Estimating Time Preferences with Structure: Crop Rotations in Agriculture

Allen Peters*
University of British Columbia

Working Paper
CLICK HERE FOR THE MOST RECENT VERSION

October 28, 2023

Abstract

Time preferences are omnipresent, but they are difficult to measure in the contexts in which they are applied. In agriculture, farmers' time preferences drive choices that impact food security, industry sustainability, and the environment. I structurally estimate the discount rate of farm operators in Alberta, Canada using a dynamic discrete choice model of crop rotation decisions. My estimation strategy leverages the finite temporal dependence of expected yields on crop history and builds on a recent identification result for dynamic discrete choice models. My estimates suggest a strong present bias, somewhat in line with experimental estimates and in contrast to common modelling assumptions.

*Email: apeters@protonmail.com. I would like to thank Derek Nalder from Alberta Agriculture & Forestry for providing me with an irrigation districts map file. Thank you as well to Frederik Noack and Sumeet Gulati for allowing me to present at the UBC Faculty of Land and Food Systems and to James Vercammen and Henry Siu for valuable feedback. I am further grateful to Paul Schrimpf for his supervision, direction, and help with error calculations and to Elyse Adamic for many discussions and endless support. Any errors or omissions are entirely my own.

1 Introduction

How individuals and groups of individuals value the future versus the present is a fundamental question in a wide range of contexts. It is particularly important where long term outcomes or sustainability are important, however any economic model involving an intertemporal decision depends on the discount factor. Unfortunately, discounting is difficult to measure and is therefore simply assumed in many models, often leading to controversial outcomes. Perhaps the most notable example of this is in climate-economy models, where a change in the discount factor of a few percentage points can alter the estimated social cost of greenhouse gas emissions by a factor of two or more. This paper looks at discounting in the context of agricultural practices, where it is key to food security, long term profitability, local environmental protection, and global climate stewardship. I present a novel approach to measuring discounting among farm operators, which leverages crop rotation decisions in a dynamic structural model to estimate context-specific time preferences.

A large literature exists that takes as given that *individuals* may have non-zero pure rates of time preference, and attempts to measure them. These estimates vary widely, but often imply significant present bias (Frederick et al., 2002). In contrast, economic models typically assume that *firms* have a zero rate of pure time preference. For example, in dynamic discrete choice models that are widely used to study firm decisions, the discount factor is often fixed at 5% per year, equal to the long-run interest rate (e.g., Collard-Wexler, 2013; Dunne et al., 2013). However, there are reasons to question this assumption. Especially in large firms, intertemporal choices represent a complex aggregation of individuals' time preferences, and heterogeneity among individuals further complicates aggregation (Ebert et al., 2018; Frederick et al., 2002; Gollier & Zeckhauser, 2005). For small firms, aggregation may be less of a factor, however the distinction between an individual and a firm becomes unclear.

In this paper, I study a particular class of small firms, namely farms, and elicit their discount factor from a key intertemporal decision. I construct a dynamic discrete choice (henceforth DDC) model of operational decisions and build on a recent identification result from Abbring and Daljord (2020) to structurally estimate the discount factor. While this approach requires structural assumptions regarding the operational environment of the firm, it elicits context-specific values and offers a transparent identification channel.

My identification strategy leverages decisions regarding crop rotation, which is the practice of varying the type of crop that is grown on a plot of land from one growing season to the next. Leaving intervals between growing a particular crop is considered good practice for long-term soil management and prevents yields from declining over time due to crop-specific diseases, pests, weeds, and moisture depletion. As such, one season's crop choice impacts not only that season's yield and profit, but also future profit through its impact on crop yield in subsequent seasons. The discount factor is identified from crop rotation decisions via an exclusion restriction on average expected profit in different states. The intuition behind this is that choices that shift expected profit between the present season and the future provide direct information on the farmer's relative valuation of present and future. I take the average expected profit difference between planting any two crops as known and the variance of the profit difference as a nuisance parameter to be estimated.

I estimate the discount factor separately for three distinct regions using a rich dataset of plot-level crop choices derived from satellite imagery. In contrast, data on expected yields and profits is aggregated at the regional-level. To overcome this limitation, I compute more granular expected yields using yield penalties from non-ideal crop rotations that are drawn from region-specific crop rotation experiments in the scientific agriculture literature.

I find evidence of strong present bias, with estimates ranging from 0.33 to 0.8. This suggests that farm operators may behave more like individuals than canonical firms—contrary to the typical modeling assumption. This has important policy implications, as the time preferences of farm operators drive land-use choices that affect key issues such as food security and environmental sustainability. Additionally, I demonstrate a novel application of Abbring and Daljord's identification result for dynamic discrete choice estimation frameworks and employ it with partially aggregate data. To my knowledge, this is only the second application of dynamic discrete choice models to estimate discounting outside a *consumption* choice setting. In doing so, I explore the potential of this technique for establishing revealed context-specific discount rates in a wide range of settings and without the requirement for extremely granular data.

2 Background

2.1 Related Literature

This paper relates to three strands of literature in terms of both methodology and topic. First, this work relates to a vast literature on estimating time preferences as well as more specific strands on both firm and farmer time preferences. The methodology I use builds on a recent identification strategy by Abbring and Daljord (2020) and other approaches that use dynamic discrete choice models to study consumers and firms. Finally, there is a small literature applying discrete choices model to study land-use, including a model by Scott (2014) that I build from.

Both bottom-up and top-down approaches have been applied to estimate time preferences. An extensive microeconomics literature exists in which time preferences of individuals are elicited from lab or field experiments (e.g., Dean and Sautmann, 2014). These estimates may suffer from questionable external validity due to the artificial lab setting or the typically small stakes involved and can only be applied to individuals, not groups. At the other end of the spectrum, aggregate discount factors for entire populations are often estimated from macroeconomic models using the Euler equation (e.g. Cagetti, 2003; Gourinchas and Parker, 2002). This approach abstracts from the complex set of interactions through which individual or investor preferences manifest themselves in firm behaviour. Perhaps not surprisingly, these different approaches elicit strikingly different estimates of discount rates, as has been thoroughly documented by Frederick et al. (2002).

The literature that directly focuses on firm discounting is sparse. A recent strand of management literature studies short-termism but does not directly estimate discounting (e.g., Anderson et al., 2012). One exception is Harris and Siebert (2017), who estimate discount factors for individual semiconductor firms using a structural model of mergers. Their estimates range as low as 0.73, with the bulk of firms in the range 0.93-0.96.

A handful of authors has looked at the discount factors of farmers. Duquette et al. (2012), Hermann et al. (2015), and Harrison et al. (2002) use experimental evidence that elicits the personal discount factors of farmers. All obtain values in a range from 0.7 to 0.78. Hermann et al. find a weak correlation between discount rate and farm size. Structural estimates include Lence (2000)[1] and Abdulkadri and Langemeier (2000), which both use consumption and investment data for farmers to estimate discount factors directly from the Euler equation. Their estimates vary from below 0.95 to above 0.99.

Recently, a number of authors have estimated *consumer* discount factors using dynamic discrete choice models (e.g. De Groote and Verboven, 2019; Einav et al., 2015; Rossi, 2018). To the best of my knowledge, Bollinger (2015) is the only study that estimates the discount factor of *firms* in a discrete choice setting. This is in the context of green technology adoption in the garment cleaning industry and uses exogenously changing environmental policies that affect future but not current profits to identify the discount factor. He finds a discount factor of 0.94.

A number of authors have applied discrete choice models to land-use choices, however this is typically to investigate land conversion in and out of agriculture or between cash crops and pasture. Claassen and Tegene (1999) is an early example of a static discrete choice setup for land-use choices, while Lubowski et al. (2008) is an example of a dynamic treatment. Of particular note is Scott (2014), as I draw from his model setup in the current work. He develops a dynamic discrete choice model to investigate farm operator land-use decisions in the United States. His focus is on an alternative estimation technique, and he does not estimate the discount factor. Wu et al. (2004) applies a discrete choice model to crop and tillage decisions, but in a static model. The present work is the first to elicit time preferences from crop rotation decision using structural modeling and estimation.

2.2 Agriculture and Crop Rotations in Canada

Agriculture is a key industry in many parts of Canada, particularly in the west-central Prairie region, comprised of the provinces of Alberta, Saskatchewan, and Manitoba. This region constitutes one of the world's major producers and exporters of grains and oilseeds. The industry exhibits a number of features that make it amenable to modeling with the single-agent dynamic discrete choice framework.

The agricultural industry in Canada features competitive markets with a large number of small firms. As of 2012, when the Canadian Wheat Board's Single Desk marketing power ended, all major crops are sold in an open, often global market. The 2016 census reported 13,451 oilseed and grain farms (Statistics Canada, 2020c). The average land area of a farm is 1,237 acres (Statistics Canada, 2020b), however a single field is typically around 160 acres, so an average a farm consists of less than ten fields. Profit margins are thin, with the median income of *families* with oilseed and grain farms around $80,865 in 2010 (Statistics Canada, 2016).

Farms in Canada are primarily sole proprietorships (52%), with only 2% classified as non-family

[1] The estimates in Lence (2000) likely suffer from small sample bias.

corporations and the remainder being partnerships or family corporations.[2] This adds another dimension to the question of time preferences in this particular context, as it is unclear whether to expect their behaviour to more closely resemble a canonical firm, discounting at the interest rate, or an individual, who might tend to be more present biased. As in many developed nations, there is a trend toward corporatization (Magnan, 2015), which is controversial in part due to how it may affect environmental stewardship, a topic closely related to time preferences.

Farm operators face a discrete and relatively small set of viable crops from which to choose. The most commonly grown crops in Alberta are wheat, canola, barley, and field peas. The per-farm area of land seeded for just wheat and canola was 602 acres and 559 acres in 2020 (Statistics Canada, 2020b). Crop decisions are made roughly simultaneously due to climatic constraints. Canada's climate permits a single growing season, so apart from winter wheat, which makes up less than 3% of all wheat seeded in Alberta,[3] crop decisions are synchronized in late winter.

