process to provide structure, we find consistent results using individual agricultural output data. As a robustness check, we fix the difference in the ability between the two types to be $\hat{\kappa} \cdot 66.12$ to reflect the difference in the value of farm output from the census between purchasers and homesteaders, and then re-estimate the Bayesian model. Table 6 reports these results. The estimates are nearly

identical. The estimate for the ability for the low type, μ_B , is slightly smaller when using the agricultural census data (1.211 versus 1.237 under the Bayesian model without the agricultural census data), and the standard deviation of the prior, σ_{prior} , is also slightly smaller (0.378 versus 0.443 under the Bayesian model without the agricultural census data). Likewise, the probability of abandoning in the first five years and the counterfactual probability of farming in a world without homesteading also have very similar results using the agricultural census data and estimating the model without it. (These counterfactuals are discussed in the next section.) Together, these results show that our model is externally consistent with other data sources.

6.3 Counterfactuals from the Bayesian Model

We consider two separate counterfactual scenarios using our Bayesian model. First, what fraction of homesteaders would have bought land in a counterfactual scenario in which homesteading was not an option? Second, what fraction of homesteaders would still have abandoned their farm if they had begun with the information they learned through farming? Specifically, in order to understand the importance of the information learned by farmers, we estimate a counterfactual scenario in which they begin with the information they learned after (i) five and (ii) ten years.

6.3.1 Farming without Homesteading

The concept of the Homestead Act had been hotly debated in Congress for several decades before its passage (Gates (1968)). It was by no means clear that such an Act would pass; therefore, we consider a counterfactual scenario of westward expansion without homesteading. Without the Homestead Act, we assume buying the land was a requirement for farming, and some fraction of homesteaders may not have done that. We estimate that fraction. Formally, this counterfactual is the conditional probability of buying the land given that one homesteaded in the data but homesteading is no longer available. We estimate this counterfactual using our Bayesian structural model and we use a nested logit model to allow for correlation between the propensity to buy the land and to homestead.

Homesteaders acquire the title to their land for a nominal fee after five years. Further, they do have the option to acquire the title earlier by paying \$200 to the Department of the Interior. In particular, at $t = 0$ a homesteader could buy the land for \$200. When deciding

between buying and homesteading, the farmer will also take into account that life events may force a sale of the farm, and in our data about 20% of homesteaders sell their land in the first five years. In the case of such a sale, the farmer will still have to pay \$200 to acquire the title. Therefore, the farmer will take the discounted expected cost of acquiring the farm into account at $t = 0$. This causes the difference in cost between purchasing and homesteading to be somewhat less than \$200. In particular, the discounted expected cost of acquiring the title of the farm at $t = 0$ is

$$\$200 - \$200 \cdot \sum_{t=1}^9 \delta^t \Pr(\text{homestead } s = 1, \dots, t-1, \text{ and sell at } t).$$

In our data, this discounted expected cost is \$190.73 in 1870 dollars. We only need to sum over periods 1 through 9 since the homesteaders acquire the title at $t = 10$. Using μ_A and μ_B , we calculate this payment separately for type A and type B . Call these Bayesian discounted expected costs Ψ_A and Ψ_B . In our data, the discounted expected cost of acquiring the title to their farm is \$185.67 for type A and \$183.23 for type B (in 1870 dollars). As discussed in the previous section, the low ability farmers are more likely to sell their farm than the high ability farmers; therefore, type B has a somewhat lower discounted expected cost of acquiring the title. This Bayesian expectation of the payment is lower for type B because of the more negative bias.

To calculate the fraction of individuals who would have bought land in a counterfactual situation in which homesteading was not available, we use the following nested logit model, where $(1 - \rho_{learn})$ is the dependence between purchasing and homesteading and γ_{learn} is the disutility to move to and farm in the counties under consideration for people outside those counties. Consider

$$\Pr(\text{buy}|\text{homestead or buy}) = \frac{1}{1 + \exp\{\kappa \cdot \Psi_A / \rho_{learn}\}} \Pr(A) + \frac{1}{1 + \exp\{\kappa \cdot \Psi_B / \rho_{learn}\}} \Pr(B),$$

and $0 < \rho_{learn} \leq 1$, $\kappa > 0$

where $\Pr(A) = \Pr(A|HS) \Pr(HS) + \Pr(A|Purchase) \Pr(Purchase)$ and $\Pr(B) = 1 - \Pr(A)$. The probability $\Pr(\text{buy}|\text{homestead or buy})$ can be estimated using our data; in particular,

we use

$$\Pr(\widehat{\text{buy}}|\widehat{\text{homestead or buy}}) = \frac{\#sales \text{ at time } = 0}{\#sales \text{ at time } = 0 + \#\widehat{\text{homestead at time}} = 0}.$$

Note that κ , Ψ_A , and Ψ_B are identified, so the only remaining parameter, ρ_{learn} , is identified as well. Further, $\Omega_A = 200 - \Psi_A$ and $\Omega_B = 200 - \Psi_B$.

Next, consider the probability of not farming in our Kansas counties,

$$\Pr(\text{no farming}) = \frac{\Pr(A)}{1 + [\exp(\gamma_{learn}/\rho_{learn}) + \exp\{\gamma_{learn}/\rho_{learn} - \kappa \cdot \Psi_A/\rho_{learn}\}]^{\rho_{learn}}} + \frac{\Pr(B)}{1 + [\exp(\gamma_{learn}/\rho_{learn}) + \exp\{\gamma_{learn}/\rho_{learn} - \kappa \cdot \Psi_B/\rho_{learn}\}]^{\rho_{learn}}},$$

$$\text{and } 0 < \rho_{learn} \leq 1, \kappa > 0.$$

The last equation identifies γ_{learn} and we use the U.S. census to calculate $\Pr(\text{no farming})$. We then use $\hat{\gamma}_{learn}$ and $\hat{\rho}_{learn}$ to calculate the counterfactual probability of not farming given that homesteading is no longer available. Using the above estimates, we can calculate the counterfactual probability of farming in a situation with no homesteading by removing a choice in the equation for $\Pr(\text{no farming})$,

$$\Pr(\widehat{\text{no farming}}|\widehat{\text{no homesteading}}) = \frac{\widehat{\Pr(A)}}{1 + [\exp\{\hat{\gamma}_{learn} - (\hat{\kappa} \cdot \hat{\Psi}_A)\}]} + \frac{\widehat{\Pr(B)}}{1 + [\exp\{\hat{\gamma}_{learn} - (\hat{\kappa} \cdot \hat{\Psi}_B)\}]}.$$

The probability of buying conditional on homesteading or buying is

$$\Pr(\text{buy}|\text{homestead or buy}) = \frac{1}{1 + \exp\{\kappa \cdot \Psi_A/\rho_{learn}\}} \Pr(A) + \frac{1}{1 + \exp\{\kappa \cdot \Psi_B/\rho_{learn}\}} \Pr(B),$$

$$\text{and } 0 < \rho_{learn} < 1, \kappa > 0,$$

where $\Pr(A) = \Pr(A|HS) \cdot \Pr(HS) + \Pr(A|Purchase) \cdot \Pr(Purchase)$ and $\Pr(B) = 1 - \Pr(A)$. Details on the calculation of the counterfactual are in the online appendix. We use the nested logit model to obtain the following probability,

$$\Pr(\widehat{\text{no farming}}|\widehat{\text{no homesteading}}) = \frac{\widehat{\Pr(A)}}{1 + [\exp\{\hat{\gamma}_{learn} - (\hat{\kappa} \cdot \hat{\Psi}_A)\}]} + \frac{\widehat{\Pr(B)}}{1 + [\exp\{\hat{\gamma}_{learn} - (\hat{\kappa} \cdot \hat{\Psi}_B)\}]}.$$

This counterfactual is the conditional probability of not farming given that one homesteaded in the data but homesteading is no longer available.

We find the dependence between homesteading and purchasing to be fairly large ($(1 - \hat{\rho}_{learn}) = 0.410$ (s.e. $9.9 \cdot 10^{-2}$)) in the Bayesian setting, meaning that homesteading and purchasing land are reasonably close substitutes for each other.¹⁴ The disutility to farm in the counties in Kansas under consideration for people outside that area is large ($\hat{\gamma}_{learn} = -8.971$ (s.e. 0.105)). This disutility is equivalent to \$15,738.60 in 1870 dollars, meaning that people would have to have been compensated approximately \$16,000 to move to these counties and start farming, if they did not already live there. These estimates give the counterfactual probability of not farming in a situation with no homesteading. This probability is 0.689 (s.e. $8.6 \cdot 10^{-2}$), meaning that in a counterfactual world in which the U.S. government had not offered homesteading, about 31% of homesteaders would have bought the land, while 69% would have chosen not to farm. Given that more than 270 million acres of land were homesteaded in the U.S., this reduction is highly economically significant. Western expansion had far-reaching impacts on the United States' relationship with Native Americans, racial divisions, the Dust Bowl, and the balance of partisan power in Congress. Our results indicate that without the Homestead Act, the speed of this expansion would have been reduced by about 37.5% because, while fewer homesteaders would have acquired land, those who originally purchased land would still have done so. Results are given in Table 4 of the appendix.

However, this counterfactual about the reduction of the number of farms without the Homestead Act does not mean that agricultural production would have been reduced by the same amount. We find that farmers who originally purchased land produced a higher value of livestock and crops than did farmers who originally homesteaded land. Therefore, in a counterfactual scenario where homesteading was not possible, agricultural output would have decreased by less than the reduction in the number of farmers. We provide a back-of-the-envelope calculation for bounds on agricultural production in Kansas if homesteaders had only been able to purchase land.¹⁵ This calculation bounds the agricultural output from new

¹⁴Note that $\rho = 0$ means complete dependence in McFadden's (1978) model.

¹⁵Since nearly all high ability farmers are purchasers and nearly all low ability farmers are homesteaders, we calculate the lower bound of agricultural production in Kansas in a counterfactual world without homesteading as the average farm production in 1870 for purchasers multiplied by the number of purchasers in Kansas plus the average farm production in 1870 for homesteaders multiplied by the number of homesteaders in Kansas multiplied by 0.31, the fraction of homesteaders who would have purchased in this counterfactual.

farms in Kansas in a counterfactual situation in which homesteading did not exist between \$179,290,833 and \$180,008,761 (1870 dollars). Since the total agricultural production in Kansas in 1870 was \$458,530,522, which includes resold farms, this estimate is reasonable. We also calculated this counterfactual in a full information setting. This result is given in the appendix.

6.3.2 Homesteading Under Different Beliefs

A useful feature of our model is that it allows constructing counterfactuals using different beliefs. Since we include prior beliefs that are updated over time and allow these beliefs to be biased, we can create counterfactual situations that change these beliefs and biases. For example, in many settings, agents may be optimistic or pessimistic, and researchers or policy-makers may be interested in learning how behavior would change if the agents hold unbiased beliefs. Agents make choices based on their beliefs, and beliefs are not necessarily fixed over time, so our model can be used to create a counterfactual of how agents' decisions would change if their information set evolved differently. Since we include unobserved heterogeneity in types, counterfactuals calculated from our model can also vary by type.

We consider counterfactual situations in which homesteaders began with all the information and beliefs they would have acquired later. These types of counterfactuals address a more general policy question about incentivizing learning. Our parameter estimates show that people may hold highly mistaken beliefs. The counterfactuals which we create demonstrate how long it takes to learn enough information to change decisions based on mistaken beliefs. Calculating this type of counterfactual is relevant to a number of policy settings in which policy-makers allocate a resource about which individuals have information uncertainty. If individuals have highly mistaken beliefs about the resource, they may not use the resource long enough to overcome those mistake beliefs. Understanding how long it may take individuals to overcome their mistaken beliefs can be of interest to policy-makers in structuring administrative requirements about resource allocation.

Consider $\pi_{t=0}(\mu_m)$, the prior on μ_m at $t = 0$, which has mean $\bar{\mu}_m + \eta_m$ and variation

Likewise, we calculate the upper bound in the same manner, except that we assume the most productive homesteaders would have purchased land in this counterfactual scenario, so we multiply the number of homesteaders by the average farm production in 1870 for purchasers.

σ_{prior}^2 . That is,

$$\pi_{t=0}(\mu_m) = \frac{1}{\sqrt{2\pi\sigma_{prior}^2}} \exp\left\{-\frac{1}{2\sigma_{prior}^2}(\mu_m - \bar{\mu}_m - \eta_m)^2\right\}.$$

The expected value of homesteading at time $t = 0$ is

$$V_{m,0} = \ln\{\exp(\bar{\pi}_{mt} + \delta V_{m,1}) + \exp(Sell_{notitle}) + 1\} + 0.577.$$

We can calculate counterfactuals of the above expectation, i.e., the expected values of homesteading at time $t = 0$ under different assumptions.

Suppose the homesteaders would have had the information that they have at $t = 20$ when they first acquire land. What fraction of homesteaders of each type would abandon in the first five years, if they had begun with all information they learned after (i) five years and (ii) ten years? Abandoning a homestead represents a policy failure, so this counterfactual measures the impact that additional information has on the rate of failure. The context for this counterfactual is that historians estimate that 55% of homesteads in the U.S. were abandoned (Gates (1968)). In our two counties, we observe a relatively lower abandonment rate because eastern Kansas represents the ideal farming scenario for homesteaders due to its high-quality soil and abundant water. Therefore, our estimates represent a lower bound on the value of learning to farmers in this time period. In locations where the margin for error was narrower for homesteaders, like Arizona, learning was like to be even more important.

In our setting, if individuals had begun with the information they had by period 20, significantly fewer of them would abandon or sell their farms and more of them would continue to farm. About 7% of individuals abandon their farm in the first five years in the Bayesian model in which they learn information each period. However, in the counterfactual scenario in which they have all information from period 20, only about 2.6% of individuals abandon their farm in the first five years. This result, shown in Figure 6, indicates that learning later-period information influences individuals to change their abandonment decisions a significant amount in our setting. If they had begun with the beliefs they held at $t = 20$, 37.9% fewer homesteaders who abandoned in the Bayesian scenario would have abandoned in the first five years. This means that 62.1% of the farmers who abandoned their farms in the first five years would not have done so if they knew the information they learned after 10 years.