Importantly for this paper, crop decisions have a strong intertemporal aspect. Diverse rotations lead to increased long-term profitability due to the reduction of pests, diseases, and weeds and the retention of residual nutrients and moisture (Zentner et al., 2002). For example, wheat has been shown to benefit from increased soil available nitrogen from peas and reduced weeds after canola (Gill, 2018), while canola is particularly susceptible to a disease known as Blackleg in continuous cropping (Harker et al., 2018). With the adoption of reduced tillage practices, crop-specific diseases and weeds have become an even stronger motivation for crop rotation (Kutcher et al., 2011).

Crop rotation is currently a particularly salient issue for canola given its recent rise in popularity. The amount of canola seeded each year has grown enormously since the early 2000s and has come to rival the traditional mainstay wheat crop in many regions. Its high profitability puts pressure on farmers to shorten crop rotations despite the known yield impacts. Harker et al. (2018) reports that canola was grown every second year in 40% of the Canadian Prairies in 2015, while 5% was continuous-cropped.

3 Model

I model farm operator crop choices using a stationary single-agent dynamic discrete choice model (e.g. Rust (1994)). In accordance with the fundamental independence assumption required of this class of models, I assume that a single farmer owns a single field, and therefore the terms field, farm, farmer, farm operator, and agent are used interchangeably. Fields within a region r are assumed to be of homogeneous type; they may differ due to their crop history and idiosyncratic shocks but are similar in every other sense, such as size, soil quality and expected weather. Since I analyze regions independently, I omit the r subscript throughout the descriptions of model setup and identification.

As per the typical growing season in the Canadian Prairies, farmers make crop decisions once per year. Time is discrete with an infinite horizon. Farm operators act to maximize expected discounted

[2] Figures calculated based on data from Statistics Canada (2020d)

[3] Calculated based on data for 2011-2019 from Statistics Canada (2020a)

payoffs, which are determined entirely by and are therefore identical to expected profits. Formally, I specify the single-period payoff for field i at the time of planting as:

$$U(j, x_{it}, \nu_{it}) = \pi(j, x_{it}) + \nu_{jit}, \tag{1}$$

where the crop choice is $j \in \mathcal{J} = \{0, 1, ..., J-1\}$ where J is the number of crop choices considered. x_{it} denotes the vector of observable state variables, to be further defined below. The error term ν_{jit} represents idiosyncratic variations that are assumed to be distributed identically and independently across fields and time. ν_{jit} is observed by the agent but not the econometrician. It allows for non-persistent heterogeneity in expected profits and is unrelated to the agent's uncertainty in payoffs. $\pi(j, x_{it})$ is the average expected profit from planting crop j in state x_{it}.

This payoff specification embodies a number of assumptions that are typical of dynamic discrete choice models. First, it implies farm operators are risk neutral. While some studies emphasize the importance of uncertainty and risk aversion for farmers (e.g. Hermann et al. (2015) and Lence (2000)), I choose to model them as one would typically model firms and to look for a departure from that model in terms of the discount rate. Risk neutrality is a typical assumption in the discrete choice literature and is likely a reasonable approximation given the typical magnitude of the difference in expected profit between crop choices.[4] Also implicit in the specification is that payoffs are additively separable in the observable component $\pi(j, x_{it})$ and unobservable component ν_{jit}.

The standard assumption of independent error terms implies that unobserved differences between fields are independent. This precludes economies of scale, although these likely play a role in farm operators' actual decisions, as evidenced by the current trend of increasing farm size. Identically distributed error terms across individuals and time implies there is no persistence in the unobservable component of payoffs. I assume a Gumbel distribution with mean zero and scale parameter α. In imposing a common scale parameter, I further assume the error terms are identically distributed across choices.[5]

I assume the discount rate β is constant over time and common across farm operators. Revenue and costs are not discounted *within* periods, for example between planting and harvest. Given stationarity, the Bellman equation for this dynamic choice problem can be written as follows. I omit the i subscripts for the remainder of this section for brevity, denoting period $t+1$ with a prime.

$$V(x, \nu_j) = \max_{j \in \mathcal{J}} \left\{ U(j, x, \nu_j) + \beta \int V(x', \nu'_{j'}) dF(x', \nu'_{j'} \mid j, x, \nu_j) \right\}, \tag{2}$$

with $V(x, \nu_j)$ being the optimal decision rule and $F(\cdot)$ the Markov transition distribution function describing the farm operator's belief about future states. Due to the additively separable payoff specification, we can write the choice specific value function as $v(j, x, \nu_j) = v(j, x) + \nu_j$. The choice-specific value function prior to the realization of the idiosyncratic shock, which I refer to as the *expected* choice-specific value function, is therefore:

[4] It is unclear whether the risk aversion parameter would be identified in a specification with absolute risk aversion. Allowing for relative risk aversion would require complete farm income data.

[5] It may be possible to relax this assumption in subsequent work. This is explained further in Section 4.

$$v(j,x) = \pi(j,x) + \beta \int V(x', \nu'_{j'}) dF(x', \nu'_{j'} \mid j, x, \nu_j). \tag{3}$$

Taking the future state to be independent of the current error term, conditional on the current state and choice, and assuming the set of states $\mathcal{X}$ to be discrete and finite with number of states X, we can rewrite the expected choice-specific value function as:

$$v(j,x) = \pi(j,x) + \beta \sum_{x'} \mathbb{E}_{\nu'} \left[\max_{j' \in \mathcal{J}} \left\{ v(j', x') + \nu'_{j'} \right\} \right] q(x' \mid j, x). \tag{4}$$

3.1 State Space

I discuss two different state space definitions. The first is a larger, more intuitive space, which explicitly incorporates the effects declining yield with shorter crop break intervals. The second compresses the state space into a version more suitable for estimation with a short time series of aggregate data.

The first state space consists of the individual field state k_{it} and the regional market state ω_t:

$$\tilde{x}_{it} = (k_{it}, \omega_t) \tag{5}$$

The field state is defined as the crop history of a given field. Assuming that the field state displays a finite dependence of two years (i.e. crop history more than 2 years prior is irrelevant to current expected outcomes), we can write it as

$$k_{it} = (k^1_{it}, k^2_{it}), \tag{6}$$

where $k^1_{it} = j_{i,t-1}$ and $k^2_{it} = j_{i,t-2}$ are the crops planted on field i in the previous year and two years prior, respectively. As such, the evolution of the field state is deterministic. Most crop rotation studies do not examine the effects of crop rotation beyond three year rotations (corresponding to two year dependence), as the effects are generally considered to be most pronounced over the first two years.[6] The field state subsumes all disease, pest, weed, and soil moisture effects of crop rotations.

I define the market state as the set of expected benchmark profits for each crop choice (indicated by superscript):[7]

$$\omega_t = (\pi^{*0}_t, \pi^{*1}_t, ..., \pi^{*J-1}_t). \tag{7}$$

Expected benchmark profits are defined as the expected profits in the absence of any decreased yield due to insufficient crop break intervals. The market state is assumed to evolve via a stationary Markov process.

[6] See for example Harker et al. (2015) and Kutcher et al. (2013). Wilcox (2012) is one of the few sources that show longer dependence, but is not based on an experimental setup.

[7] An alternative, more granular market state definition would incorporate the components of expected profit (prices, benchmark yields, and costs) separately. However, this would lead to a very large state space with thousands of states.

This specification embodies a few notable assumptions. First, weather is not included as a state variable. I assume that farmers take weather as unpredictable, meaning that the weather at the time of planting is not expected to be predictive of the weather throughout the growing season. Related to this, I account for soil moisture effects only through the crop rotation effect embodied in the field state variable. As will be shown formally in Section 5.2, input costs do not depend on the individual field state. For example, I abstract from the potential use of more pesticide in continuous cropping. I further assume that crop rotation has the same effect on yields in all regions.

In this state specification, the future market state is assumed to be independent of the current choice. This is consistent with the competitive nature of commodity markets, as the vast majority of farms are small firms subject to globally determined prices.

3.1.1 Compressing the State Space

The state space definition outlined above results in a very large state space, even with an extremely simple representation of choices and profit levels.[8] Given the limited length of the time series data available, it is unlikely to observe all the relevant states. Therefore, I propose to compress the state space to the vector of *individual* expected profits:

$$x_{it} = (\pi_{it}^0, \pi_{it}^1, ..., \pi_{it}^{J-1}) \tag{8}$$

Individual expected profits combine the individual field state and regional market state into a single individual state variable (per choice). This simplification collapses the two-period dependence in the first specification to a one-period dependence[9] and makes the crop rotation effect less explicit. However, the benefit is that the state space is greatly reduced,[10] which allows for more transitions between states to be observed, making estimation more feasible.

4 Identifying the Discount Factor

I estimate the discount factor by building slightly on the identification result from Abbring and Daljord (2020). This result allows for estimation of the discount factor from choice probabilities, expected profits, and state transition probabilities. It employs an exclusion restriction on primitive utility, in this case average expected profit. In the generalized specification presented by Abbring and Daljord, the difference between the expected profits for two choice-state combinations is taken as known and the variance of the idiosyncratic error is normalized to one. This is equivalent to knowing the expected profit difference scaled by the variance of the idiosyncratic error term. With a full panel dataset this quantity could be estimated. I present a derivation in which only the average expected profit difference is known, and the variance is a parameter to be estimated. In the following,

[8]For example, with three crop choices and two levels (high/low) for expected profits there are 72 possible states.
[9]Note that I employ the two-period state dependence in the synthesis of expected yields, as described in Section 5.
[10]The example with three crop choices and two expected profit bins is reduced to eight possible states.

I will follow closely Abbring and Daljord, but without normalizing the error variance. I omit i, t, and r subscripts for brevity.

First, we can rewrite (4) as

$$v(j,x) = \pi(j,x) + \beta \sum_{x'} \mathbb{E}_{\nu'}\left[m(x') + v(j_0,x')\right] q(x' \mid j,x), \tag{9}$$

where $m(x') = \mathbb{E}_{\nu'}\left[\max_{j' \in \mathcal{J}}\left\{v(j',x') - v(j_0,x') + \nu'_{j'}\right\}\right]$ is the expected "excess" surplus of choice j over a reference choice j_0. Given the assumption that the error terms are drawn independently from a Gumbel distribution with mean zero and scale parameter α, we have:[11]

$$\mathbb{E}_{\nu'}\left[\max_{j' \in \mathcal{J}}\left\{v(j',x') - v(j_0,x') + \nu'_{j'}\right\}\right] = \alpha \ln\left(\sum_{j' \in \mathcal{J}} \exp\left(\frac{v(j',x') - v(j_0,x')}{\alpha}\right)\right). \tag{10}$$

Similarly, we can derive a form of the Hotz-Miller inversion (Hotz & Miller, 1993) where the variance is not normalized. This equation relates the difference in value functions to the choice probabilities $p(j,x)$:[12]

$$\alpha \ln\left(\frac{p(j,x)}{p(j_0,x)}\right) = v(j,x) - v(j_0,x), \quad j \in \mathcal{J}/\{j_0\}. \tag{11}$$

Combining (10) and (11), we have, simply:

$$m(x') = -\alpha \ln(p(j_0,x')). \tag{12}$$

It is useful to write (9) in matrix form. I move the choice to a subscript in vector forms for brevity. Let $\boldsymbol{v}_j$, $\boldsymbol{p}_j$, $\boldsymbol{\pi}_j$, and $\boldsymbol{m}$ be $X \times 1$ vectors stacking $v(j,x)$, $p(j,x)$, $\pi(j,x)$, and $m(x)$ for all states $x \in \mathcal{X}$. Vectors for the reference choice j_0 are indicated by subscript 0, e.g. $\boldsymbol{v}_0$. Let $\boldsymbol{Q}_j$ be the $X \times X$ matrix of transition probabilities where rows correspond to x and columns correspond to x'. $\boldsymbol{Q}(j,x)$ denotes a row of the transition matrix.

$$v(j,x) = \pi(j,x) + \beta \boldsymbol{Q}(j,x)[\boldsymbol{m} + \boldsymbol{v}_0] \tag{13}$$

Reproducing the derivation in Abbring and Daljord (2020) for exposition, we can assemble instances of this equation into a fully matrix representation, difference the corresponding equation for the reference choice, and apply (11):

$$\begin{aligned} \boldsymbol{v}_j &= \boldsymbol{\pi}_j + \beta \boldsymbol{Q}_j[\boldsymbol{m} + \boldsymbol{v}_0] \\ \boldsymbol{v}_j - \boldsymbol{v}_0 &= \boldsymbol{\pi}_j - \boldsymbol{\pi}_0 + \beta[\boldsymbol{Q}_j - \boldsymbol{Q}_0][\boldsymbol{m} + \boldsymbol{v}_0] \\ \alpha(\ln(\boldsymbol{p}_j) - \ln(\boldsymbol{p}_0)) &= \boldsymbol{\pi}_j - \boldsymbol{\pi}_0 + \beta[\boldsymbol{Q}_j - \boldsymbol{Q}_0][\boldsymbol{m} + \boldsymbol{v}_0]. \end{aligned} \tag{14}$$

[11]see Appendix A.1 for derivation.
[12]See Appendix A.2 for derivation.

Next, note that the first equation above also holds for $j = j_0$. Solving this instance for $\boldsymbol{v}_0$ gives:

$$\boldsymbol{v}_0 = [\boldsymbol{I} - \beta \boldsymbol{Q}_0]^{-1}[\boldsymbol{\pi}_0 + \beta \boldsymbol{Q}_0 \boldsymbol{m}]. \tag{15}$$

Substituting this and the vector form of (12) into (14) and simplifying, we have:

$$\alpha(\ln(\boldsymbol{p}_j) - \ln(\boldsymbol{p}_0)) = \boldsymbol{\pi}_j - \boldsymbol{\pi}_0 + \beta[\boldsymbol{Q}_j - \boldsymbol{Q}_0][\boldsymbol{I} - \beta \boldsymbol{Q}_0]^{-1}[\boldsymbol{\pi}_0 - \alpha \ln \boldsymbol{p}_0]. \tag{16}$$

Note that we have a set of equations in two unknowns (α and β), as the choice probabilities, average expected profits, and transition probabilities can all be estimated from the data. We could use this set of equations directly, however I difference this equation once again to make the result robust to a constant (across states) shift in the reference payoff $\boldsymbol{\pi}_0$.[13] Abbring and Daljord's exclusion restriction enumerates the valid choice-state combinations that we can use to assemble a set of equations to estimate β.[14] Take $x_1, x_2 \in \mathcal{X}$, $j \in \mathcal{J}/\{j_0\}$, and $l \in \mathcal{J}$ such that either $l \neq j$, $x_1 \neq x_2$, or both. Given that Q_0 exhibits single action finite dependence[15], I posit that β is uniquely identified by a set of two or more equations of the following form:[16]

$$\begin{aligned}\alpha\left[\ln\left(p(j,x_1)/p(j_0,x_1)\right) - \ln\left(p(l,x_2)/p(j_0,x_2)\right)\right] - \left[\pi(j,x_1) - \pi(j_0,x_1) - \pi(l,x_2) + \pi(j_0,x_2)\right] = \\ \beta[\boldsymbol{Q}(j,x_1) - \boldsymbol{Q}(j_0,x_1) - \boldsymbol{Q}(l,x_2) + \boldsymbol{Q}(j_0,x_2)][\boldsymbol{I} - \beta \boldsymbol{Q}_0]^{-1}[\boldsymbol{\pi}_0 - \alpha \ln \boldsymbol{p}_0].\end{aligned} \tag{17}$$

The scale parameter α complicates Abbring and Daljord's set identification result. However, since α and β do not enter the equation symmetrically, they are likely to be locally identified using a set of equations. The number of possible discount factors satisfying the equation is equal to the number of past periods that are required to determine the state transition probabilities for a given choice. Given the compressed state space specification, I expect point identification of the discount factor.

5 Data

The data required for estimation includes crop choice histories and state-specific expected profits. My approach involves combining separate datasets for each of these to construct a 'pseudo' panel dataset. I use the term 'pseudo' because the dataset for state-specific expected profits is synthesized from cost, price, and yield data that represent averages by year and region. To obtain state-specific expected profits, I adjust average yields by a factor based on the crop rotation history, where this factor is taken from the scientific crop rotation literature. This process is further explained below.

I focus on non-irrigated agricultural land in Alberta and choose regions to be soil zones to match

[13] See Appendix of Abbring and Daljord (2020).

[14] Note that the scale parameter α is a nuisance parameter that is not of direct interest here.

[15] See Abbring and Daljord (2020) for details and Arcidiacono and Miller (2020) for a general discussion.

[16] Exact single finite action dependence is not possible with estimated transition matrices, however local identification can be verified numerically.

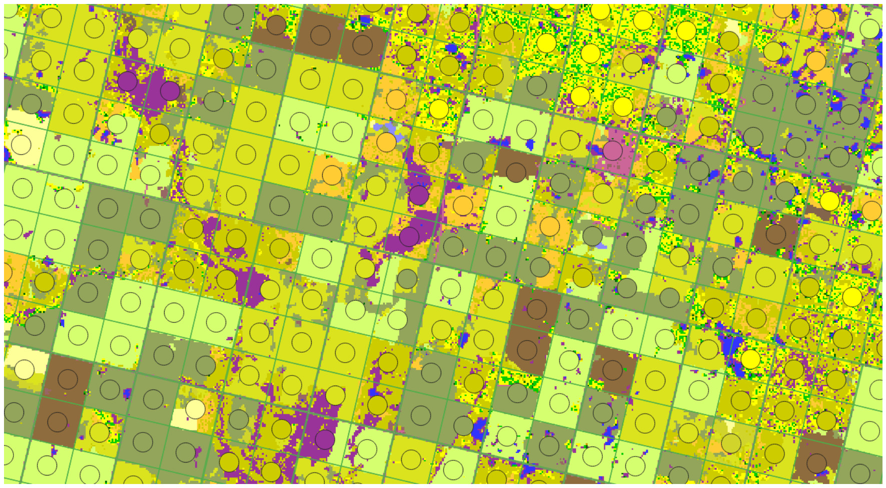


Figure 1: Sample of field crop assignment. The green grid demarcates quarter sections and the background color indicates the crop at 30 m resolution. The color of each circle indicates the assigned field crop as determined by the majority of pixels in each field.

the data for average payoffs. Fields within a soil zone typically have similar soil type and weather patterns, resulting in similar cost structure and crop choices. The data is most complete for the Brown, Dark Brown, and Black soil zones.

5.1 Crop Choice Histories

I construct crop choice histories for each quarter section of agricultural land in Alberta for the years 2011 to 2019 from the Annual Crop Inventory (Agriculture and Agri-Food Canada, 2019). This is a set of annual crop inventory maps published by Agriculture and Agri-Food Canada that are constructed from satellite imagery to provide crop classification with an accuracy of at least 85% at a resolution of 30 m for the entire Prairie Region (Fisette et al., 2014). I assign a crop choice to each quarter section (as defined by the Alberta Township System) for each year as the majority pixel value within the area. Figure 1 shows a small sample for one year, with the resulting field crop assignment indicated by the color of the circle overlaid on each quarter section. I exclude from the dataset any quarter section which was employed in a non-agricultural land-use or a pasture/forage in any year. Land is typically only used for pasture/forage if it is of marginal quality and unsuitable for cash crops, which represents a different choice problem than the one of interest here. Further simplification of crop classification is discussed in Section 5.2.4.