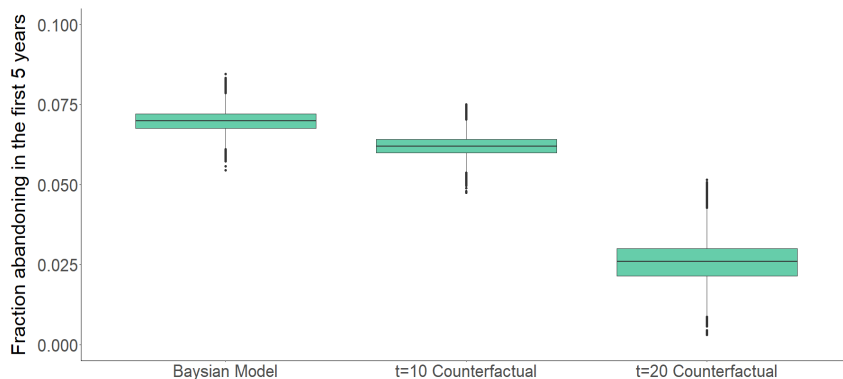


Figure 6: Fraction of homesteaders abandoning in the first 5 years under different beliefs

How much learning later-period information influences individuals to change their abandonment decisions is heterogeneous across farmers. Under the counterfactual scenario in which they begin with the information from period 20, 0.00024% of high ability homesteaders abandon, compared to 2.6% for the weighted average of both types. Similarly, 6.4% of high ability homesteaders abandon in the Bayesian model, compared to 7% for the weighted average of both types.

Likewise, we can simulate continuation, selling, and abandonment decisions under a counterfactual information set where homesteaders begin with the information they learn by period 10. In this case, 6.2% of individuals abandon their farms in the first five years. The abandonment rate for a counterfactual scenario based on $t = 10$ (6.2%) is closer to the abandonment rate for the Bayesian model (7%) than is the rate for the $t = 20$ counterfactual (2.6%) because the $t = 20$ counterfactual has farmers start with more information than the $t = 10$ counterfactual. This additional information leads to them changing their decisions to a greater extent. This result is summarized in Figure 6. Figure 6 shows a much larger difference in the fraction abandoning in the first five years between the $t = 20$ counterfactual and the Bayesian model than there is between the $t = 10$ counterfactual and the Bayesian model.

This difference in abandonment decisions between the two counterfactual scenarios is due to the information realizations in each period, X_t , and the bias terms, η_m , $m \in \{A, B\}$, which determine the beliefs in period 10 to a larger degree than they determine the beliefs in period 20. Individuals receive information realizations that are negative on average in the first ten periods, while they receive realizations that are positive on average in the last ten periods

(see Figure 12 in the online appendix). These positive information realizations eventually overcome their pessimistic bias (η_m is negative for all m), leading them to continue to farm if they have the information they learn by period 20. The effect of the initial bias reduces over time; however, the negative realizations in periods 1 through 10 add to the pessimistic bias, so relatively more homesteaders abandon their farms in a counterfactual scenario in which they begin with the information they have in period 10 than if they begin with the information they have in period 20. Simulated decisions for the counterfactual scenarios in which individuals begin with the information they learn by period 10 and period 20 are given in Table 5 of the appendix.

The percentage of farmers making a different abandonment decision given their beliefs at $t = 20$ is relatively large in our setting. These counterfactuals demonstrate how our Bayesian model allows for the estimation of the fraction of agents who make different decisions based on counterfactual beliefs, and how these counterfactuals can be heterogeneous by type. In a model without updating of Bayesian beliefs, this type of counterfactual would not be possible. This counterfactual demonstrates how information impacts the failure rate of the U.S. government’s homesteading policy.

7 Conclusion

To Americans of the time, deciding how to allocate public lands was of great importance. Shortly before the Civil War, Senator Henry Clay said in a speech to Congress: “No subject which has presented itself to the present, or perhaps any preceding congress, is of greater magnitude than that of the public lands.” The Homestead Act granted more than 270 million acres of land to individuals, or approximately 11% of the land in the United States (National Park Service). This Act impacts the balance of partisan power in Congress, wealth inequality, ecological events, and agricultural policy. It is regarded by historians to be the first major farm subsidy in the United States and encouraged the Jeffersonian ideal of the U.S. as a land of small farmers. Over 50% of the United States’ population worked in agriculture at the beginning of our time period (Gates (1968)). Both because of the importance of the agricultural sector in the nineteenth century American economy and the contention surrounding the use of public lands, our model informs our understanding of decisions faced by the majority of American workers at this time.

Information sets and beliefs change over time and dynamic optimization is inextricable from this time dimension. We extend the popular dynamic optimization model by Rust (1987) to allow the agents to change their information sets over time and to update their beliefs. We apply our model to farming in the American Midwest historically, where the U.S. government gave land away through homesteading. Individuals who homesteaded were required to farm the land. Thus, these individuals learned about their ability to farm the land in Kansas. Further, they had to decide whether to (i) acquire the title of the farm by farming for five years, (ii) abandon the farm, or (iii) sell the farm once they had the title. In our model, we allow for this updating of beliefs and dynamic optimization.

We find that farmers here initially quite pessimistic about farming in Kansas, which is corroborated by contemporaneous local news stories. Specifically, we find that farmers undervalued farming by about \$1,500. We then use our model to construct counterfactuals. If the U.S. government had not offered homesteading, then 31% of the homesteaders would have bought the land and 69% would have opted not to farm. Further, we calculate the percentage of the homesteaders that would have abandoned in the first five years in the case they had started with the information that they obtained later. We find that if individuals had begun with the information they acquired after ten years, significantly more of them would have continued to farm and fewer of them would have abandoned their farm. In particular, fewer high ability farmers would have abandoned. Our Bayesian model captures this process of learning about their own farming ability. However, farmers were initially quite pessimistic about farming in Kansas, which is corroborated by contemporaneous local news stories.

Extending a widely used dynamic discrete choice model to include Bayesian updating and unobserved types has many empirical applications beyond our setting, including insurance markets, consumer behavior, restaurants learning about their local demand for their food, plant replacement, teachers learning about their teaching ability before getting tenure, and investments that are made over a long time period, such as research and development expenses. Specifically, it can be applied in settings such as career choices that depend on beliefs that may be biased and workers who update their beliefs as they learn about their skills over time (Roy (1951)), minority students learning about their higher education opportunities (Hoxby and Turner (2015)), students learning about returns to college majors (Arcidiacono et al. (2010) and Wiswall and Zafar (2011)), and workers learning about job mobility (Topel and Ward (1988)).

8 References

- Akerberg, D. (2003): “Advertising, Learning, and Consumer Choice in Experience Good Markets: A Structural Examination.” *International Economic Review*, 44, 1007–1040.
- Allen, Douglas W. (2019): “Establishing economic property rights by giving away an Empire,” *The Journal of Law and Economics*. 62(2), 251–280.
- Allen, Douglas W. (1991): “Homesteading and Property Rights; Or, “How the West Was Really Won.”” *The Journal of Law and Economics*, 34(1), 1-23.
- Allen, Douglas W. and Bryan Leonard (2019): “Rationing by racing and the Oklahoma land rushes.” *Journal of Institutional Economics*, 16(2), 127-144.
- Allen, Douglas W. and Bryan Leonard (2020): “How many rushed during the Oklahoma land openings?” *Cliometrica*, 14, 397-416.
- Allen, Douglas W. and Bryan Leonard (2021): “Property Rights and Path Dependence: 19th Century Land Policy and Modern Economic Outcomes.” *The Economic Journal*, 131(640), 3073-3102.
- Anderson, Terry L. and Peter J. Hill (1990): “The Race for Property Rights.” *The Journal of Law and Economics*, 33(1), 177-197.
- Atack, Jeremy (2013): “On the Use of Geographic Information Systems in Economic History: The American Transportation Revolution Revisited.” *The Journal of Economic History*, 73(2), 313-338.
- Atack, Jeremy (2016): “Historical Geographic Information Systems (GIS) database of U.S. Railroads for 1850 to 1903.”
- Atack, Jeremy, Fred Batemen, and Robert A. Margo (2010): “Did Railroads Induce or Follow Economic Growth?” *Social Science History*, 34(2), 171-197.
- Banach, S. (1922): “Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales.” *Fundamenta Mathematicae*, 3: 133–181.
- Bowden, R. J. (1992): “Competitive Selection and Market Data: The Mixed-Index Problem.” *The Review of Economic Studies*, 59(3), 625–633.
- Covert, Thomas (2015): “Experiential and Social Learning in Firms: The Case of Hydraulic Fracturing in the Bakken Shale.”
- Crawford, Gregory S. and Matthew Shum (2005): “Uncertainty and Learning in Pharmaceutical Demand.” *Econometrica*, 73(4), 1137-1173.
- Darden, M. (2017): “Smoking, Expectations, and Health: A Dynamic Stochastic Model of

- Lifetime Smoking Behavior,” *Journal of Political Economy*, 125: 1245-1721.
- Duflo, Ester, Michael Kremer, and Jonathan Robinson (2008). “How high are rates of return to fertilizer?” Evidence from field experiments in Kenya. *American Economic Review*, 98(2), 482-488.
- are so poor (and what can be done about it).” *Great Plains Quarterly*, 28(3), 181-190.
- Erdem, Tulin and Michael Keane (1996). “Decision-Making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets.” *Marketing Science*, 15(1), 1-20.
- Field, Erica (2007). “Entitled to Work: Urban Property Rights and Labor Supply in Peru.” *The Quarterly Journal of Economics*, 122(4), 1561-1602.
- Field, Alfred J., Erica Field, and Maximo Torero (2006). “Property Rights and Crop Choice in Rural Peru, 1994-2004.” *International Food and Policy Research Institute*.
- Fudenberg, D. and J. Tirole (1991): “Perfect Bayesian Equilibrium and Sequential Equilibrium.” *Journal of Economic Theory* 53: 236–260.
- Gates, Paul W. (1936): “The Homestead Law in an Incongruous Land System.” *The American Historical Review*, 41(4), 652-681.
- Gates, Paul W. (1954): *Fifty Million Acres: Conflicts of Kansas Land Policy, 1854-1890*. Cornell University Press, Ithica, NY.
- Gates, Paul W. (1968): *History of Public Land Law Development*. The Public Land Law Review Commission, Washington D.C..
- Gates, Paul W. (1977): “Homesteading in the High Plains.” *Agricultural History*, 51(1): 109-133.
- Gentzkow, Matthew, Jesse M. Shapiro, and Michael Sinkinson 2014: “Competition and Ideological Diversity: Historical Evidence from US Newspapers.” *American Economic Review*, 104 (10): 3073-3114.
- Ham, J., and T. Woutersen (2022): “Confidence Sets for Continuous and Discontinuous Functions of Estimated Parameters,” working paper.
- Hansen, Zeynep K. and Gary D. Libecap (2004): “Small Farms, Externalities, and the Dust Bowl of the 1930s.” *Journal of Political Economy*, 112(3), 665-694.
- Hausman, J. A. and T. Woutersen (2014): “Estimating a Semi-Parametric Duration Model without Specifying Heterogeneity.” *Journal of Econometrics*, 178: 114-131.
- heterogeneous-agent incomplete-insurance economies.” *Journal of Economic Dynamics and*

- Control*, 17: 953-969.
- Hodgson, Charles (2019). "Information Externalities, Free Riding, and Optimal Exploration in the UK Oil Industry."
- Kamenica, Emir and Matthew Gentzkow (2011). "Bayesian persuasion." *American Economic Review*, 101(6), 2590-2615.
- Limerick, Patricia N., Charles Rankin, and Clyde A. Milner II (1991): *Trails: Toward a New Western History*. University Press of Kansas, Lawrence, KS.
- Malin, James C. (1946). "Dust Storms: 1850-1900." *Soil Science*, 63(3), 265-296.
- Martin, George W. (1870). *Union*. Junction City, Kansas, May 21, 1870.
- Mattheis, Ross and Itzchak Tzachi Raz (2021): "There's No Such Thing As Free Land: The Homestead Act and Economic Development."
- McFadden, D. (1973): *Conditional logit analysis of qualitative choice behavior*. In *Frontiers in Econometrics*, ed. P. Zarembka. Academic Press, New York.
- McFadden, D. (1978): *Modeling the choice of residential location*. *Transportation Research Record*, 672, 72-77.
- Montero, Eduardo (2022). "Cooperative property rights and development: Evidence from land reform in El Salvador." *Journal of Political Economy*, 130(1), 48-93.
- National Park Service (2021): "Homesteading by the Numbers." United States Department of the Interior.
- Nevo, Aviv, John L. Turner, and Jonathan W. Williams (2016), "Usage-Based Pricing and Demand for Residential Broadband." *Econometrica*, 84: 411-443.
- Pakes, Ariel, Jack Porter, Kate Ho, Joy Ishii (2015): "Moment inequalities and their application." *Econometrica*, 83, 315-334.
- Peck, William J. (1967): *The Palimpsest*. The State Historical Society of Iowa, Iowa City, IA.
- Business Cycle Measurement." *Carnegie-Rochester Conference Series on Public Policy*, 25, 11-44
- Ridder, G., and T. Woutersen (2003): "The Singularity of the Information Matrix of the Mixed Proportional Hazard Model," *Econometrica*, Vol. 71, 5: 1579-1589.
- Rust, John (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica*, 55(5), 999-1033.
- Shannon, Fred A. (1966): *The Farmer's Last Frontier: Agriculture, 1860 - 1897*. Holt,

Rinehart and Winston, New York, NY.

Smith, Cory (2020): “Land Concentration and Long-Run Development: Evidence from the Frontier United States.”

Stange, Kevin M. (2012): “An Empirical Investigation of the Option Value of College Enrollment.” *American Economic Journal: Applied Economics*, 4(1), 49-84.

Steck, Andrew (2022): “Industry Dynamics with Social Learning: Evidence from Hydraulic Fracturing.”

Watson, Joel (2017): “A General, Practicable Definition of Perfect Bayesian Equilibrium.” *Economics*.