The soil zone for each quarter section is assigned using the Prairie Soil Zones of Canada [17] and irrigation zones are identified based on an irrigation district map obtained from the Irrigation & Farm Water Branch of the Government of Alberta. The resulting region definitions are shown in Figure 2

[17] Map file obtained from Agriculture and Agri-Food Canada (2018).

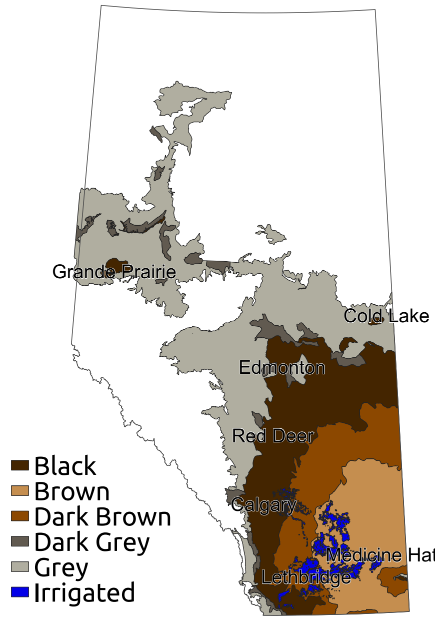


Figure 2: Regions as defined by soil zones and irrigation districts. Base map from the Government of Alberta.

5.2 Payoffs

Field state-specific average expected profits for crop choice j in region r and year t with field state k are constructed from annual soil zone-specific averages of prices p_{rt}, yields y_{rt} and costs c_{rt}: [18]

$$\pi_{rtk}(j) = p_{rt}(j) * y\left(j, k, y_{rt}(j)\right) - c_{rt}(j). \tag{18}$$

The function $y(\cdot)$ maps average expected yield to field state-specific expected yield and is described in Section 5.2.2.

5.2.1 Average Prices, Yields, and Costs

I present results from two alternative datasets for average annual prices, yields, and costs by soil zone, both obtained from reports published by the Economics and Competitiveness Branch of the Government of Alberta. Agriprofit$ Cropping Alternatives is an annual forecast published early in the year as a tool to help farmers make crop decisions (Agriculture and Forestry Alberta, 2014-2020). These forecasts are based on prior years' survey data, market trends, and expert opinion. It should be noted that the methodology is not entirely consistent over time. For example, in some years regional differences in prices only reflect differences in crop quality, but in others, they also include actual price differences between regions. This does not pose a problem if one chooses to interpret these values as average expectations in the sense of average beliefs about future payoffs rather than statistical expectations, which is plausible as they are openly available publications from an official,

[18]The assumption of competitive markets implies that there is no covariance term.

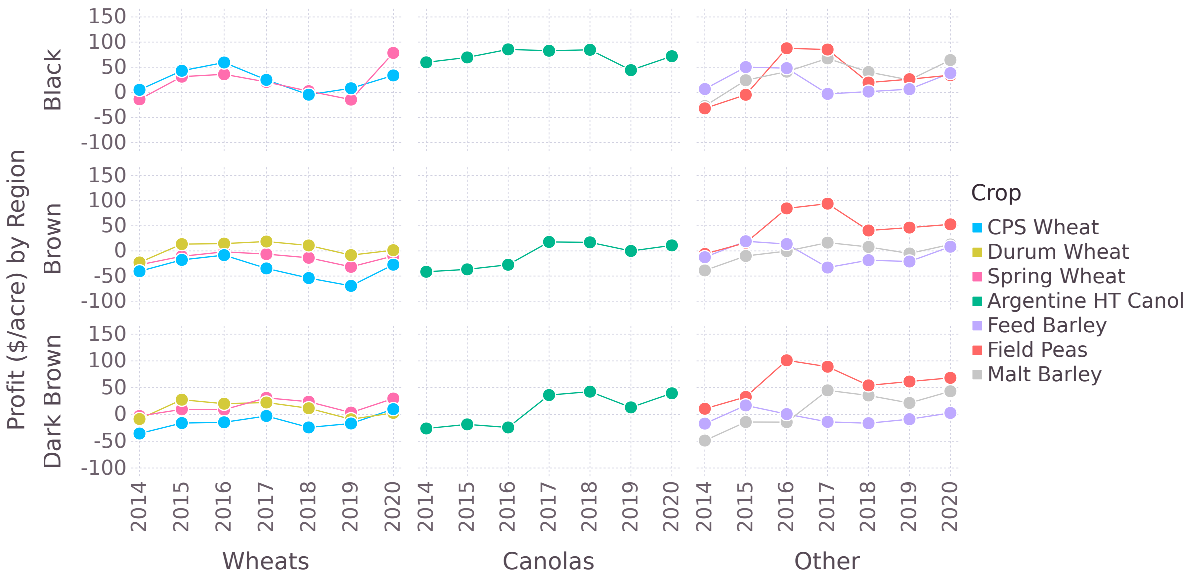


Figure 3: Profits time series from Agriprofit$ Cropping Alternatives. Adjusted to 2014 dollars.

central source. Publications span the years 2014 to 2020. The values 'Expected Yield per Acre' and 'Expected Market Price' are extracted from the PDF document tables for the averages of price and yield. To account for the mix of owned and rented land, I calculate capital costs based on 'Total Capital Costs' less half of 'Crop Share/Cash Rent'. Total costs are the sum of this value and 'Total Direct Expense.' The average expected profits calculated from these values are plotted in Figure 3 (see Appendix A.3 for the price, yield, and cost series).

Agriprofit$ Cost and Return Benchmarks for Crops and Forages is an end of season summary based on voluntary survey data (Agriculture and Forestry Alberta, 2004-2018). These reports span 2004 to 2019, with the exception of 2014. Sample sizes are generally not reported. I select the reported data average over both owned and rented land under the assumption that the sample is representative of actual tenure rates and because the disaggregated data is less complete. The values 'Yield per Acre', 'Expected Market Price', and 'Total Production Cost' are extracted from the PDF document tables for the averages of price, yield, and costs, respectively. The data also includes other crop-specific income, most importantly from crop insurance, which I add in calculating profits. The average expected profits calculated from these values are plotted in Figure 4.[19] This dataset is much less consistent in the crop varieties that are reported from year to year (see Table 3), but provides a longer series for the Black soil zone. Aggregation of crop varieties is discussed in Section 5.2.4.

5.2.2 State-Specific Expected Yield

Following crop rotation studies, I define state-specific expected yield as:

$$y\left(j, k, y_{rt}(j)\right) = \gamma_k(j) * y^*(j, y_{rt}(j), N_{krt}(j)), \tag{19}$$

[19]See Appendix A.3 for the price, yield, and cost series.

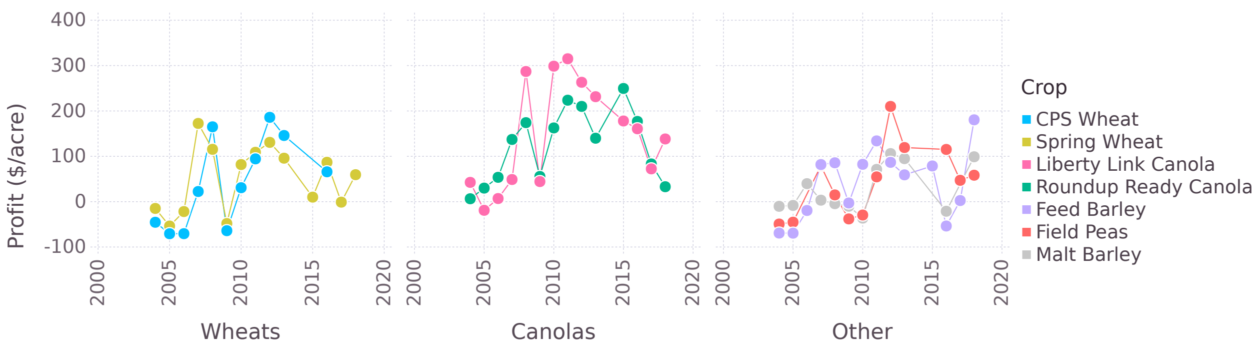


Figure 4: Profits time series from Agriprofit$ Cost and Return Benchmarks for the Black soil region. Adjusted to 2014 dollars.

Crop	**Crop Break Interval** 0	1	2	**References**
Wheat	0.11	0.00	0.00	1,2
Canola	0.18	0.02	0.00	1,3,4,5,6
Barley	0.16	0.07	0.00	2,5,7
Peas	0.14	0.11	0.00	2,9

Table 1: Yield penalties compiled from the crop rotation literature. References: 1-Gill (2018), 2-Wilcox (2012), 3-Harker et al. (2015), 4-Harker et al. (2018), 5-Williams et al. (2014), 6-Dosdall et al. (2012), 7-Arshad et al. (1999), 8-Wright (1990), 9-Nayyar et al. (2009)

where $\gamma_k(j)$ is a yield penalty based on values from crop rotation studies and $y^*(\cdot)$ is a benchmark yield in which no crop history effects are present, calculated from average yields y_{rt} and field-state populations N_{krt}. This is further explained below.

The yield penalties are provided in Table 1. These values are compiled via a thorough literature review of relatively recent crop rotation studies in all soil zones on the Canadian Prairies. Each value represents an average of reported values from at least two different sources for each crop. Due to limited data availability, the expected yield penalty for a given crop is calculated based on the last time that crop was grown (i.e. crop break interval), so each γ is not unique for every field state k. For example, $\gamma_{1,1}(j) = \gamma_{1,2}(j)$ because Crop 1 was grown in the preceding year in both states. This simplification is consistent with many scientific analyses (e.g. Wilcox (2012)), as the literature identifies disease and pests as the primary cause of decreasing yields, many of which are specific to the crop categories of interest here (Harker et al., 2015).

To obtain benchmark yields from the available average yield data, I assume that the averages are calculated from a sample with a field state distribution that is similar to that observed in the Crop Inventory data. This allows the benchmark yields to be calculated as follows:[20]

[20]Derivation in Appendix A.4.

$$y^*(j, y_{rt}(j), N_{krt}(j)) = \frac{y_{rt}(j)}{\sum_k \frac{N_{krt}(j)}{\sum_k N_{krt}(j)} * \gamma_k(j)}. \tag{20}$$

Because the datasets do not cover precisely the same years, I use the average frequency of each field state over all years in the Crop Inventory, rather than the yearly field state frequencies.

5.2.3 Synthesizing Expected Profits

In order to determine expected profits for each crop choice in each region, year, and field-state, I use the deterministic nature of field state transitions within the model. For example, a field in which Crop 1 was grown the previous year and Crop 2 was grown the year before is in field-state (1,2). Choosing Crop 3 in the current year will transition the field to state (3,1). Therefore, I can construct a synthetic population for estimating expected profits, which consists of one field in each possible field state[21] for each year and region. The expected profit of each possible choice for each field in this set is calculated using (18). For transition probabilities, I follow a similar process, but taking the Cartesian product of current field-states and current crop choices to obtain a two-period synthetic dataset of expected profits.

5.2.4 Crop Categorization

Because crops are classified with different degrees of precision between datasets, some judgment must be used to harmonize the data. For example, the Annual Crop Inventory distinguishes only between spring wheat and winter wheat, while the other sources have data on specific varieties, such as CWRS and durum. The categories in the different datasets which have reasonably complete data are listed in Table 3.[22] Furthermore, in order to maintain a tractable model I limit the choice set to three crops. Therefore, some data must be either aggregated or excluded. To begin, I consider payoff data for only wheat, canola, barley, and peas as they are clearly the most relevant crops in all soil zones, as shown in Figure 5.[23]

As it is unclear a priori whether it is more appropriate to excluded or aggregate different crop types and varieties, I present results for a number of variants, described briefly in Table 2.[24] Where data is aggregated over multiple crops, I take averages to calculate yield penalties, prices, yields, and costs. I aggregate Liberty Link and Roundup Ready canola because they are simply different proprietary varieties with distinct herbicide-compatibilities. The distinction between feed barley and malt barley is often due simply to resulting crop quality rather the seed variety planted (O'Donovan et al., 2014). I drop CPS wheat as it is a type of spring wheat. Spring wheat comprised on average 86% of all wheat grown between 2011 and 2019 in Alberta.[25] I present variants with and without

[21]There are nine states in a specification with a two-year crop history and three crop choices.
[22]I exclude winter wheat from the Wheat category as it is a relatively minor crop and it's growing season is different from other crops.
[23]The data for lentils in the Brown soil zone is not consistent over years.
[24]See Appendix A.5 for a detailed description of the variants considered.
[25]Calculated based on data for 2011-2019 from Statistics Canada (2020a).

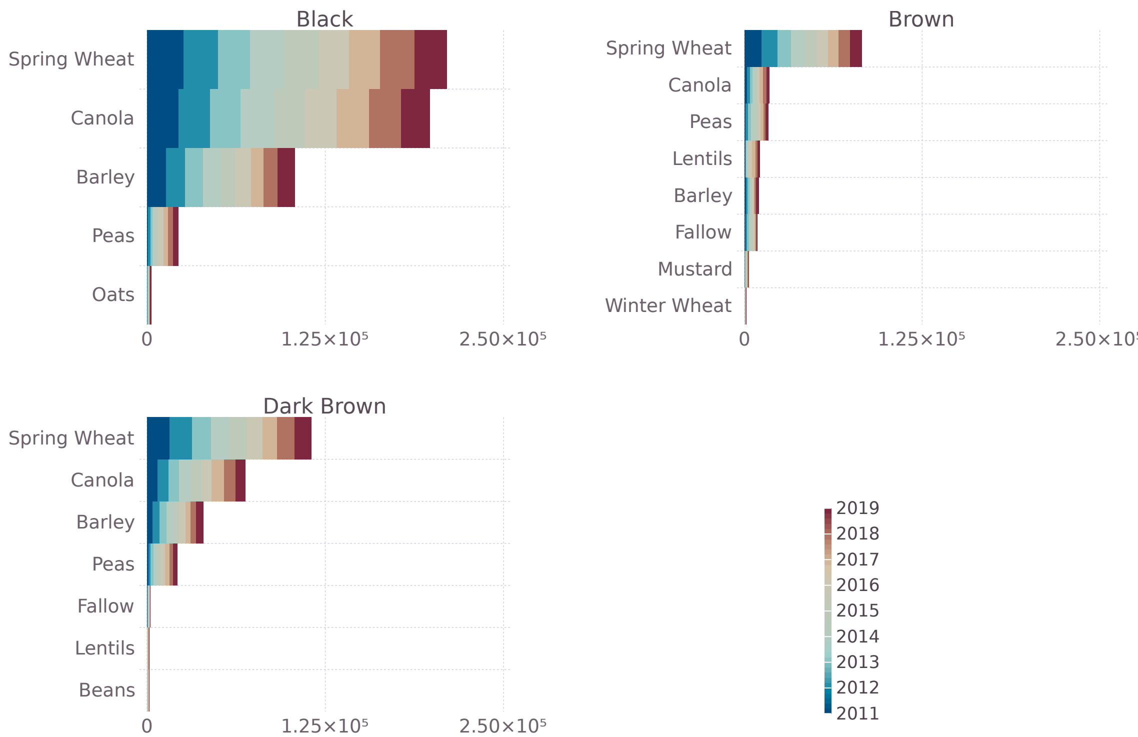


Figure 5: Frequencies of most common crops by soil zone and year. Crops excluded if less than 1% of most common crop in soil zone.

durum wheat.

Variant	Description
Preferred	Aggregates data for both barley and peas as Other category.
All Crops	Assigns all fields not labeled as Wheat or Canola to Other category.
Top 3 Crops	Uses only data from third most common crop in region for Other category.
Top 3, All	Same as Top 3 Crops, but assigns all fields not labeled as Wheat or Canola to Other.
3rd Penalty	Uses yield penalty from third most common crop for Other category.
Durum	Includes durum wheat (only for Brown and Dark Brown soil regions).

Table 2: Brief descriptions of variants considered.

5.2.5 Discretization of Profits

In order to construct discrete states, expected profit values are discretized. Due to the limited amount of data available, I bin profit values into just two levels: high (H) and low (L). I define bins separately for each crop choice, with the cutoff as the median value of the one-period synthesized profits dataset over all years and field states.[26]

[26]The median potentially creates more transitions between states than the mean.

Soil Zone	Crop Inventory 2011-2019	Cropping Alternatives 2014-2020	Cost and Return Benchmarks 2004-2013,2015-2019
Black	Spring Wheat	Spring Wheat, CPS Wheat	Spring Wheat, CPS Wheat (2015, 2017, 2018)
	Canola/Rapeseed	Argentine HT Canola	Liberty Link Canola, Roundup Ready Canola
	Barley	Feed Barley, Malt Barley	Feed Barley, Malt Barley (2015, 2017)
	Peas	Field Peas	Field Peas (2006)
Dark Brown	Spring Wheat	Spring Wheat, CPS Wheat, Durum Wheat	Spring Wheat (2016, 2017)
	Canola/Rapeseed	Argentine HT Canola	
	Barley	Feed Barley, Malt Barley	
	Peas	Field Peas	
Brown	Spring Wheat	Spring Wheat, CPS Wheat, Durum Wheat	Spring Wheat (2013)
	Canola/Rapeseed	Argentine HT Canola	
	Barley	Feed Barley, Malt Barley	
	Peas	Field Peas	
	Lentils	Lentils*, Red Lentils*	

Table 3: Crop classifications in different datasets. Years where observations are missing are in parentheses. Crop types or varieties with more than 3 years missing are excluded. *Reported data switches from Lentils to Red Lentils in 2016.

6 Estimation

Estimation consists of two stages. In the first, expected profits, choice probabilities, and transition probabilities are estimated separately. In the second, these estimates are used to determine the implied discount factor. Errors are estimated by applying the delta method at various stages. Each region is estimated separately.

Expected profits are estimated from the synthesized dataset described in Section 5.2.3 by averaging over all calculated profit values corresponding to each choice-state combination. The number of observations used to calculate each estimate varies based on the mapping from the original state specification to the compressed state specification. One implication of this is that, somewhat counter-intuitively, the expected profit of a given choice is not necessarily equal for two states with the same profit level for that choice. For example, $\pi(j_0, LLL)$ is not in general equal to $\pi(j_0, LLH)$ even though the profit state level for choice j_0 is L in both cases. A second implication of using a synthesized dataset is that the true variances and covariances of these profit estimates are not known. In order to nevertheless calculate an indication of the true standard errors, I assume zero covariance between estimates and calculate the variance of profit values in each state-choice combination. In order to report the most conservative possible estimates, I attribute the maximum variance to all estimates. This further ensures that weights in the second stage estimation are not skewed by this less than ideal error estimate.

To estimate choice probabilities, compressed states x are first assigned to the Crop Inventory data by merging it with the synthesized expected profits dataset on region, year, and field state. Choice probabilities are then estimated in two steps to allow for the calculation of standard errors. First, the unconditional state probabilities and joint state-choice probabilities are estimated as averages of indicator variables. The conditional choice probabilities are then calculated as the joint probability divided by the state probability. Standard errors are calculated using the delta method, where observations are assumed to be independent and identically distributed so that the covariance matrix of the unconditional probabilities is given by the covariance of the indicator variables divided by the number of observations.

Transition probabilities and errors are calculated in an analogous fashion to the choice probabilities, but using the synthesized dataset described in Section 5.2.3. As such, transition probabilities are based on a combination of the modeled deterministic individual field state transitions and the empirical market state evolution.[27] For the Cost and Return Benchmarks dataset, I treat the year 2015 as if it followed immediately after 2013 because 2014 is missing from the dataset.

For all first stage estimates, any states that are not observed or which have only observations into, but not out of the state (i.e. are observed only in the last year of the time series) are excluded.