A Appendix

Table 2: Parameter Estimates from the Bayesian and Full Information Models

	Bayesian Model	Full Information Model
κ	0.00057 (0.000049)	0.00057 (0.000049)
$Sell_{notitle}$	0.73242 (0.063152)	0.73242 (0.063152)
μ_{home} (homesteader ability)		0.36343 (0.016349)
$\mu_{purchase}$ (purchaser ability)		0.92475 (0.064923)
μ_A (high type ability)	1.24934 (0.088922)	
μ_B (low type ability)	1.23749 (0.081451)	
η_A (high type bias)	-1.44204 (0.146520)	
η_B (low type bias)	-1.50232 (0.114041)	
$V_{A,final}$ (high type final value)	7.62185 (0.366754)	
$V_{B,final}$ (low type final value)	7.54212 (0.342184)	
σ_{prior}	0.44308 (0.023400)	
δ	0.75942 (0.000516)	
Proportion of purchasers who are high ability	0.99000 (0.289456)	
Proportion of homesteaders who are high ability	0.01000 (0.009725)	
Observations	7,011	7,011

Standard errors in parentheses

Table 3: Number of High and Low Ability Purchasers and Homesteaders

	Purchasers	Homesteaders	Both
High Ability	3162	38	3200
Low Ability	32	3779	3811
Both	3194	3817	7,011

Table 4: Counterfactual Estimates of Farming without Homesteading

	Bayesian Model	Full Information Model
$1-\rho$ (dependence between purchase and HS)	0.41037 (0.098726)	0.41380 (0.098617)
γ (disutility of being outside the county)	-8.97116 (0.105182)	-8.96835 (0.066119)
Pr(no farming no homesteading)	0.68931 (0.086012)	0.68189 (0.085881)
Observations	7,011	7,011

Standard errors in parentheses

Table 5: Counterfactual Estimates of Farming Under Different Beliefs

	Bayesian Model	$t = 10$ Counterfact.	$t = 20$ Counterfact.
Pr(abandon) in first 5 years	0.06974 (0.006587)	0.06218 (0.003096)	0.02645 (0.006497)
Observations	7,011	7,011	7,011

Standard errors in parentheses

Table 6: Parameter Estimates from the Agricultural Census Data

	Bayesian Model using Census Data
κ	0.00057 (0.000049)
$\text{Sell}_{notitle}$	0.73242 (0.063152)
μ_A (high type ability)	1.24934 (0.088922)
μ_B (low type ability)	1.21112 (0.008018)
η_A (high type bias)	-1.17185 (0.210966)
η_B (low type bias)	-1.12638 (0.207969)
$V_{A,final}$ (high type final value)	7.62185 (0.366754)
$V_{B,final}$ (low type final value)	7.48990 (0.388688)
σ_{prior}	0.37821 (0.063521)
δ	0.76035 (0.001094)
Proportion of purchasers who are high ability	1.00000 (0.572111)
Proportion of homesteaders who are high ability	0.01000 (0.009434)
Pr(no farming no homesteading)	0.68245 (0.086300)
Pr(abandon) in first 5 years	0.06758 (0.003647)
Observations	7,011

Standard errors in parentheses

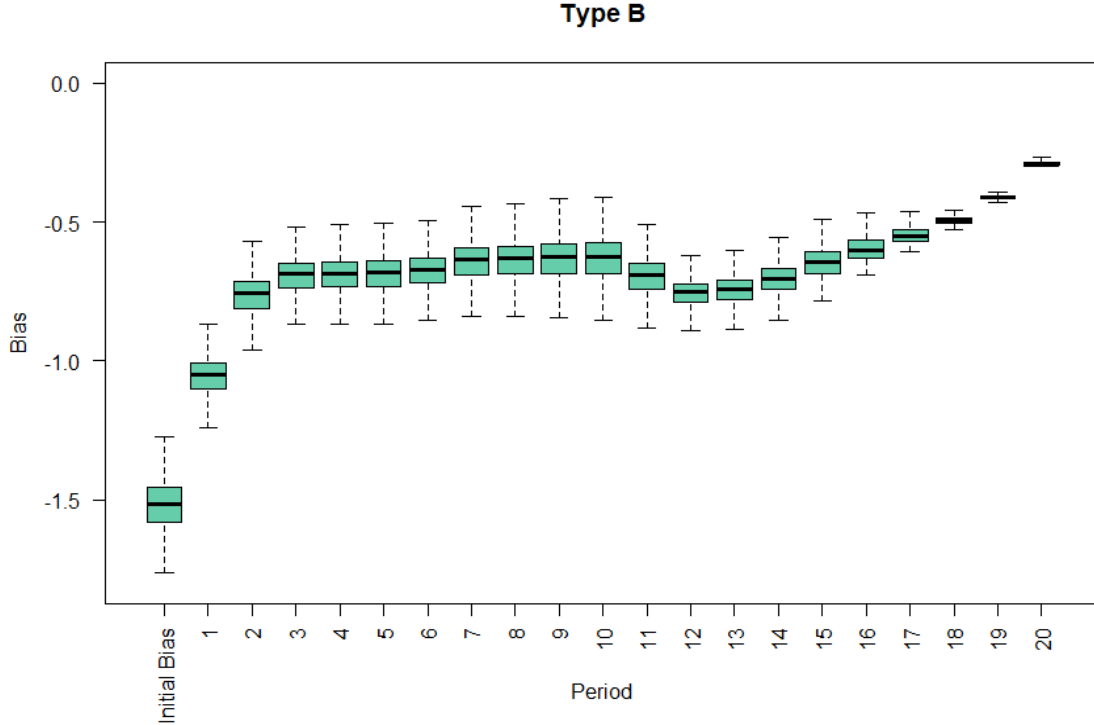


Figure 7: The bias of type B decreases toward zero over the 20 periods; the bias of type A is slightly smaller

A.1 Proof of Lemma 1: Bayesian Plausibility

Kamenica and Gentzkow (2011) introduce the concept of Bayesian Persuasion and make the assumption of Bayesian plausibility. They state that a distribution of posteriors is Bayes plausible if the expected posterior probability equals the prior. We now show that this property holds in our model. As in the main text, the random variable X_t is observed at time t by the farmer. The farmer observes such a shock every period and let these shocks be i.i.d. over time. As in the text, $X_t|\mu$ is normally distributed with mean μ and variance σ_{Data}^2 for all t . The parameter μ is unknown and let the farmer's prior on μ be normally distributed with mean $\bar{\mu} + \eta$ and variance σ_{Prior}^2 where $\bar{\mu}$ is the true value of μ and η is the bias. After observing X_1 , the farmer update their prior. If X_1 is not observed then the farmer could derive the predictive distribution for X_1 . In particular, the predictive distribution is the $X_1|\mu$ integrated with respect to $\pi_{t=0}(\mu)$, the prior on μ at $t = 0$. This yields a normal distribution with mean $\bar{\mu} + \eta$ and variance $\sigma_{\text{Prior}}^2 + \sigma_{\text{Data}}^2$. Integrating the updated prior $\pi(\mu|X_1)$ with respect to this predictive distribution yields $\pi_{t=0}(\mu)$. See the

online appendix for more details.

A.2 Proof of Lemma 2: Contraction Mapping and Geometric Convergence

We first show that $V_{mt}(V)$ is a contraction mapping and that the conditions for Banach's (1922) fixed point are satisfied. Consider

$$\begin{aligned} V_{mt}(V) = E[& \sup_{d_{it} \in \{C,S,A\}} [1(d_{it} = C) \cdot \{\varepsilon_{C,it} + \bar{\pi}_{mt} + \delta V\} + \\ & + 1(d_{it} = S) \cdot (Sell_{it} + \varepsilon_{S,it}) + 1(d_{it} = A) \cdot \varepsilon_{A,it}]]. \end{aligned}$$

Note that $V_{mt}(V)$ is non-decreasing in its argument and consider \dot{V} and \ddot{V} . Without loss of generality we assume that $\dot{V} \geq \ddot{V}$. This gives

$$V_{mt}(\dot{V}) - V_{mt}(\ddot{V}) \leq \delta(\dot{V} - \ddot{V})$$

so $V_{mt}(V)$ is a contraction mapping. Further, the set $[0, \bar{V}]$ is non-empty, since $\bar{V} > 0$, and is bounded, since $\bar{V} < \infty$. Finally, the set $[0, \bar{V}]$ is convex by inspection. Thus, the assumptions of Banach's (1922) fixed point are satisfied, so the solution to the equation $V_{mt}(V) = V$ is unique.

Next, we show that the rate of convergence is geometric in the number of iterations. As in the text, $0 \leq \delta < 1$ and $V_{mt,\text{fixed}}$ denotes the fixed point for all m and t , i.e., $V_{mt}(V_{mt,\text{fixed}}) = V_{mt,\text{fixed}}$ for all m and t . Let $V = V_{mt,\text{fixed}} + \zeta$, and first consider $\zeta > 0$, so V is larger than $V_{mt,\text{fixed}}$. This gives

$$\begin{aligned} V_{mt}(V_{mt,\text{fixed}} + \zeta) &= E[\sup_{d_{it} \in \{C,S,A\}} [1(d_{it} = C) \cdot \{\varepsilon_{C,it} + \bar{\pi}_{mt} + \delta(V_{mt,\text{fixed}} + \zeta)\} + \\ & + 1(d_{it} = S) \cdot (Sell_{it} + \varepsilon_{S,it}) + 1(d_{it} = A) \cdot \varepsilon_{A,it}]] \\ &\leq V_{mt,\text{fixed}} + \delta\zeta. \end{aligned}$$

Further, note that $V_{mt}(V_{mt,\text{fixed}} + \zeta)$ is non-decreasing in ζ . This gives

$$0 \leq V_{mt}(V_{mt,\text{fixed}} + \zeta) - V_{mt,\text{fixed}} \leq \delta\zeta$$

where $\zeta > 0$. Thus,

$$|V_{mt}(V_{mt,\text{fixed}} + \zeta) - V_{mt,\text{fixed}}| \leq \delta|\zeta|$$

where $\zeta > 0$. Similarly, for the second case where $\zeta < 0$, we have $V_{mt}(V_{mt,\text{fixed}} + \zeta) \leq V_{mt,\text{fixed}}$ and $V_{mt}(V_{mt,\text{fixed}} + \zeta) \geq V_{mt,\text{fixed}} + \delta\zeta$. Thus,

$$V_{mt,\text{fixed}} + \delta\zeta \leq V_{mt}(V_{mt,\text{fixed}} + \zeta) \leq V_{mt,\text{fixed}}.$$

Subtracting V_{mt} yields

$$\delta\zeta \leq V_{mt}(V_{mt,\text{fixed}} + \zeta) - V_{mt,\text{fixed}} \leq 0.$$

This gives

$$|V_{mt}(V_{mt,\text{fixed}} + \zeta) - V_{mt,\text{fixed}}| \leq \delta|\zeta|$$

for any ζ so that $V_{mt}(V)$ is a contraction mapping. Another way to write the last equation is

$$V_{mt}(V_{mt,\text{fixed}} + \zeta) = V_{mt,\text{fixed}} + \zeta^* \text{ where } \zeta^* \leq \delta|\zeta|$$

Now consider the starting value $V_{R=1}$ and let

$$\begin{aligned} V_{R=2} &= V_{mt}(V_{R=1}) \text{ and} \\ V_{R=r} &= V_{mt}(V_{R=r-1}) \text{ for } r > 1. \end{aligned}$$

Further define

$$\zeta_r = V_{R=r} - V_{mt,\text{fixed}}.$$

Note that

$$\begin{aligned} V_{R=2} &= V_{mt}(V_{R=1}) = V_{mt,\text{fixed}} + \zeta_2 \text{ where } |\zeta_2| \leq \delta|V_{R=1} - V_{mt,\text{fixed}}| = \delta|\zeta_1|, \\ V_{R=3} &= V_{mt}(V_{R=2}) = V_{mt,\text{fixed}} + \zeta_3 \text{ where } |\zeta_3| \leq \delta|\zeta_2| \leq \delta^2|\zeta_1|. \end{aligned}$$

Repeating these steps yields

$$V_{R=r} = V_{mt}(V_{R=r-1}) = V_{mt,\text{fixed}} + \zeta_r \text{ where } |\zeta_r| \leq \delta|\zeta_{r-1}| \leq \delta^r|\zeta_1| \text{ and } r > 1, \text{ and}$$

$$|V_{mt,R=r} - V_{mt,\text{fixed}}| \leq \delta^r |V_{mt,R=1} - V_{mt,\text{fixed}}| \text{ where } r > 1.$$

A.3 Counterfactuals from the Full Information Model

Because of the scale of the Homestead Act, an important policy question is what would have happened to land acquisition if the U.S. government had not allowed homesteading and had only sold the land? Specifically, what fraction of homesteaders would have chosen to acquire land if their only method for doing so was purchasing it? We previously discussed this counterfactual under the assumptions of our Bayesian model. Here, we also provide a counterfactual under a full information case.

In the full information case without unobserved types, the nested logit model for calculating the counterfactual is given by

$$\Pr(\text{buy}|\text{homestead or buy}) = \frac{1}{1 + \exp\{\kappa \cdot \Psi/\rho\}}, \text{ and } 0 < \rho < 1, \kappa > 0.$$

$$\Pr(\text{no farming}) = \frac{1}{1 + [\exp(\gamma/\rho) + \exp\{\gamma/\rho - \kappa \cdot \Psi/\rho\}]^\rho}, \text{ and } 0 < \rho < 1, \kappa > 0.$$

More information on the calculation of this counterfactual in the full information and Bayesian settings is found in the online appendix.

The disutility to farm in the counties in Kansas under consideration for people outside that area is large, $\hat{\gamma} = -8.968$ (s.e. $6.6 \cdot 10^{-2}$), and homesteading and buying are not perfect substitutes, $(1 - \hat{\rho}) = 0.414$ (s.e. $9.9 \cdot 10^{-2}$) in a full information setting, similar to their substitutability in the Bayesian setting. The disutility is comparable to \$15,733.33 (in 1870 dollars). These parameter estimates are shown in Table 4. The parameter estimates yield the following counterfactual: If the U.S. government had not offered homesteading, then 32% of the homesteaders would have bought the land and 68% would opt not to farm. The standard error of this counterfactual is $8.6 \cdot 10^{-2}$ and the 95% confidence interval¹⁶ for homesteaders

¹⁶Ham and Woutersen (2022) propose methods for confidence sets for nonlinear counterfactual functions and we checked the validity conditions. We use 100,000 nonparametric bootstrap replications to calculate the standard errors in this subsection.

opting not to farm is [50.8%, 85.2%]. Results are given in Table 5 of the appendix.

The estimates from the Bayesian model and the full information model are quite similar for γ , the disutility to farm in these counties in Kansas given that an individual is not currently living in these counties. In both the Bayesian and full information models, about two-thirds of the homesteaders choose not to farm in a situation without homesteading. However, the heterogeneous types in the Bayesian model may better capture what a marginal farmer, i.e., a low ability farmer, would do.