The second stage estimation of β is achieved with a two-step efficient minimum distance estimator. The residual of each equation is calculated as:

[27] They are not based on the crop choice observations as this would inaccurately weight the transition probabilities based on the number of fields in each individual profit state.

$$\begin{aligned}\varepsilon_{r,(l,j,j_0,x_1,x_2)}(\alpha,\beta) = {} & \alpha\left(\ln(p(j,x_1)/p(j_0,x_1) - \ln(p(l,x_2)/p(j_0,x_2)\right) - \\ & (\pi(j,x_1) - \pi(j_0,x_1) - \pi(l,x_2) + \pi(j_0,x_2)) - \\ & \beta[\boldsymbol{Q}_j(x_1) - \boldsymbol{Q}_0(x_1) - \boldsymbol{Q}_l(x_2) + \boldsymbol{Q}_0(x_2)] \times [\boldsymbol{I} - \beta\boldsymbol{Q}_0]^{-1}[\boldsymbol{\pi}_0 - \alpha \ln \boldsymbol{p}_0].\end{aligned} \tag{21}$$

Let $\theta_r = (\alpha_r, \beta_r)$ denote the parameters to be estimated and let η_r denote the vector of expected profits, choice probabilities and transition probabilities, which is estimated in the first stage. $\hat{\theta}$ is estimated with weight matrix W by:

$$\hat{\theta} = \arg\min_{\theta} \boldsymbol{\varepsilon_r}(\theta_r, \hat{\eta}_r)' W \boldsymbol{\varepsilon_r}(\theta_r, \hat{\eta}_r), \tag{22}$$

where $\boldsymbol{\varepsilon_r}(\theta_r, \hat{\eta}_r)$ is the column vector of residuals from (21) using all valid combinations of $l, j, j_0 \in \mathcal{J}$ and $x_1, x_2 \in \mathcal{X}$, as discussed in Section 4. This includes all possible assignments of the reference choice j_0 and all numerically unique permutations. For example, setting $j = Canola, x_1 = (HHH)$ and $l = Other, x_2 = (HHL)$ is not numerically different from $j = Other, x_1 = (HHL)$ and $l = Canola, x_2 = (HHH)$ so only one of these is included in the estimation. This results in 744 equations per region if all possible states are observed. In the first step I use an identity weight matrix and in the second I use the efficient weight matrix $\left(\nabla_\eta \varepsilon_r(\hat{\theta}_r, \hat{\eta}_r)\hat{\Omega}\nabla_\eta \varepsilon_r(\hat{\theta}_r, \hat{\eta}_r)'\right)^{-1}$, where $\hat{\Omega}$ is the estimated covariance matrix for η.[28] Standard errors are calculated using the delta method.[29] It should be stressed that due to the use of aggregate data for profits, the error estimates are a best attempt to indicate the degree of uncertainty, but may be an underestimate.

7 Results

This section first discusses my preferred estimate in detail, followed by comparisons with the other variants considered. My preferred estimate uses the Cropping Alternatives forecast values and aggregates data for barley and peas in the Other category, excluding fields planted with any other crop. I prefer this dataset to the actual realized data from the Cost and Return Benchmarks dataset as I consider it to more accurately reflect farmers' expectations, which is the important parameter in determining their decisions. In addition, the Cropping Alternatives data does not suffer from missing observations in any year for any of the crops considered. In comparison with other variants using the same data, I prefer including as many other crops as possible in the Other category as it more accurately represents the potential payoffs from planting an alternative crop to either wheat or canola.

[28] I use the Moore-Penrose inverse to ensure invertability.

[29] See Appendix A.6 for a derivation of the covariance matrix and the efficient weighting matrix.

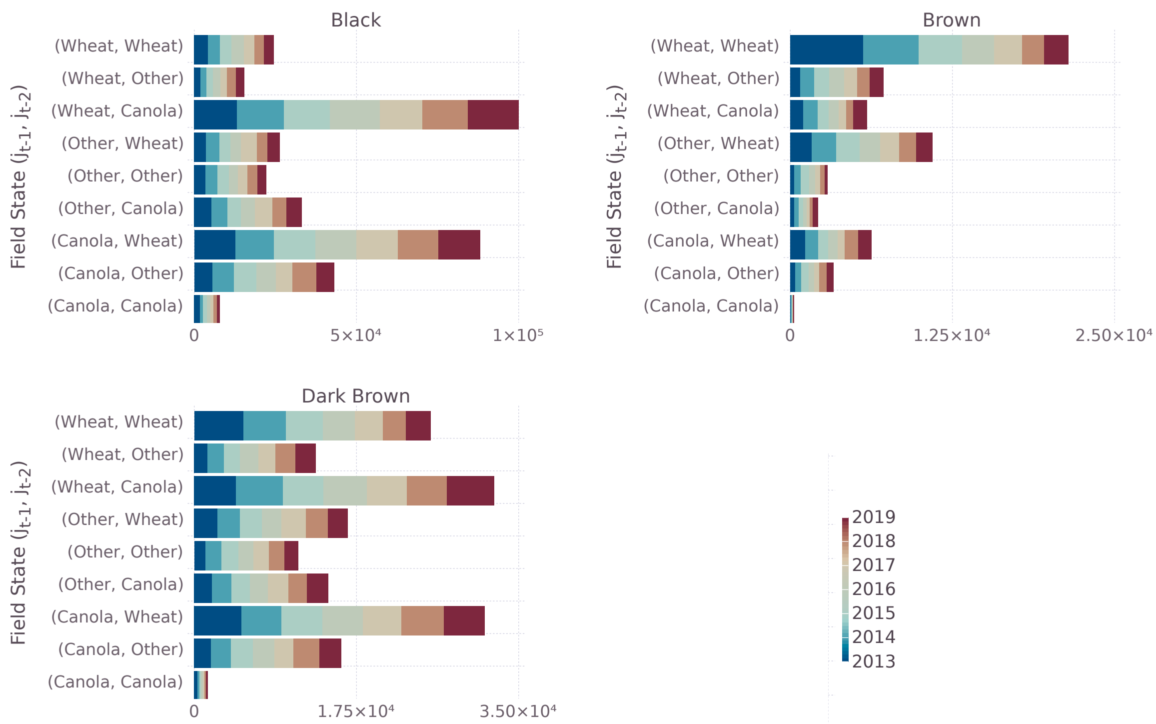


Figure 6: Frequencies of individual field states by region and year for the preferred variant. Scales vary between figures.

7.1 Crop Rotations

The frequencies of individual field states as defined for the preferred variant are shown in Figure 6. In the Black and Dark Brown regions, the most popular crop rotation by far alternates between wheat and canola, while alternating canola with either barley or peas is also common. Continuous cropping of wheat is by far the most common cropping system in the Brown region, and is a close second in the Dark Brown region. Continuous cropping of canola is rare in all the regions.

7.2 First Stage Estimates

Expected profits from the first stage estimation are presented in Figure 7. As noted previously, expected profits for a given crop are not identical between states with the same profit level for that crop. However, this variation is typically not large. Profits are notably larger and exhibit more variation between states in the Black soil region than in other regions. In all regions, the Other category is relatively highly profitable, which is somewhat surprising given that wheat and canola are more commonly grown. It is also somewhat surprising how many expected profits are negative. This may indicate the tight margins faced by farmers and the importance of cropping decisions. As discussed previously, the errors are not indicative of the true uncertainty in the profit estimates, but are used to obtain a conservative estimate of the error in the discount factor estimation.

Choice probabilities are shown in Figure 8. The most notable observation is that in the Brown

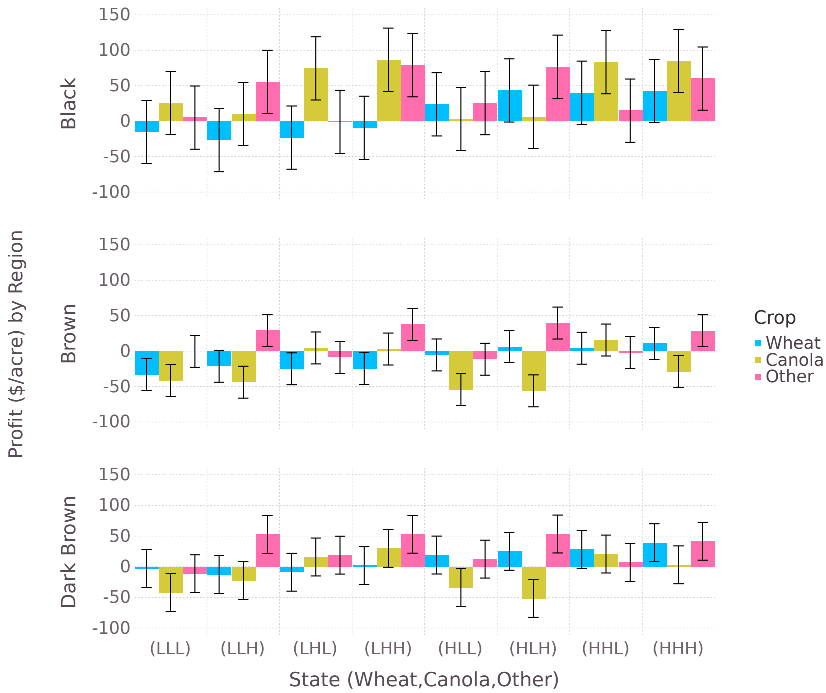


Figure 7: Estimated expected profits by state and region for the preferred variant, using the Cropping Alternatives data. Error bars indicate standard errors, calculated from the maximum profit variance per region (see Section 6).

soil region the ranking of choices is the same in all states except one. Comparing visually to the ranking of profits in each state, choices in the Brown soil region appear to be unrelated to the relative profitability of the crops. In contrast, in the Black soil region, there is an obvious correlation between profitability and choice probability for almost all states. The Dark Brown soil region exhibits a similar, but less consistent correlation. This may indicate that the model is less well suited to the Brown soil region, possibly due to the greater variety of crops and the fact that I do not have consistent data for lentils, which is one of the more popular secondary crops in that region.

Transition probabilities are depicted in Figure 9. The transition matrices are quite sparse, meaning that many transitions are not observed in the data. This is likely due to the limited size of the synthesized dataset. Peculiarities such as the guaranteed transition to a low profit state after growing canola in any state are observed with most variants when using the Cropping Alternatives data. This phenomenon is much less pronounced with the Cost and Return Benchmarks data due to the longer time span of the dataset and the greater volatility of actual profits compared to forecast profits.

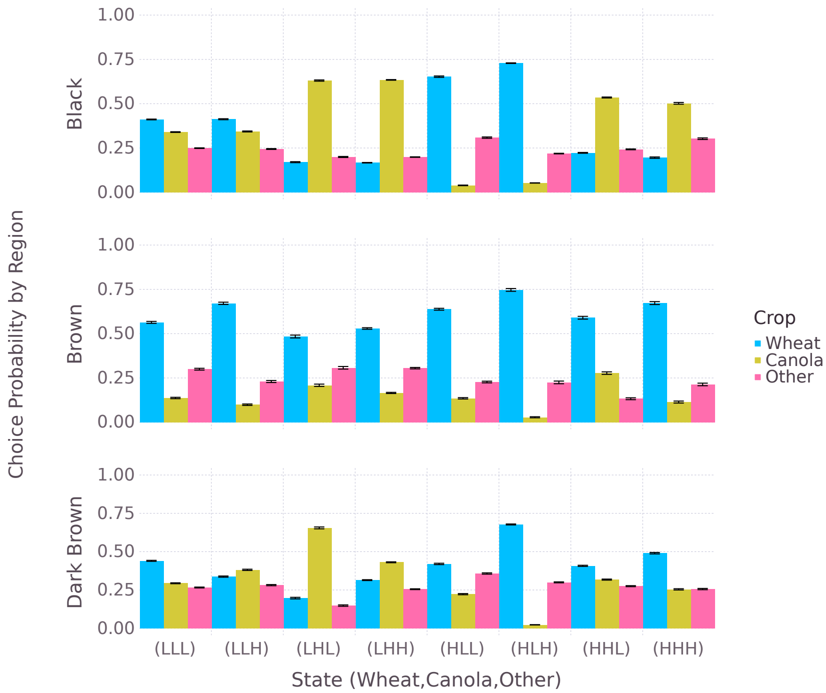


Figure 8: Estimated choice probabilities by state and region for the preferred variant, using the Cropping Alternatives data. Error bars indicate standard errors.

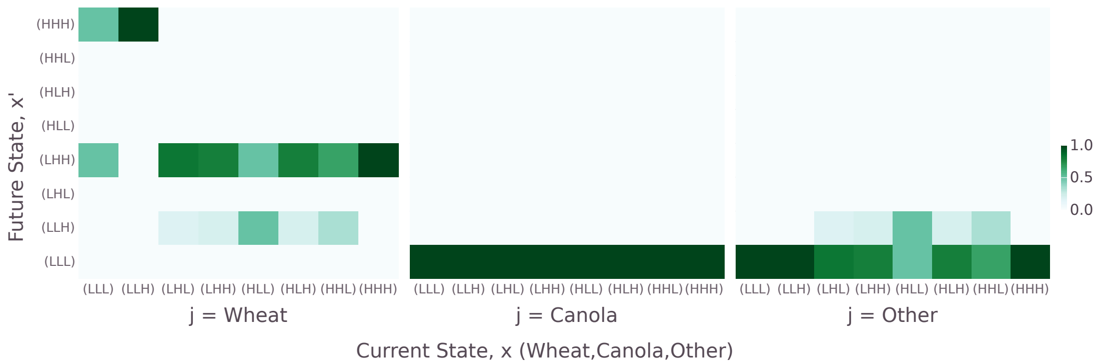
7.3 Discount Factor

The second stage two-step estimation results for all the variants considered are listed in Table 4.[30] I begin by discussing the results for the preferred variant. Using the Cropping Alternatives forecast data, the discount factor estimates range from 0.33 for the Dark Brown soil region up to 0.8 for the Brown soil region. The error term scale parameters are between 15.5 and 24.49, which correspond to standard deviations of 19.9 and 31.41, respectively. This can be interpreted very loosely as an indication of how well the model explains the data, in the sense that an implausibly large estimate would imply a large amount of variation in expected profits that is not explained by the model. The estimates for the Black soil zone from the two datasets agree within error,[31] with point estimates of 0.65 and 0.598. Unfortunately, the data was incomplete for the other regions, so no comparison is available. The variation of estimates between regions is rather unexpected, and may indicate variation in data quality or in how well the model describes the data rather than variation in farmer preferences. The larger estimate in the Brown soil region as compared to the Black soil region is consistent with the previous observation of greater correlation between current period expected profits and choice probabilities for the Black soil region. However, a similar consistency does not hold for the Dark Brown soil region.

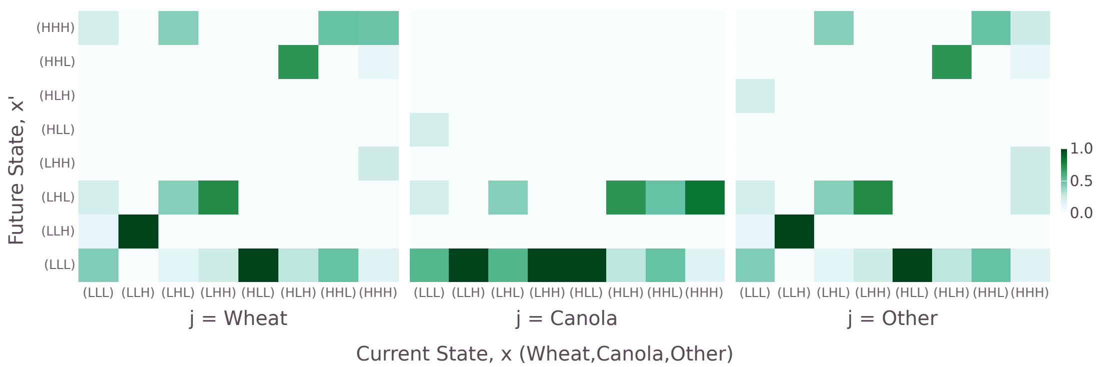
The preferred estimates are robust to minor variations in analysis parameters for some regions

[30]First step point estimates are very similar to the second step estimates, as shown in Appendix A.7. The second step acts primarily to reduce error estimates.

[31]Due to the use of aggregate data for costs, yields, and prices, errors do not reflect the total uncertainty of the estimates, but are intended as a loose indication of the uncertainty.



(a)



(b)

Figure 9: Estimated transition probabilities $q(x' \mid j, x)$ for the Black soil region for the preferred variant for a) the Cropping Alternatives dataset and b) the Cost and Return Benchmarks dataset. 'H' and 'L' indicate high and low expected profits, respectively.

Data	Variant	Region	Discount Factor	Error Scale Param.
Cropping Alternatives	Preferred	Black	0.65 (0.09)	24.49 (0.09)
		Brown	0.8 (0.1)	15.5 (0.1)
		Dark Brown	0.33 (0.04)	24.2 (0.04)
	All Crops	Black	0.65 (0.09)	25.14 (0.09)
		Brown	0.7 (0.1)	17.1 (0.1)
		Dark Brown	0.32 (0.04)	27.73 (0.04)
	Top 3 Crops	Black	-0.121 (0.005)	16.844 (0.005)
		Brown	2.3 (0.2)	11.3 (0.2)
		Dark Brown	-0.19 (0.02)	14.41 (0.02)
	Top 3, All	Black	-0.091 (0.003)	20.601 (0.003)
		Brown	1.5 (0.1)	15.9 (0.1)
		Dark Brown	0.36 (0.03)	17.14 (0.03)
	3rd Penalty	Black	0.65 (0.09)	24.2 (0.09)
		Brown	0.8 (0.1)	15.3 (0.1)
		Dark Brown	0.33 (0.04)	24.45 (0.04)
	Durum	Brown	0.27 (0.02)	19.4 (0.02)
		Dark Brown	0.38 (0.03)	24.48 (0.03)
C&R Benchmarks	Preferred	Black	0.598 (0.001)	12.975 (0.001)
	All Crops	Black	0.5997 (0.0005)	13.3713 (0.0005)
	Top 3 Crops	Black	0.637 (0.005)	22.383 (0.005)
	Top 3, All	Black	0.69 (0.03)	21.64 (0.03)
	3rd Penalty	Black	0.5934 (0.0009)	13.5578 (0.0009)

Table 4: Estimation results for all variants considered. Standard errors in parentheses.

and are somewhat less robust for others. The 'All Crops' variant assigns the uncategorized fields to the Other category when calculating the choice probabilities. This has virtually no effect on the estimates for either dataset, which is expected given the dominance of the primary crops. The '3rd Penalty' variant gives an indication of the sensitivity of the estimates to the yield penalties by applying the yield penalty for the third most common crop rather than the average yield penalties for barley and peas. Given that this is a small change,[32] it is unsurprising that the discount factor estimates are virtually unaffected.

The 'Top 3 Crops' variant uses only data from the three most common crops in the region for both profits and choice probabilities. This variant produces infeasible estimates for the discount factor when using the Cropping Alternatives data. This is somewhat surprising, but may be because the alternatives to the two primary crops are not accurately represented solely by the third most common choice. The 'Top 3, All' variant is similar, but assigns all fields not seeded with wheat or canola to the Other category. This change does not provide feasible estimates for most regions. With the Cost and Return Benchmarks data, the effect of these two variants is much smaller, and estimates remain similar though slightly higher (0.637 and 0.69, respectively, versus 0.598). This suggests that the lack of robustness with Cropping Alternatives data may be related to the limited amount of data available, as the Cost & Return Benchmarks data covers almost twice as many years. Lastly, the Durum variant includes durum wheat, which is a less common type of wheat, in the Wheat category. This has little effect on the estimate for the Dark Brown soil region, but dramatically decreases the Brown soil region estimate.