B Online Appendix

B.1 Matching Method

Matching between the tract books and the resale deeds does not necessarily create a one-to-one matching. Individual i may acquire several adjacent pieces of land (multiple tract book observations) but combine them as one farm and sell them as a single unit (one deed observation). Conversely, individual i may acquire one piece of land (one tract book observation) but split it into multiple pieces when selling it (multiple deed observations). A matching like this one is inherently order-dependent, but the fact that these matches are many-to-many decreases this problem significantly. We further outline how we decrease the order-dependency below.

In order to provide a matching between the tract books (initial land acquisition) and the deeds (land resale), we use a multi-step approach. First, we remove abandoned homesteads from the dataset to be matched because abandoned homesteads by definition were not resold. Second, we locate perfect matches between the initial land acquisition data and the resale deeds and exclude these observations from the datasets to be matched using the process below. Perfect matches are defined by observations with the exact same first, middle, and last name and the exact same location (township-range, section, and aliquot) in the tract book data and the deed data. In our dataset, about 10% of the tract books match perfectly to deeds.

Second, we use the Stata modules *matchit* and *soundex*, which join two datasets based on a string variable which does not necessarily need to be exactly the same: they are fuzzy matching tools. The *matchit* module matches based on character similarities and the *soundex* module matches based on characters with similar sounds. We create a string variable in the tract book data which combines the name (first, middle, and last) and the location (township-range, section, and aliquot). Similarly, we create a string variable in the deeds which combines the name (first, middle, and last) and the location (township-range, section, and aliquot). Stata's *matchit* and *soundex* modules create a many-to-many matching and assign each match a similarity score between 0 and 1 based on the similarity of the two string variables. Similarity scores of 1 indicate a perfect match, and *matchit* and *soundex* automatically drop any matches less than 0.5 and allow the user to create a cutoff above that. Any cutoff still results in a many-to-many match for the reasons described

earlier. We use an initial cutoff of 0.7, and then hand-match the data as described below.

To obtain the best matches, we search through each potential match several times manually and delete incorrect matches, becoming stricter each time to reduce endogeneity in the matching. Matches are evaluated based on the name, location, acres, and date of acquisition and date of resale. Because the data were digitized from handwritten, historical records, hand-matching allows us to utilize matches that would be difficult for a computer to make. Hand-matching is still regarded to be the gold standard for historical datasets under 15,000 observations. Finally, each match in the full set of matches is manually evaluated against the unmatched tract book observations, the unmatched deed observations, and the other matched observations to determine if it is the best possible match. This matching method provides a dataset containing 7,011 observations (all the tract book observations before 1874).

We use a similar method to match the BLM tract books and the Kansas agricultural censuses (and to the U.S. population censuses and soil quality data). Not every farm was recorded in the Kansas agricultural census, so we match 1,471 census records to the 7,011 BLM observations. We use this smaller dataset only to estimate production results as a robustness check, not our Bayesian model.

B.2 Value Functions in the Full Information Model

The value functions for both homesteaders and purchasers depend on the values in the next period. We first consider homesteaders, and for a given $Sell_{it}$ we have

$$V_{home,t} = E\left[\sup_{d_{it} \in \{C,S,A\}} [1(d_{it} = C) \cdot \{\varepsilon_{C,it} + \mu + \delta V_{home,t+1}\} + 1(d_{it} = S) \cdot (Sell_{it} + \varepsilon_{S,it}) + 1(d_{it} = A) \cdot \varepsilon_{A,it}] \right], \text{ where } t = 1, 2, \dots \quad (6)$$

In our application we assume that the random shocks $\varepsilon_{C,it}$, $\varepsilon_{S,it}$, and $\varepsilon_{A,it}$ given $\bar{\pi}_{mt}$ and X_t are realizations from a type I extreme value distribution.¹⁷ This gives

$$V_{home,t} = \ln(e^{\mu + \delta V_{home,t+1}} + e^{Sell_{it}} + 1) + 0.577, \text{ where } t = 1, 2, \dots \quad (7)$$

¹⁷The cumulative distribution function of this extreme value distribution is $F(\varepsilon) = \exp(-\exp(-\varepsilon))$.

From period $t = 10$ onward the homesteaders have the title to the farm. For the full information model only, we assume that the sale price for a farm that has the title is the same across individuals and time periods. This gives

$$V_{home,t} = \ln(e^{\mu+\delta V_{home,t+1}} + e^{Sell} + 1) + 0.577, \text{ where } t = 10, \dots \quad (8)$$

Purchasers have the title to the land from the beginning so the last equation holds for them for all periods (without the subscript *home*). Next define the function $V_{full}(V)$ as

$$V_{full}(V) = \ln(e^{\mu+\delta V} + e^{Sell} + 1) + 0.577.$$

Below we show that $V_{full}(V)$ is a contraction mapping.

B.2.1 Contraction Mapping Full Information Model

For the full information model we have the function

$$V_{full}(V) = \ln(e^{\mu+\delta V} + e^{Sell} + 1) + 0.577.$$

The function $V_{full}(V)$ is a contraction mapping for $0 \leq \delta < 1$. To see this, note that $V_{full}(V)$ is non-decreasing in its argument and consider \dot{V} and \ddot{V} . First, let $\dot{V} \geq \ddot{V}$,

$$\begin{aligned} V_{full}(\dot{V}) - V_{full}(\ddot{V}) &= \ln(e^{\mu+\delta\dot{V}} + e^{Sell} + 1) - \ln(e^{\mu+\delta\ddot{V}} + e^{Sell} + 1) \\ &= \ln\left(\frac{e^{\mu+\delta\dot{V}} + e^{Sell} + 1}{e^{\mu+\delta\ddot{V}} + e^{Sell} + 1}\right) \\ &= \ln\left(\frac{e^{\mu+\delta\ddot{V}+\delta(\dot{V}-\ddot{V})} + e^{Sell} + 1}{e^{\mu+\delta\ddot{V}} + e^{Sell} + 1}\right) \\ &\leq \ln\left\{\frac{e^{\mu+\delta\ddot{V}+\delta(\dot{V}-\ddot{V})} + e^{Sell+\delta(\dot{V}-\ddot{V})} + e^{\delta(\dot{V}-\ddot{V})}}{e^{\mu+\delta\ddot{V}} + e^{Sell} + 1}\right\} \end{aligned}$$

since $e^{\delta(\dot{V}-\ddot{V})} \geq 1$. This yields

$$\begin{aligned} V_{full}(\dot{V}) - V_{full}(\ddot{V}) &\leq \ln\left\{e^{\delta(\dot{V}-\ddot{V})} \left(\frac{e^{\mu+\delta\ddot{V}} + e^{Sell} + 1}{e^{\mu+\delta\ddot{V}} + e^{Sell} + 1}\right)\right\} \\ &= \delta(\dot{V} - \ddot{V}). \end{aligned}$$

The argument on the left hand side and on the right hand side are both non-negative. Thus, taking the absolute values on both sides is straightforward and yields

$|V_{full}(\dot{V}) - V_{full}(\ddot{V})| \leq \delta|\dot{V} - \ddot{V}|$ for $\dot{V} \geq \ddot{V}$. Next, consider $\dot{V} < \ddot{V}$ and note that \dot{V} and \ddot{V} are arbitrary values so that, after relabeling, we have $\dot{V} > \ddot{V}$ so the argument above applies and we have $|V_{full}(\dot{V}) - V_{full}(\ddot{V})| \leq \delta|\dot{V} - \ddot{V}|$ for any \dot{V} and \ddot{V} . Therefore, $V_{full}(V)$ is a contraction mapping since $0 \leq \delta < 1$.

B.3 Value Functions in the Bayesian Model

The value functions for both homesteaders and purchasers depend on the values in the next period. Note that if individual i is of type m then $\bar{\pi}_{it} = \bar{\pi}_{mt}$. We first consider homesteaders and for a given $Sell_{it}$ we have

$$V_{home,mt} = E\left[\sup_{d_{it} \in \{C,S,A\}} [1(d_{it} = C) \cdot \{\varepsilon_{C,it} + \bar{\pi}_{mt} + \delta V_{home,m,t+1}\} + \right. \quad (9)$$

$$\left. + 1(d_{it} = S) \cdot (Sell_{it} + \varepsilon_{S,it}) + 1(d_{it} = A) \cdot \varepsilon_{A,it} \right],$$

where $m \in \{A, B\}$. When the error terms $\varepsilon_{C,it}$, $\varepsilon_{S,it}$, and $\varepsilon_{A,it}$ given $\bar{\pi}_{mt}$ and X_t are independent realizations from a type I extreme value distribution, then we can write $V_{home,mt}$ as

$$V_{home,mt} = \ln(e^{\bar{\pi}_{mt} + \delta V_{home,m,t+1}} + e^{Sell_{it}} + 1) + 0.577,$$

where $m \in \{A, B\}$. In the text, we discuss the stationarity assumption, $E(Sell_{i,t+s}) = Sell_{it}$ for all $s \geq 0$, and this assumption yields

$$V_{home,mt} = \ln(e^{\bar{\pi}_{mt} + \delta V_{home,m,t+1}} + e^{Sell_{it}} + 1) + 0.577,$$

where $m \in \{A, B\}$. The last equation also holds for purchasers when we change the *home* subscript. Next, consider the function $V_{mt}(V)$,

$$V_{mt}(V) = E\left[\sup_{d_{it} \in \{C,S,A\}} [1(d_{it} = C) \cdot \{\varepsilon_{C,it} + \bar{\pi}_{mt} + \delta V\} + \right. \quad (10)$$

$$\left. + 1(d_{it} = S) \cdot (Sell_{it} + \varepsilon_{S,it}) + 1(d_{it} = A) \cdot \varepsilon_{A,it} \right],$$

where $m \in \{A, B\}$. When the error terms $\varepsilon_{C,it}$, $\varepsilon_{S,it}$, and $\varepsilon_{A,it}$ given $\bar{\pi}_{mt}$ and X_t are independent realizations from a type I extreme value distribution, we can write $V_{mt}(V)$ as

$$V_{mt}(V) = \ln(e^{\bar{\pi}_{mt} + \delta V} + e^{Sell_{it}} + 1) + 0.577,$$

where $m \in \{A, B\}$. Below we show that $V_{mt}(V)$ is a contraction mapping. This function is relevant for purchasers for all periods and for homesteaders for $t = 10, 11, \dots$

A feature of our model that we did not use in our identification analysis is that Bayesian beliefs converge over time. In particular, $\bar{\pi}_{mt}$ converges to μ_m . This feature motivates grouping later periods together as the priors will be very similar. Specifically, we use the following fixed point

$$V_{m,final} = \ln\{\exp(\bar{\mu}_m + \delta V_{m,final}) + \exp(Sell_{it}) + 1\} + 0.577.$$

We use this fixed point in our analysis. In particular, we use the moments

$$g_{m, \text{fixed point}} = 4 \cdot [\ln\{\exp(\bar{\mu}_m + \delta V_{m,final}) + \exp(Sell_{notitle} + \kappa \cdot 200) + 1\} + 0.577 - V_{m,final}],$$

where $m \in \{A, B\}$. Further, dynamic selection causes that the type with higher ability, type A , is more likely to continue relative to the low ability type, type B . This feature yields that we only observe the high ability type at the end or that both types are about as able. Note that two types can be about as able while having different priors at $t = 0$. The reason that we do not use this feature in our identification analysis is that we follow the farmers for a very long time while in other applications the individuals or firms may be followed for a shorter period. We use this feature of our model in our empirical analysis.

B.4 Probabilities for Selling and Abandoning

Here we detail the value functions and probabilities for continuing, selling, and abandoning for homesteaders of type m . The value function is

$$V_{home,mt} = \ln[\exp\{\bar{\pi}_{mt} + \delta V_{home,mt}\} + \exp\{Sell_{notitle} + 1(t \geq 10)\kappa \cdot 200\} + 1] + 0.577.$$

This value function gives us the following probabilities for continuing in periods $t = \{1, \dots, 20\}$,

$$\Pr_{m,home}(d_{it} = C) = 1 - \frac{1 + \exp\{Sell_{notitle} + 1(t \geq 10)\kappa \cdot 200\}}{1 + \exp\{\bar{\pi}_{mt} + \delta V_{mt}\} + \exp\{Sell_{notitle} + 1(t \geq 10)\kappa \cdot 200\}}.$$

The probability of selling in periods $t = \{1, \dots, 20\}$ is

$$\Pr_{m,home}(d_{it} = S) = \frac{\exp\{Sell_{notitle} + 1(t \geq 10)\kappa \cdot 200\}}{1 + \exp\{\bar{\pi}_{mt} + \delta V_{mt}\} + \exp\{Sell_{notitle} + 1(t \geq 10)\kappa \cdot 200\}}.$$

Similarly, the probability of abandoning the farm at time $t = \{1, \dots, 20\}$ for each type is calculated from the value functions as

$$\Pr_{m,home}(d_{it} = A) = \frac{1}{1 + \exp\{\bar{\pi}_{mt} + \delta V_{mt}\} + \exp\{Sell_{notitle} + 1(t \geq 10)\kappa \cdot 200\}}.$$

B.5 Proof of Lemma 1: Bayesian Plausibility

Kamenica and Gentzkow (2011) introduce the concept of Bayesian persuasion and make the assumption of Bayesian plausibility. They state that a distribution of posteriors is Bayes plausible if the expected posterior probability equals the prior. We now show that this property holds in our model. Let $\pi_{t=0}(\mu)$ denote the prior on μ at $t = 0$ and let $\pi_{t=0}(\mu)$ be the density of a normally distributed random variable with mean $\bar{\mu} + \eta$ and variance σ_{prior}^2 . The random variable X_t is observed at time t by the farmer and let X_t be normally distributed with mean μ and variance σ_{data}^2 . The farmer observes such a shock every period and let these shocks be i.i.d. over time. Let $\bar{\mu} + \eta$ be denoted by φ . This gives the following

prior predictive distribution for any t ,

$$\begin{aligned}
p(x_t) &= \int p(x_t|\mu)\pi_{t=0}(\mu)d\mu \\
&= \int \frac{1}{\sqrt{2\pi\sigma_{data}^2}} \exp\left\{-\frac{1}{2\sigma_{data}^2}(x_t - \mu)^2\right\} \cdot \frac{1}{\sqrt{2\pi\sigma_{prior}^2}} \exp\left\{-\frac{1}{2\sigma_{prior}^2}(\mu - \varphi)^2\right\} d\mu \\
&= \int \frac{1}{\sqrt{2\pi\sigma_{data}^2}} \frac{1}{\sqrt{2\pi\sigma_{prior}^2}} \exp\left[-\frac{1}{2\sigma_{data}^2\sigma_{prior}^2}\{(\sigma_{data}^2 + \sigma_{prior}^2)\mu^2 - 2\mu(\sigma_{prior}^2x_t + \sigma_{data}^2\varphi) \right. \\
&\quad \left. + \sigma_{prior}^2x_t^2 + \sigma_{data}^2\varphi^2\right\}]d\mu \\
&= \int \frac{1}{2\pi\sqrt{\sigma_{data}^2\sigma_{prior}^2}} \exp\left[-\frac{\sigma_{data}^2 + \sigma_{prior}^2}{2\sigma_{data}^2\sigma_{prior}^2}\{\mu^2 - 2\mu\frac{\sigma_{prior}^2x_t + \sigma_{data}^2\varphi}{\sigma_{data}^2 + \sigma_{prior}^2} + \frac{\sigma_{prior}^2x_t^2 + \sigma_{data}^2\varphi^2}{\sigma_{data}^2 + \sigma_{prior}^2}\right\}]d\mu \\
&= \int \frac{1}{2\pi\sqrt{\sigma_{data}^2\sigma_{prior}^2}} \exp\left[-\frac{\sigma_{data}^2 + \sigma_{prior}^2}{2\sigma_{data}^2\sigma_{prior}^2}\{\mu^2 - 2\mu\frac{\sigma_{prior}^2x_t + \sigma_{data}^2\varphi}{\sigma_{data}^2 + \sigma_{prior}^2} + \left(\frac{\sigma_{prior}^2x_t + \sigma_{data}^2\varphi}{\sigma_{data}^2 + \sigma_{prior}^2}\right)^2\right\}]d\mu \cdot \\
&\quad \cdot \exp\left[-\frac{\sigma_{data}^2 + \sigma_{prior}^2}{2\sigma_{data}^2\sigma_{prior}^2}\left\{-\left(\frac{\sigma_{prior}^2x_t + \sigma_{data}^2\varphi}{\sigma_{data}^2 + \sigma_{prior}^2}\right)^2 + \frac{\sigma_{prior}^2x_t^2 + \sigma_{data}^2\varphi^2}{\sigma_{data}^2 + \sigma_{prior}^2}\right\}\right] \\
&= \frac{1}{\sqrt{2\pi}\sqrt{\sigma_{data}^2 + \sigma_{prior}^2}} \int \frac{\sqrt{\sigma_{data}^2 + \sigma_{prior}^2}}{\sqrt{2\pi}\sqrt{\sigma_{data}^2\sigma_{prior}^2}} \exp\left[-\frac{\sigma_{data}^2 + \sigma_{prior}^2}{2\sigma_{data}^2\sigma_{prior}^2}\left\{\mu^2 - 2\mu\frac{\sigma_{prior}^2x_t + \sigma_{data}^2\varphi}{\sigma_{data}^2 + \sigma_{prior}^2} \right. \right. \\
&\quad \left. \left. + \left(\frac{\sigma_{prior}^2x_t + \sigma_{data}^2\varphi}{\sigma_{data}^2 + \sigma_{prior}^2}\right)^2\right\}\right]d\mu \cdot \exp\left[-\frac{\sigma_{data}^2 + \sigma_{prior}^2}{2\sigma_{data}^2\sigma_{prior}^2}\left\{-\left(\frac{\sigma_{prior}^2x_t + \sigma_{data}^2\varphi}{\sigma_{data}^2 + \sigma_{prior}^2}\right)^2 + \frac{\sigma_{prior}^2x_t^2 + \sigma_{data}^2\varphi^2}{\sigma_{data}^2 + \sigma_{prior}^2}\right\}\right].
\end{aligned}$$

Note that

$$\frac{\sqrt{\sigma_{data}^2 + \sigma_{prior}^2}}{\sqrt{2\pi}\sqrt{\sigma_{data}^2\sigma_{prior}^2}} \exp\left[-\frac{\sigma_{data}^2 + \sigma_{prior}^2}{2\sigma_{data}^2\sigma_{prior}^2}\left\{\mu^2 - 2\mu\frac{\sigma_{prior}^2x_t + \sigma_{data}^2\varphi}{\sigma_{data}^2 + \sigma_{prior}^2} + \left(\frac{\sigma_{prior}^2x_t + \sigma_{data}^2\varphi}{\sigma_{data}^2 + \sigma_{prior}^2}\right)^2\right\}\right]$$

is the density of a normal distribution so that it integrates to one. This gives

$$\begin{aligned}
p(x_t) &\propto \exp\left[-\frac{\sigma_{data}^2 + \sigma_{prior}^2}{2\sigma_{data}^2\sigma_{prior}^2} \left\{ -\left(\frac{\sigma_{prior}^2 x_t + \sigma_{data}^2 \varphi}{\sigma_{data}^2 + \sigma_{prior}^2}\right)^2 + \frac{\sigma_{prior}^2 x_t^2 + \sigma_{data}^2 \varphi^2}{\sigma_{data}^2 + \sigma_{prior}^2} \right\}\right] \\
&\propto \exp\left[-\frac{1}{2\sigma_{data}^2\sigma_{prior}^2} \left\{ -\frac{(\sigma_{prior}^2 x_t + \sigma_{data}^2 \varphi)^2}{\sigma_{data}^2 + \sigma_{prior}^2} + \sigma_{prior}^2 x_t^2 + \sigma_{data}^2 \varphi^2 \right\}\right] \\
&\propto \exp\left[-\frac{1}{2\sigma_{data}^2\sigma_{prior}^2} \frac{1}{\sigma_{data}^2 + \sigma_{prior}^2} \left\{ -(\sigma_{prior}^4 x_t^2 + 2x_t \sigma_{data}^2 \sigma_{prior}^2 \varphi + \sigma_{data}^4 \varphi^2) \right. \right. \\
&\quad \left. \left. + (\sigma_{data}^2 \sigma_{prior}^2 + \sigma_{prior}^4) x_t^2 + (\sigma_{data}^4 + \sigma_{data}^2 \sigma_{prior}^2) \varphi^2 \right\}\right] \\
&\propto \exp\left[-\frac{1}{2\sigma_{data}^2\sigma_{prior}^2} \frac{1}{\sigma_{data}^2 + \sigma_{prior}^2} \left\{ (\sigma_{data}^2 \sigma_{prior}^2) x_t^2 - 2x_t \sigma_{data}^2 \sigma_{prior}^2 \varphi \right\}\right] \\
&\propto \exp\left[-\frac{1}{2} \frac{1}{\sigma_{data}^2 + \sigma_{prior}^2} \{x_t^2 - 2x_t \varphi\}^2\right],
\end{aligned}$$

Thus, the posterior predictive distribution is normal with mean $\varphi = \bar{\mu} + \eta$ and variation $\sigma_{data}^2 + \sigma_{prior}^2$. The posterior after observing X_1 is the same as the prior conditional on X_1 and let this be denoted $\pi(\mu|X_1)$. Thus, we have

$$\begin{aligned}
\pi(\mu|X_1) &\propto p(X_1 = x_1|\mu)\pi_{t=0}(\mu) \\
&\propto \frac{1}{\sqrt{2\pi\sigma_{data}^2}} \exp\left\{-\frac{1}{2\sigma_{data}^2}(X_1 - \mu)^2\right\} \cdot \frac{1}{\sqrt{2\pi\sigma_{prior}^2}} \exp\left\{-\frac{1}{2\sigma_{prior}^2}(\mu - \varphi)^2\right\} \\
&\propto \exp\left[-\frac{1}{2\sigma_{data}^2\sigma_{prior}^2} \left\{ (\sigma_{data}^2 + \sigma_{prior}^2)\mu^2 - 2\mu(\sigma_{prior}^2 X_1 + \sigma_{data}^2 \varphi) + \sigma_{prior}^2 X_1^2 + \sigma_{data}^2 \varphi^2 \right\}\right] \\
&\propto \exp\left[-\frac{\sigma_{data}^2 + \sigma_{prior}^2}{2\sigma_{data}^2\sigma_{prior}^2} \left\{ \mu^2 - 2\mu \frac{\sigma_{prior}^2 X_1 + \sigma_{data}^2 \varphi}{\sigma_{data}^2 + \sigma_{prior}^2} + \frac{\sigma_{prior}^2 X_1^2 + \sigma_{data}^2 \varphi^2}{\sigma_{data}^2 + \sigma_{prior}^2} \right\}\right].
\end{aligned}$$

The last equation states that the posterior, or prior conditional on X_1 , is the density of the normal distribution with mean $\frac{\sigma_{prior}^2 X_1 + \sigma_{data}^2 \varphi}{\sigma_{data}^2 + \sigma_{prior}^2}$ and variance $\frac{\sigma_{data}^2 \sigma_{prior}^2}{\sigma_{data}^2 + \sigma_{prior}^2} = \frac{1}{\frac{1}{\sigma_{data}^2} + \frac{1}{\sigma_{prior}^2}}$. Next, we integrate out X_1 so that we obtain the unconditional prior. Note that $p(x_1) = p(x_t)$

because of the i.i.d. property. The prior $\pi_{t=1}(\mu)$ is the same as in the text. Note that

$$\begin{aligned}
\pi_{t=1}(\mu) &= \int \pi(\mu|x_1)p(x_1)dx_1 \\
&= \int \frac{1}{\sqrt{2\pi}} \sqrt{\frac{\sigma_{data}^2 + \sigma_{prior}^2}{\sigma_{data}^2 \sigma_{prior}^2}} \exp\left\{-\frac{\sigma_{data}^2 + \sigma_{prior}^2}{2\sigma_{data}^2 \sigma_{prior}^2} \left(\mu - \frac{\sigma_{prior}^2 x_1 + \sigma_{data}^2 \varphi}{\sigma_{data}^2 + \sigma_{prior}^2}\right)^2\right\} \\
&\quad \cdot \frac{1}{\sqrt{2\pi(\sigma_{data}^2 + \sigma_{prior}^2)}} \exp\left\{-\frac{1}{2} \frac{1}{\sigma_{data}^2 + \sigma_{prior}^2} (x_1 - \varphi)^2\right\} dx_1 \\
&= \int \frac{1}{\sqrt{2\pi}} \sqrt{\frac{\sigma_{data}^2 + \sigma_{prior}^2}{\sigma_{data}^2 \sigma_{prior}^2}} \exp\left[-\frac{1}{2} \left\{ \frac{\sigma_{data}^2 + \sigma_{prior}^2}{\sigma_{data}^2 \sigma_{prior}^2} \mu^2 - 2\mu \frac{x_1}{\sigma_{data}^2} - 2\mu \frac{\varphi}{\sigma_{prior}^2} \right. \right. \\
&\quad \left. \left. + \frac{1}{\sigma_{data}^2 \sigma_{prior}^2} \frac{(\sigma_{prior}^2 x_1 + \sigma_{data}^2 \varphi)^2}{\sigma_{data}^2 + \sigma_{prior}^2} \right\}\right] \\
&\quad \cdot \frac{1}{\sqrt{2\pi(\sigma_{data}^2 + \sigma_{prior}^2)}} \exp\left\{-\frac{1}{2} \frac{1}{\sigma_{data}^2 + \sigma_{prior}^2} (x_1 - \varphi)^2\right\} dx_1
\end{aligned}$$

This gives

$$\begin{aligned}
\pi_{t=1}(\mu) &\propto \int \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} \left\{ \frac{\sigma_{data}^2 + \sigma_{prior}^2}{\sigma_{data}^2 \sigma_{prior}^2} \mu^2 - 2 \frac{x_1}{\sigma_{data}^2} \left(1 - \frac{\sigma_{prior}^2}{\sigma_{data}^2 (\sigma_{data}^2 + \sigma_{prior}^2)}\right) 2x_1 \varphi \frac{\sigma_{data}^2}{\sigma_{prior}^2} \right. \right. \\
&\quad \left. \left. - 2\mu \frac{\varphi}{\sigma_{prior}^2} + \frac{x_1^2}{\sigma_{data}^2} + \frac{\sigma_{prior}^2}{\sigma_{data}^2 (\sigma_{data}^2 + \sigma_{prior}^2)} \frac{\sigma_{data}^4}{\sigma_{prior}^4} \varphi^2 \right\}\right] \\
&\quad \cdot \exp\left\{-\frac{1}{2} \frac{1}{\sigma_{data}^2 + \sigma_{prior}^2} (-2x_1 \varphi + \varphi^2)\right\} dx_1 \\
&\propto \int \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} \left\{ \frac{1}{\sigma_{data}^2} (x_1^2 - 2x_1 \mu + 2x_1 \frac{\sigma_{data}^2}{\sigma_{data}^2 + \sigma_{prior}^2} \varphi) + \frac{\sigma_{data}^2 + \sigma_{prior}^2}{\sigma_{data}^2 \sigma_{prior}^2} \mu^2 - 2\mu \frac{\varphi}{\sigma_{prior}^2} \right. \right. \\
&\quad \left. \left. + \frac{\sigma_{prior}^2}{\sigma_{data}^2 (\sigma_{data}^2 + \sigma_{prior}^2)} \frac{\sigma_{data}^4}{\sigma_{prior}^4} \varphi^2 \right\}\right] \cdot \exp\left(-\frac{1}{2} \frac{1}{\sigma_{data}^2 + \sigma_{prior}^2} \varphi^2\right) dx_1 \\
&\propto \int \frac{1}{\sqrt{2\pi \sigma_{data}^2}} \exp\left[-\frac{1}{2} \frac{1}{\sigma_{data}^2} \left(x_1 - \mu + \frac{\sigma_{data}^2}{\sigma_{data}^2 + \sigma_{prior}^2} \varphi\right)^2\right] dx_1 \\
&\quad \cdot \exp\left[-\frac{1}{2} \left\{ -\frac{1}{\sigma_{data}^2} \left(\mu - \frac{\sigma_{data}^2}{\sigma_{data}^2 + \sigma_{prior}^2} \varphi\right)^2 \right. \right. \\
&\quad \left. \left. + \frac{\sigma_{data}^2 + \sigma_{prior}^2}{\sigma_{data}^2 \sigma_{prior}^2} \mu^2 - 2\mu \frac{\varphi}{\sigma_{prior}^2} + \frac{1}{(\sigma_{data}^2 + \sigma_{prior}^2)} \frac{\sigma_{data}^2}{\sigma_{prior}^2} \varphi^2 \right\}\right] \cdot \exp\left(-\frac{1}{2} \frac{1}{\sigma_{data}^2 + \sigma_{prior}^2} \varphi^2\right)
\end{aligned}$$

since $\frac{1}{\sqrt{2\pi\sigma_{data}^2}} \exp[-\frac{1}{2}\frac{1}{\sigma_{data}^2}(x_1 - \mu + \frac{\sigma_{data}^2}{\sigma_{data}^2 + \sigma_{prior}^2}\varphi)^2]$ is a density and, therefore, integrates to one. This gives

$$\begin{aligned}
\pi_{t=1}(\mu) &= \int \frac{1}{\sqrt{2\pi}} \sqrt{\frac{\sigma_{data}^2 + \sigma_{prior}^2}{\sigma_{data}^2 \sigma_{prior}^2}} \exp[-\frac{1}{2}\{\frac{\sigma_{data}^2 + \sigma_{prior}^2}{\sigma_{data}^2 \sigma_{prior}^2} \mu^2 - 2\mu \frac{x_1}{\sigma_{data}^2} - 2\mu \frac{\varphi}{\sigma_{prior}^2} \\
&\quad + \frac{1}{\sigma_{data}^2 \sigma_{prior}^2} \frac{\sigma_{prior}^4 x_1^2 + 2x_1 \varphi \sigma_{data}^2 \sigma_{prior}^2 + \sigma_{data}^4 \varphi^2}{\sigma_{data}^2 + \sigma_{prior}^2}\}] \\
&\quad \cdot \frac{1}{\sqrt{2\pi(\sigma_{data}^2 + \sigma_{prior}^2)}} \exp\{-\frac{1}{2}\frac{1}{\sigma_{data}^2 + \sigma_{prior}^2}(x_1 - \varphi)^2\} dx_1 \\
&= \int \frac{1}{\sqrt{2\pi}} \sqrt{\frac{\sigma_{data}^2 + \sigma_{prior}^2}{\sigma_{data}^2 \sigma_{prior}^2}} \exp[-\frac{1}{2}\{\frac{\sigma_{data}^2 + \sigma_{prior}^2}{\sigma_{data}^2 \sigma_{prior}^2} \mu^2 - 2\mu \frac{x_1}{\sigma_{data}^2} - 2\mu \frac{\varphi}{\sigma_{prior}^2} \\
&\quad + \frac{\sigma_{prior}^2}{\sigma_{data}^2(\sigma_{data}^2 + \sigma_{prior}^2)}(x_1^2 + 2x_1 \varphi \frac{\sigma_{data}^2}{\sigma_{prior}^2} + \frac{\sigma_{data}^4}{\sigma_{prior}^4} \varphi^2)\}] \\
&\quad \cdot \frac{1}{\sqrt{2\pi(\sigma_{data}^2 + \sigma_{prior}^2)}} \exp\{-\frac{1}{2}\frac{1}{\sigma_{data}^2 + \sigma_{prior}^2}(x_1^2 - 2x_1 \varphi + \varphi^2)\} dx_1 \\
&\propto \int \frac{1}{\sqrt{2\pi}} \exp[-\frac{1}{2}\{\frac{\sigma_{data}^2 + \sigma_{prior}^2}{\sigma_{data}^2 \sigma_{prior}^2} \mu^2 - 2\mu \frac{x_1}{\sigma_{data}^2} - 2\mu \frac{\varphi}{\sigma_{prior}^2} + \frac{x_1^2}{\sigma_{data}^2} \\
&\quad + \frac{\sigma_{prior}^2}{\sigma_{data}^2(\sigma_{data}^2 + \sigma_{prior}^2)}(2x_1 \varphi \frac{\sigma_{data}^2}{\sigma_{prior}^2} + \frac{\sigma_{data}^4}{\sigma_{prior}^4} \varphi^2)\}] \\
&\quad \cdot \exp\{-\frac{1}{2}\frac{1}{\sigma_{data}^2 + \sigma_{prior}^2}(-2x_1 \varphi + \varphi^2)\} dx_1
\end{aligned}$$

$$\begin{aligned}
\pi_{t=1}(\mu) &\propto \exp\left[-\frac{1}{2}\left\{-\frac{1}{\sigma_{data}^2}\left(\mu - \frac{\sigma_{data}^2}{\sigma_{data}^2 + \sigma_{prior}^2}\varphi\right)^2 + \frac{\sigma_{data}^2 + \sigma_{prior}^2}{\sigma_{data}^2\sigma_{prior}^2}\mu^2\right.\right. \\
&\quad \left.\left.- 2\mu\frac{\varphi}{\sigma_{prior}^2} + \frac{1}{(\sigma_{data}^2 + \sigma_{prior}^2)}\frac{\sigma_{data}^2}{\sigma_{prior}^2}\varphi^2 + \frac{1}{\sigma_{data}^2 + \sigma_{prior}^2}\varphi^2\right\}\right] \\
&\propto \exp\left[-\frac{1}{2}\left\{-\frac{\mu^2}{\sigma_{data}^2} + 2\mu\frac{1}{\sigma_{data}^2 + \sigma_{prior}^2}\varphi - \frac{\sigma_{data}^2}{(\sigma_{data}^2 + \sigma_{prior}^2)^2}\varphi^2 + \frac{\sigma_{data}^2 + \sigma_{prior}^2}{\sigma_{data}^2\sigma_{prior}^2}\mu^2\right.\right. \\
&\quad \left.\left.- 2\mu\frac{\varphi}{\sigma_{prior}^2} + \frac{1}{(\sigma_{data}^2 + \sigma_{prior}^2)}\frac{\sigma_{data}^2}{\sigma_{prior}^2}\varphi^2 + \frac{1}{\sigma_{data}^2 + \sigma_{prior}^2}\varphi^2\right\}\right] \\
&\propto \exp\left[-\frac{1}{2}\left\{-\frac{\mu^2}{\sigma_{data}^2} + \frac{\sigma_{data}^2 + \sigma_{prior}^2}{\sigma_{data}^2\sigma_{prior}^2}\mu^2 + 2\mu\frac{1}{\sigma_{data}^2 + \sigma_{prior}^2}\varphi - 2\mu\frac{\varphi}{\sigma_{prior}^2}\right.\right. \\
&\quad \left.\left.- \frac{\sigma_{data}^2}{(\sigma_{data}^2 + \sigma_{prior}^2)^2}\varphi^2 + \frac{1}{(\sigma_{data}^2 + \sigma_{prior}^2)}\frac{\sigma_{data}^2}{\sigma_{prior}^2}\varphi^2 + \frac{1}{\sigma_{data}^2 + \sigma_{prior}^2}\varphi^2\right\}\right] \\
&\propto \exp\left[-\frac{1}{2}\left\{\frac{\mu^2}{\sigma_{prior}^2} + 2\mu\varphi\left(\frac{1}{\sigma_{data}^2 + \sigma_{prior}^2} - \frac{1}{\sigma_{prior}^2}\right) - \varphi^2\frac{\sigma_{data}^2}{(\sigma_{data}^2 + \sigma_{prior}^2)^2}\right.\right. \\
&\quad \left.\left.+ \varphi^2\left(\frac{1}{(\sigma_{data}^2 + \sigma_{prior}^2)}\frac{\sigma_{data}^2}{\sigma_{prior}^2} + \frac{1}{\sigma_{data}^2 + \sigma_{prior}^2}\right)\right\}\right] \\
&\propto \exp\left[-\frac{1}{2}\left\{\frac{\mu^2}{\sigma_{prior}^2} + 2\mu\varphi\left(\frac{1}{\sigma_{data}^2 + \sigma_{prior}^2} - \frac{1}{\sigma_{prior}^2}\right) - \varphi^2\frac{\sigma_{data}^2}{(\sigma_{data}^2 + \sigma_{prior}^2)^2} + \frac{\varphi^2}{\sigma_{prior}^2}\right\}\right] \\
&\propto \exp\left[-\frac{1}{2}\left\{\frac{\mu^2}{\sigma_{data}^2} - 2\frac{1}{\sigma_{data}^2 + \sigma_{prior}^2}\varphi + \frac{\sigma_{data}^2}{(\sigma_{data}^2 + \sigma_{prior}^2)^2}\varphi^2 + \frac{\sigma_{data}^2 + \sigma_{prior}^2}{\sigma_{data}^2\sigma_{prior}^2}\mu^2\right.\right. \\
&\quad \left.\left.- 2\mu\frac{\varphi}{\sigma_{prior}^2} + \frac{1}{(\sigma_{data}^2 + \sigma_{prior}^2)}\frac{\sigma_{data}^2}{\sigma_{prior}^2}\varphi^2 + \frac{1}{\sigma_{data}^2 + \sigma_{prior}^2}\varphi^2\right\}\right] \\
&\propto \exp\left[-\frac{1}{2}\left\{\frac{\mu^2}{\sigma_{prior}^2} - 2\mu\frac{\varphi}{\sigma_{prior}^2} + \varphi^2\right\}\right] \\
&\propto \exp\left[-\frac{1}{2\sigma_{prior}^2}(\mu - \varphi)^2\right].
\end{aligned}$$

Thus, $\pi_{t=1}(\mu) = \pi_{t=0}(\mu)$ so that Bayesian plausibility is satisfied.

B.6 Estimation

We now describe the estimation procedure for $Sell_{notitle}$ and κ . Both the full information model and the Bayesian model use this same method of estimating these two parameters in a first stage.

We estimate κ directly from the probabilities of each choice in the data using method of moments. Consider the moments,

$$\tilde{g}_{home} = \frac{\sum_{t=5}^{10} \sum_i 1(i \text{ sells at } t | \text{homestead at } t = 0)}{\sum_{t=5}^{10} \sum_i 1(i \text{ abandons at } t | \text{homestead at } t = 0)} - \exp(1280 \cdot \hat{\kappa})$$

$$\tilde{g}_{purchase} = \frac{\sum_{t=5}^{10} \sum_i 1(i \text{ sells at } t | \text{purchase at } t = 0)}{\sum_{t=5}^{10} \sum_i 1(i \text{ abandons at } t | \text{purchase at } t = 0)} - \exp(1475 \cdot \hat{\kappa}).$$

The average resale price if the farm was homesteaded at $t = 0$ is \$1,480 (less \$200 from acquiring the title before $t = 10$) and \$1,475 is the average resale price if the farm was purchased at $t = 0$. We use the above two moments to estimate κ and use that to calculate $\widehat{Sell}_{notitle}$ as

$$\widehat{Sell}_{notitle} = M \cdot \hat{\kappa}.$$

Note that $\widehat{Sell}_{notitle}$ measures utility. The parameter κ converts utility into dollars. Let $\widehat{Sell}_{title} = (M + 200) \cdot \hat{\kappa}$. We set $M = 1,280$ so that $M + 200$ is the average resale price for homesteaders. An additional implication of our model is that dynamic selection and a discrete mixture implies that, after some time, only the most productive farmers continue to farm. That is, the researcher could assume that after 10 years they observe only one type or the types are equally able (with potentially different priors). This is not essential for identification of our model but this feature can be used to estimate the features of the high productivity type, type A in our case, using data after 10 years of farming.

Using the estimated $\hat{\kappa}$ from the moment conditions above, we calculate $V_{A,final}$ and μ_A from the data. Let

$$\omega = \frac{1}{2} \left(\frac{\sum_{t=20}^{24} \Pr(\text{continue at } t | \text{purchase at } t = 0)}{\sum_{t=20}^{24} \Pr(\text{sell at } t | \text{purchase at } t = 0)} \right) + \frac{1}{2} \left(\frac{\sum_{t=20}^{24} \Pr(\text{continue at } t | \text{homestead at } t = 0)}{\sum_{t=20}^{24} \Pr(\text{sell at } t | \text{homestead at } t = 0)} \right).$$

Using ω , we identify $V_{A,final}$ as

$$V_{A,final} = \ln[\omega \cdot \exp\{(M + 200) \cdot \kappa\} + \exp\{(M + 200) \cdot \kappa\} + 1],$$

and likewise identify μ_A using

$$\mu_A = \ln(\omega) + (M + 200) \cdot \kappa - \delta V_{A,final}.$$

$V_{A,final}$ and μ_A are functions of quantities that we can estimate, and let $\hat{V}_{A,final}$ and $\hat{\mu}_A$

denote the corresponding estimators. We then use $\hat{V}_{A,final}$, $\hat{\mu}_A$, $\hat{\kappa}$, and $\widehat{Sell}_{notitle}$ as inputs to estimate the Bayesian model.

In estimation, we use maximum likelihood and moments for the fixed points. The log-likelihoods for homesteaders and for purchasers are calculated from the survival functions. We use two additional moments, one for the final fixed points for each type.

We use the conditional choice probabilities and survival functions to match the probabilities in the data for purchasers and homesteaders to the probabilities for purchasers and homesteaders in the model. Consider the matrix $R_{m,home}$, which gives the probability implied by the Bayesian model of each choice (continue, sell, or abandon) in period t for type m , given that the farm has survived (chosen to continue) up to t .

$$R_{m,home} = \begin{pmatrix} H_{m,home}(1) & H_{m,home}(2) & \dots & H_{m,home}(24) \\ \Pr(d_{m,1}=S|d_{m,0}=home) & H_{m,home}(1) \Pr(d_{m,2}=S|d_{m,0}=home) & \dots & H_{m,home}(23) \Pr(d_{m,24}=S|d_{m,0}=home) \\ \Pr(d_{m,1}=A|d_{m,0}=home) & H_{m,home}(1) \Pr(d_{m,2}=A|d_{m,0}=home) & \dots & H_{m,home}(23) \Pr(d_{m,24}=A|d_{m,0}=home) \end{pmatrix}$$

Call $R_{data,home}$ the analogous matrix for the data. Note that the rows are *continue*, *sell*, and *abandon* respectively. Therefore, the first row of $R_{m,home}$ gives the survival probability for each period (i.e., the probability of continuing to farm in that period multiplied by the probabilities of continuing to farm in every previous period). Rows two and three are the probability of selling and abandoning respectively multiplied by the probabilities of continuing to farm in every previous period.

Recall that $H_{m,home,t}$ denotes the probability of continuing through at least time t for homesteaders of type m , and $H_{A,home,t=0}$ denotes the proportion of homesteaders who are type A . Then let

$$\tilde{R}_{home} = \ln\{H_{A,home,t=0} \cdot R_{A,home} + (1 - H_{A,home,t=0}) \cdot R_{B,home}\} \cdot R_{data,home}.$$

The log-likelihood for homesteaders is

$$LL_{home} = \sum_{t=1}^{20} \left(\tilde{R}_{home}(2, t) + \tilde{R}_{home}(3, t) \right) + \tilde{R}_{home}(1, 20).$$

We use only the last period for the conditional probability of continuing ($\tilde{R}_{home}(1, 20)$) in the log-likelihood because the information in $\tilde{R}_{home}(1, t)$ for $t < 20$ is contained in the conditional choice probabilities for selling and abandoning ($\tilde{R}_{home}(2, t)$ and $\tilde{R}_{home}(3, t)$). Note that in order to sell or abandon in period t , a farmer must have chosen to continue in every period

prior to t . We create $LL_{purchase}$, the analogous log-likelihood for purchasers.

The two fixed point moments (one for each type) are the same for homesteaders and purchasers and are given by

$$g_{m,\text{fixed point}} = 4 \cdot [\ln\{\exp(\mu_m + \delta V_{m,\text{final}}) + \exp(\widehat{Sell}_{notitle} + \hat{\kappa} \cdot 200) + 1\} + 0.577 - V_{m,\text{final}}].$$

Using the log-likelihood and the fixed point moments, the objective function Q can be written

$$Q = -g_{A,\text{fixed point}}^2 - g_{B,\text{fixed point}}^2 + LL_{home} + LL_{purchase}.$$

We bootstrap the data using 100,000 replications to provide standard errors (discussed in the next section). Lastly, we restrict the value functions as follows:

1. $V_{m,\text{home},t} \leq V_{m,\text{home},t+1}$ for $t \in \{1, \dots, 9\}$.
2. $|V_{m,\text{purchase},t} - V_{m,\text{home},t}| \leq 0.1$ for $t \in \{11, \dots, 20\}$.
3. $V_{m,\text{final}} - V_{m,\text{purchase},20} \leq 1$.
4. $V_{m,\text{purchase},20} \leq V_{m,\text{final}}$.
5. $V_{B,\text{home},t} \leq V_{A,\text{home},t}$ for $t \in \{1, \dots, 20\}$.
6. $V_{B,\text{purchase},t} \leq V_{A,\text{purchase},t}$ for $t \in \{1, \dots, 20\}$.

These restrictions follow from the institutional details of our setting. Restriction (1) means that the homesteads are increasing in value until they get the title. Restriction (2) means that once the homesteaders have received the title, the value functions for the homesteaders and the purchasers cannot become arbitrarily far apart, since both have the title. Restriction (3) means that the final fixed points for type A and type B should be moderately close to the second-to-last fixed points (in period 20) for each type. Restriction (4) means that the final fixed points for each type should be larger than the second-to-last fixed points. Restrictions (5) and (6) mean that the value for the high ability type should be larger in each period than the value for the low type. Restriction (4) is not binding in our estimation.

Additionally, we normalize the X_t in the estimation by constraining it to have mean zero and setting the second moment equal to one, i.e., $\sum_{t=1}^{20} X_t = 0$ and $\frac{\sum_{t=1}^{20} X_t^2}{20} = 1$. One

could use the X_t as the unobserved state, but we use the prior means and variances as state variables.

B.7 Bootstrap

We estimate both the Bayesian model and the full information model in Julia, which we also use to create bootstrapped standard errors.

We observe the histories of 7,011 farmers. We then sample with replacement to create a new bootstrap sample of 7,011. From that sample, we create the probabilities of abandoning, selling, and continuing in each period. Because of this, our bootstrap resembles the original data generating process. We bootstrap the data using 100,000 replications to provide standard errors.

In estimating our counterfactual fraction of homesteaders who would have still acquired land in a scenario without homesteading, we use a large sample to estimate the parameters γ and ρ . In particular, the 1870 U.S. census gives the U.S. population as 38,558,371. In such a case the nonparametric bootstrap is not practical. A feasible approach is to use the normal approximation of the estimators and use the parametric bootstrap. That is,

$$BootstrapFraction_{home} = \Pr(\text{home}|\text{farmer}) + Z_1 \cdot (\text{s.e.}(\Pr(\text{home}|\text{farmer}))),$$

where Z_1 is a realization from the standard normal, $\Pr(\text{home}|\text{farmer}) = \Pr(\text{home}|\text{home or purchase})$, and $\text{s.e.}(\Pr(\text{home}|\text{farmer}))$ is given by

$$\text{s.e.}(\Pr(\text{home}|\text{farmer})) = \sqrt{(\Pr(\text{home}|\text{farmer})) \cdot \frac{1 - \Pr(\text{home}|\text{farmer})}{Population}}.$$

We calculate the analogous fraction for buying, $BootstrapFraction_{purchase}$, as given by

$$BootstrapFraction_{purchase} = \Pr(\text{purchase}|\text{farmer}) + Z_2 \cdot (\text{s.e.}(\Pr(\text{home}|\text{farmer}))),$$

where Z_2 is a realization from the standard normal and Z_1 and Z_2 are independent of each other.

We generate 100,000 estimates of the probabilities to homestead and 100,000 estimates of the probabilities to buy. We then calculate the counterfactual 100,000 times. This method

is valid as the counterfactual is approximately a linear function of the parameters.

B.8 Counterfactual: No Homestead Act

To calculate the fraction of individuals who would have bought land in a counterfactual situation in which homesteading was not available, we use the following nested logit model, where $(1 - \rho)$ is the dependence between purchasing and homesteading and γ is the disutility to move to and farm in the counties under consideration for people outside those counties. The nested logit is McFadden's extension of his conditional logit model, and we use this nested logit model for both our full information and Bayesian models. Both the conditional logit and the nested logit models let the choices depend on utilities rather than money, meaning we can only estimate the nested logit model because we used the structural model to estimate κ . We first consider the full information model.

$$\Pr(\text{buy}|\text{homestead or buy}) = \frac{1}{1 + \exp\{\kappa \cdot \Psi/\rho\}} \text{ and } 0 < \rho \leq 1, \kappa > 0.$$

$$\Pr(\text{no farming}) = \frac{1}{1 + [\exp(\gamma/\rho) + \exp\{\gamma/\rho - \kappa \cdot \Psi/\rho\}]^\rho}, \text{ and } 0 < \rho \leq 1, \kappa > 0.$$

We use

$$\Pr(\text{buy}|\widehat{\text{homestead or buy}}) = \frac{\#sales \text{ at time } = 0}{\#sales \text{ at time } = 0 + \#\widehat{\text{homestead at time}} = 0}.$$

$\Pr(\text{buy}|\text{homestead or buy})$ can be rewritten as

$$\rho = \frac{\kappa \cdot \Psi}{\ln\left\{\frac{1}{P(\text{buy}|\text{homestead or buy})} - 1\right\}}.$$

This suggests the estimator

$$\hat{\rho} = \frac{\hat{\kappa} \cdot \hat{\Psi}}{\ln\left\{\frac{1}{\widehat{P(\text{buy}|\text{homestead or buy})}} - 1\right\}},$$

where $\hat{\kappa}$ comes from the estimation of the structural model. Likewise, $\Pr(\text{no farming})$ can be rewritten as

$$\gamma = \ln\left[\frac{1}{\Pr(\text{no farming})} - 1\right] - \rho \ln[1 + \exp\{-(\kappa \cdot \Psi)/\rho\}],$$

where γ represents the disutility from farming. This suggests the estimator

$$\hat{\gamma} = \ln\left[\frac{1}{\widehat{\Pr(\text{no farming})}} - 1\right] - \hat{\rho} \ln[1 + \exp\{-(\hat{\kappa} \cdot \hat{\Psi})/\hat{\rho}\}].$$

After calculating $\hat{\rho}$ and $\hat{\gamma}$ as above, we use the nested logit model to obtain the following probability

$$\begin{aligned} \Pr(\text{no farming}|\widehat{\text{no homesteading}}) &= \frac{1}{1 + [\exp\{\hat{\gamma}/\hat{\rho} - (\hat{\kappa} \cdot \hat{\Psi})/\hat{\rho}\}]^{\hat{\rho}}} \\ &= \frac{1}{1 + [\exp\{\hat{\gamma} - (\hat{\kappa} \cdot \hat{\Psi})\}]}. \end{aligned}$$

This counterfactual is the conditional probability of not farming given that homesteading is no longer available.

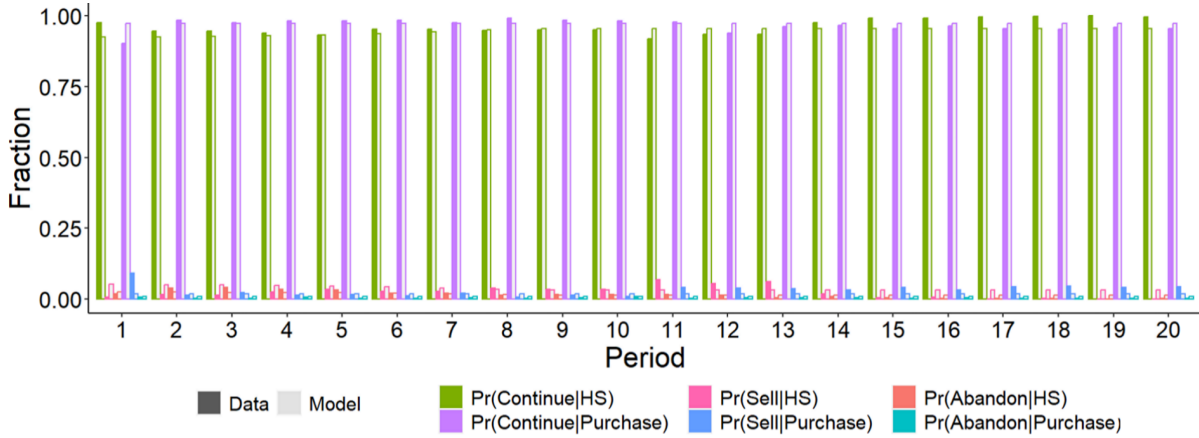


Figure 8: Fit of the Full Information model

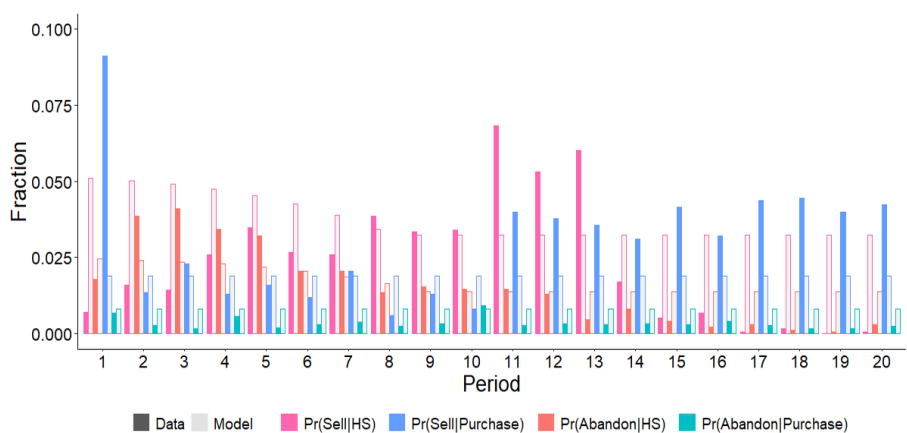


Figure 9: Fit of the Full Information model: abandoning and selling only

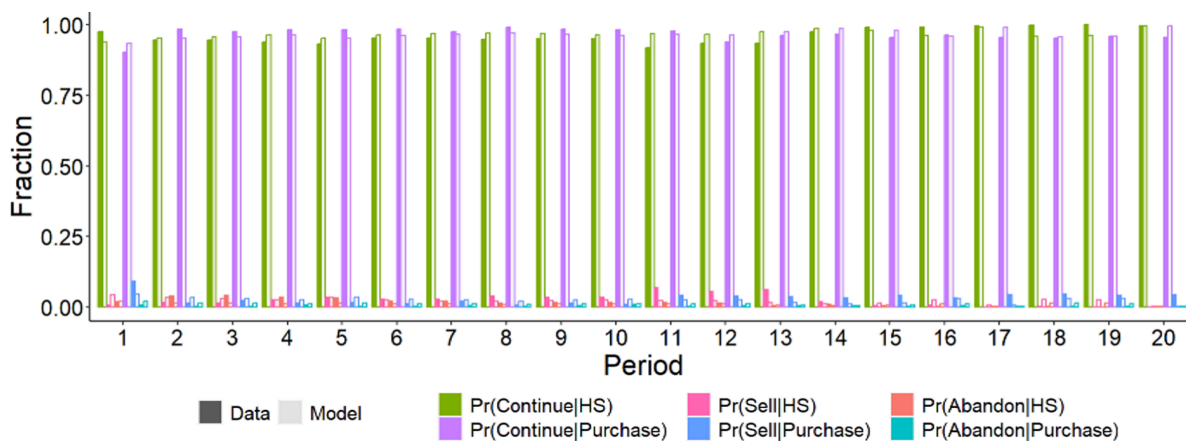


Figure 10: Fit of the Bayesian model

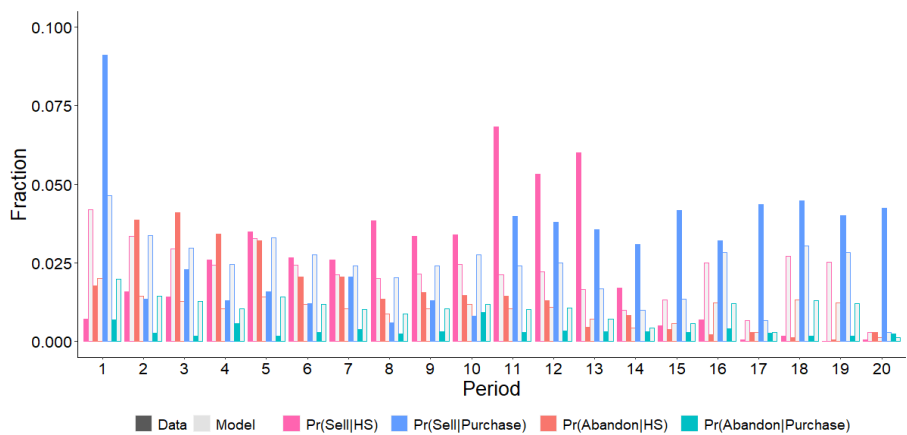


Figure 11: Fit of the Bayesian model: abandoning and selling only

Table 7: Information Realizations and Bias Estimates

	Information Realizations	Bias for Type <i>A</i>	Bias for Type <i>B</i>
Period 1	1.30682 (0.188809)	-0.99094 (0.110855)	-1.04133 (0.082695)
Period 2	1.00572 (0.243388)	-0.70947 (0.106206)	-0.75276 (0.082046)
Period 3	-0.17807 (0.119778)	-0.64382 (0.095560)	-0.68176 (0.074765)
Period 4	-0.65352 (0.086183)	-0.64488 (0.094071)	-0.67865 (0.075393)
Period 5	-0.66063 (0.090771)	-0.64644 (0.092496)	-0.67687 (0.075627)
Period 6	-0.56302 (0.122124)	-0.63892 (0.089950)	-0.66660 (0.074679)
Period 7	-0.18915 (0.211434)	-0.60173 (0.095108)	-0.62712 (0.081052)
Period 8	-0.59586 (0.113099)	-0.60128 (0.095102)	-0.62474 (0.082068)
Period 9	-0.59209 (0.159666)	-0.60063 (0.097693)	-0.62242 (0.085516)
Period 10	-0.62473 (0.122455)	-0.60223 (0.099286)	-0.62257 (0.087868)
Period 11	-1.64573 (0.207200)	-0.66707 (0.084776)	-0.68615 (0.074038)
Period 12	-1.77472 (0.299524)	-0.73187 (0.064853)	-0.74983 (0.054756)
Period 13	-0.57998 (0.174565)	-0.72347 (0.063778)	-0.74044 (0.054238)
Period 14	0.02653 (0.171250)	-0.68419 (0.063864)	-0.70027 (0.054897)
Period 15	0.45104 (0.189104)	-0.62770 (0.064681)	-0.64298 (0.056323)
Period 16	0.30818 (0.184590)	-0.58333 (0.054957)	-0.59789 (0.047244)
Period 17	0.42894 (0.280091)	-0.53751 (0.040828)	-0.55141 (0.033927)
Period 18	0.70137 (0.376991)	-0.48387 (0.023735)	-0.49716 (0.018457)
Period 19	1.44696 (0.282214)	-0.40373 (0.012128)	-0.41647 (0.010286)
Period 20	2.38195 (0.178552)	-0.29272 (0.007096)	-0.30495 (0.006369)
Observations	7,011	7,011	7,011

Standard errors in parentheses

Percent of Total Farms Abandoned	
January	16.7%
February	4.5
March	6.9
April	9.5
May	7.7
June	5.9
July	11.5
August	4.5
September	5.7
October	5.4
November	13.4
December	8.3
Total	100.0
$N = 1,111$	

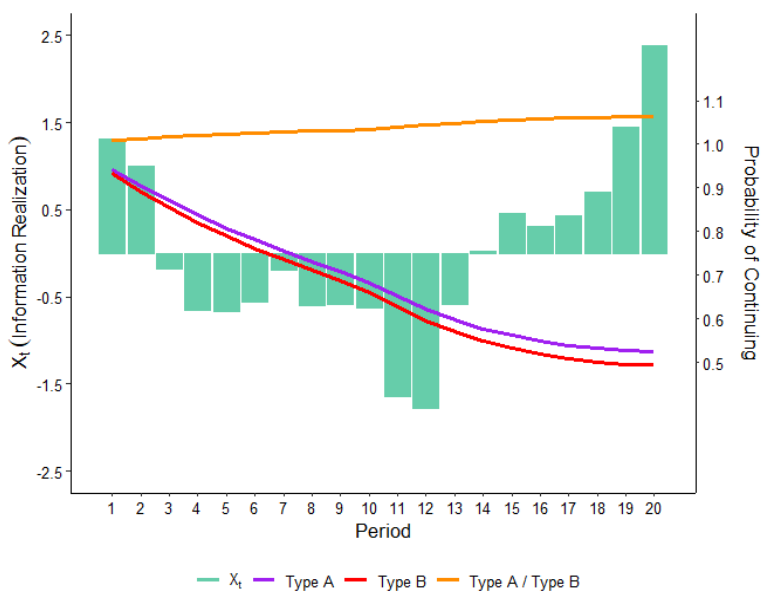


Figure 12: The probability of continuing to farm varies by type

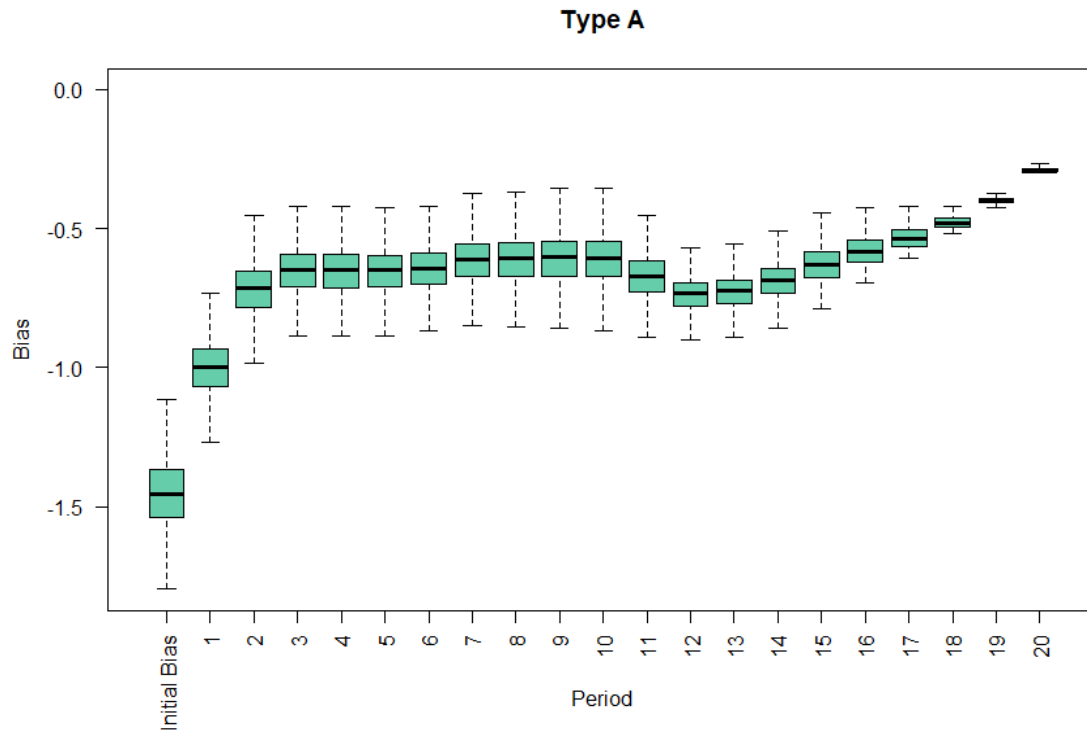


Figure 13: The bias of type *A* decreases toward zero over the 20 periods; the bias of type *B* is slightly larger

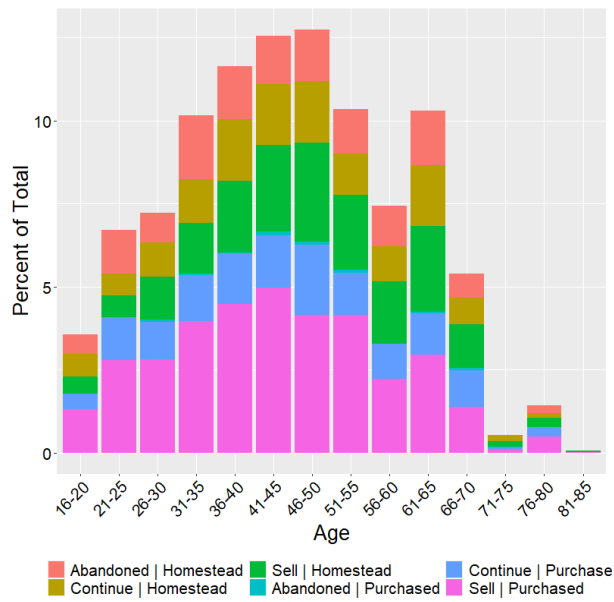


Figure 14: The distribution of land acquisition choice is consistent over all age brackets

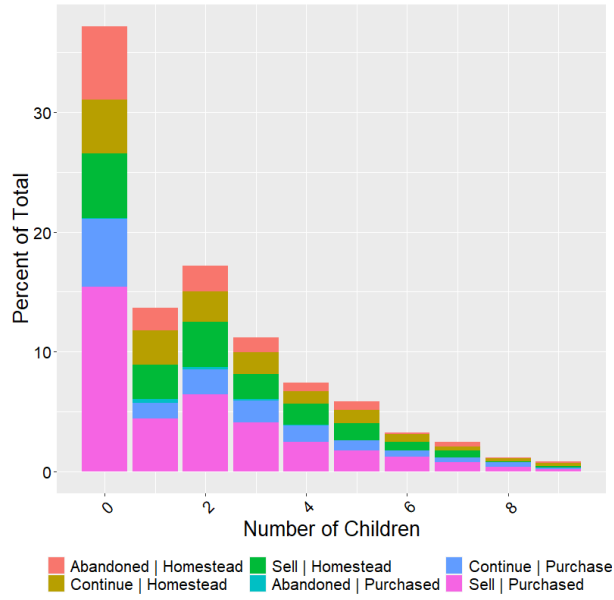


Figure 15: The distribution of land acquisition choice is consistent across farmer's number of children

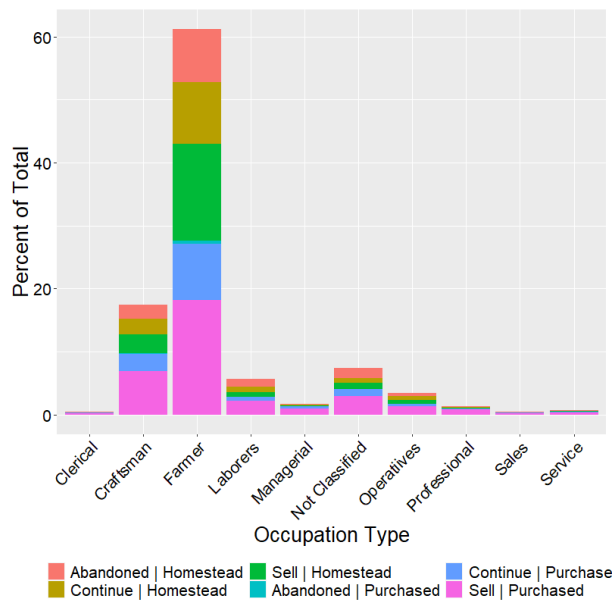


Figure 16: The distribution of land acquisition choice is consistent over all occupations of the owner

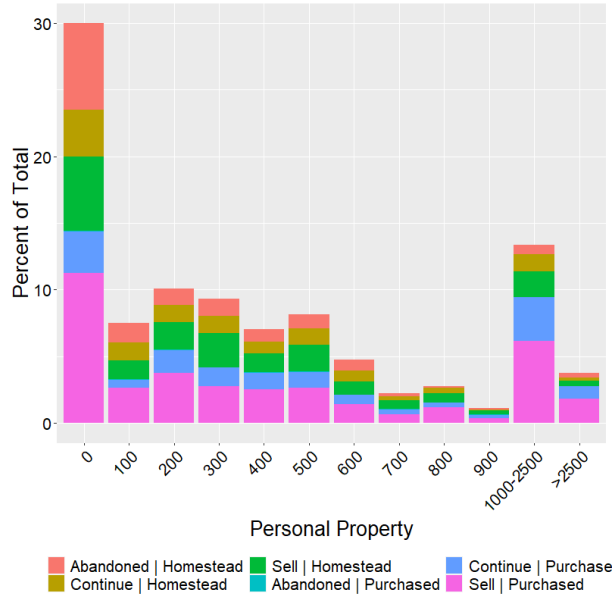


Figure 17: The distribution of land acquisition choice is consistent over all wealth brackets of the owner

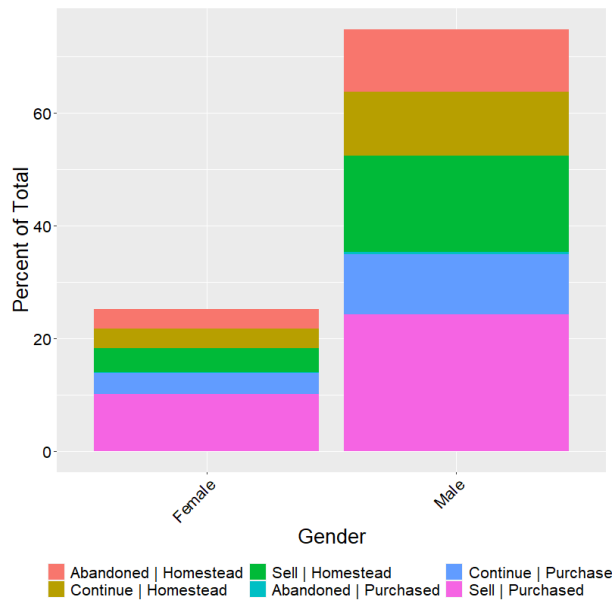


Figure 18: The distribution of land acquisition choice is consistent over genders

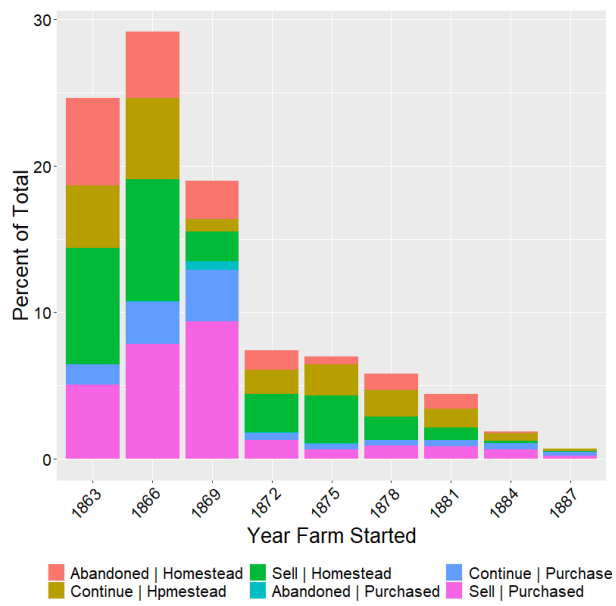


Figure 19: The distribution of land acquisition choice is consistent over all years which the farm was initially acquired