Comparing to previously reported values from the literature, which use very different techniques, the discount factor estimates from both datasets are somewhat low. The higher estimates here are in the same range as experimental measures of farmer discounting, which range from 0.7 to 0.78. Both are much lower than those obtained from estimating Euler equations on income and investment data. Taken at face value, the estimates here suggest a strong present bias present in farm operators' cropping decisions, with discounting clearly not matching interest rates. However, the data quality certainly plays a role in the current estimates. Furthermore, the unexpected variation between soil regions casts some doubt on the suitability of the model in some regions.

8 Conclusion

Time preferences play a crucial role in a wide range of contexts, and yet they are often overlooked, in part because they are difficult to measure. Common measurement methods include experimental setups and estimation of the Euler equation, but both have drawbacks, especially when considering firm behaviour. Time preferences in agriculture are important for analyzing the implications of policies regarding food security, sustainability, and environmental stewardship. I evaluate a new technique for estimating discount factors from dynamic discrete choice models by applying it to crop rotation decisions of farm operators. I find evidence suggesting a strong present bias, somewhat

[32]See Table 1.

in line with experimental estimates. One caveat to this result is that the lack of granular profit data poses a challenge. Due to the lack of panel data, I synthesize a pseudo-panel dataset for prices, yields, and costs from region-based aggregate values and crop break yield penalties from agricultural science literature. Nevertheless, I demonstrate a novel approach for estimating context-specific time preferences in agriculture, which is set to become even more promising as richer satellite data becomes available for crop-level yields.

References

Abbring, J. H., & Daljord, Ø. (2020). Identifying the discount factor in dynamic discrete choice models. *Quantitative Economics*, *11*(2), 471–501.

Abdulkadri, A. O., & Langemeier, M. R. (2000). Using farm consumption data to estimate the intertemporal elasticity of substitution and relative risk aversion coefficients. *Agricultural Finance Review*.

Agriculture and Agri-Food Canada. (2018). *Prairie soil zones of canada* [[Online; accessed 13. Apr. 2020]]. https://open.canada.ca/data/en/dataset/ac6a1e51-9c70-43ab-889f-106838410473

Agriculture and Agri-Food Canada. (2019). *Annual crop inventory* [[Online; accessed 03. Mar. 2020]]. https://open.canada.ca/en/apps/aafc-crop-inventory

Agriculture and Forestry Alberta. (2004-2018). *Agriprofit$: Cost and return profiles for selected crops* [[Online; accessed 13. Jul. 2020]]. https://open.alberta.ca/publications

Agriculture and Forestry Alberta. (2014-2020). *Agriprofit$: Cropping alternatives* [[Online; accessed 20. May. 2020]]. https://open.alberta.ca/publications/agriprofits-cropping-alternatives

Anderson, R. C., Duru, A., & Reeb, D. M. (2012). Investment policy in family controlled firms. *Journal of Banking & Finance*, *36*(6), 1744–1758.

Arcidiacono, P., & Miller, R. A. (2020). Identifying dynamic discrete choice models off short panels. *Journal of Econometrics*, *215*(2), 473–485.

Arshad, M., Franzluebbers, A., & Gill, K. (1999). Improving barley yield on an acidic boralf with crop rotation, lime, and zero tillage. *Soil and Tillage Research*, *50*(1), 47–53.

Bollinger, B. (2015). Green technology adoption: An empirical study of the southern california garment cleaning industry. *Quantitative Marketing and Economics*, *13*(4), 319–358.

Cagetti, M. (2003). Wealth accumulation over the life cycle and precautionary savings. *Journal of Business & Economic Statistics*, *21*(3), 339–353.

Claassen, R., & Tegene, A. (1999). Agricultural land use choice: A discrete choice approach. *Agricultural and Resource Economics Review*, *28*(1), 26–36.

Collard-Wexler, A. (2013). Demand fluctuations in the ready-mix concrete industry. *Econometrica*, *81*(3), 1003–1037.

De Groote, O., & Verboven, F. (2019). Subsidies and time discounting in new technology adoption: Evidence from solar photovoltaic systems. *American Economic Review*, *109*(6), 2137–72.

Dean, M., & Sautmann, A. (2014). Credit constraints and the measurement of time preferences. *Available at SSRN 2423951*.

Dosdall, L., Harker, K., O'donovan, J., Blackshaw, R., Kutcher, H., Gan, Y., & Johnson, E. (2012). Crop sequence effects on root maggot (diptera: Anthomyiidae: Delia spp.) infestations in canola. *Journal of economic entomology*, *105*(4), 1261–1267.

Dunne, T., Klimek, S. D., Roberts, M. J., & Xu, D. Y. (2013). Entry, exit, and the determinants of market structure. *The RAND Journal of Economics*, *44*(3), 462–487.

Duquette, E., Higgins, N., & Horowitz, J. (2012). Farmer discount rates: Experimental evidence. *American Journal of Agricultural Economics*, *94*(2), 451–456.

Ebert, S., Wei, W., & Zhou, X. Y. (2018). Discounting, diversity, and investment. *Diversity, and Investment (February 9, 2018)*.

Einav, L., Finkelstein, A., & Schrimpf, P. (2015). The response of drug expenditure to nonlinear contract design: Evidence from medicare part d. *The quarterly journal of economics*, *130*(2), 841–899.

Fisette, T., Davidson, A., Daneshfar, B., Rollin, P., Aly, Z., & Campbell, L. (2014). Annual space-based crop inventory for canada: 2009–2014. *2014 IEEE Geoscience and Remote Sensing Symposium*, 5095–5098.

Frederick, S., Loewenstein, G., & O'donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of economic literature*, *40*(2), 351–401.

Gill, K. S. (2018). Crop rotations compared with continuous canola and wheat for crop production and fertilizer use over 6 yr. *Canadian journal of plant science*, *98*(5), 1139–1149.

Gollier, C., & Zeckhauser, R. (2005). Aggregation of heterogeneous time preferences. *Journal of political Economy*, *113*(4), 878–896.

Gourinchas, P.-O., & Parker, J. A. (2002). Consumption over the life cycle. *Econometrica*, *70*(1), 47–89.

Harker, K., Hartman, M., Tidemann, B., O'Donovan, J., Turkington, T., Lupwayi, N., Smith, E., & Mohr, R. (2018). Attempts to rescue yield loss in continuous canola with agronomic inputs. *Canadian journal of plant science*, *98*(3), 703–716.

Harker, K., O'Donovan, J., Turkington, T., Blackshaw, R., Lupwayi, N., Smith, E., Johnson, E., Gan, Y., Kutcher, H., Dosdall, L., et al. (2015). Canola rotation frequency impacts canola yield and associated pest species. *Canadian Journal of Plant Science*, *95*(1), 9–20.

Harris, J., & Siebert, R. (2017). Firm-specific time preferences and postmerger firm performance. *International Journal of Industrial Organization*, *53*, 32–62.

Harrison, G. W., Lau, M. I., & Williams, M. B. (2002). Estimating individual discount rates in denmark: A field experiment. *American economic review*, *92*(5), 1606–1617.

Hermann, D., Mußhoff, O., & Rüther, D. (2015). *Measuring farmers' time preference: A comparison of methods* (tech. rep.). Diskussionsbeitrag.

Hotz, V. J., & Miller, R. A. (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies*, *60*(3), 497–529.

Kutcher, H., Brandt, S., Smith, E., Ulrich, D., Malhi, S., & Johnston, A. (2013). Blackleg disease of canola mitigated by resistant cultivars and four-year crop rotations in western canada. *Canadian Journal of Plant Pathology*, *35*(2), 209–221.

Kutcher, H., Johnston, A., Bailey, K., & Malhi, S. (2011). Managing crop losses from plant diseases with foliar fungicides, rotation and tillage on a black chernozem in saskatchewan, canada. *Field Crops Research*, *124*(2), 205–212.

Lence, S. H. (2000). Using consumption and asset return data to estimate farmers' time preferences and risk attitudes. *American Journal of Agricultural Economics*, *82*(4), 934–947.

Lubowski, R. N., Plantinga, A. J., & Stavins, R. N. (2008). What drives land-use change in the united states? a national analysis of landowner decisions. *Land Economics*, *84*(4), 529–550.

Magnan, A. (2015). The financialization of agri-food in canada and australia: Corporate farmland and farm ownership in the grains and oilseed sector. *Journal of Rural Studies*, *41*, 1–12.

Nayyar, A., Hamel, C., Lafond, G., Gossen, B. D., Hanson, K., & Germida, J. (2009). Soil microbial quality associated with yield reduction in continuous-pea. *Applied Soil Ecology*, *43*(1), 115–121.

O'Donovan, J. T., Grant, C. A., Blackshaw, R. E., Harker, K., Johnson, E. N., Gan, Y., Lafond, G. P., May, W. E., Turkington, T. K., Lupwayi, N. Z., et al. (2014). Rotational effects of legumes and non-legumes on hybrid canola and malting barley. *Agronomy Journal*, *106*(6), 1921–1932.

Rossi, F. (2018). Lower price or higher reward? measuring the effect of consumers' preferences on reward programs. *Management Science*, *64*(9), 4451–4470.

Rust, J. (1994). Structural estimation of markov decision processes. *Handbook of econometrics*, *4*, 3081–3143.

Scott, P. (2014). Dynamic discrete choice estimation of agricultural land use.

Statistics Canada. (2016). *Highlights and analysis* [[Online; accessed 25. Aug. 2020]]. https://www.statcan.gc.ca/eng/ca2011/ha

Statistics Canada. (2020a). *Table 32-10-0002-01 estimated areas, yield and production of principal field crops by small area data regions, in metric and imperial units* [[Online; accessed 25. Aug. 2020]]. https://doi.org/10.25318/3210000201-eng

Statistics Canada. (2020b). *Table 32-10-0153-01 total area of farms and use of farm land, historical data* [[Online; accessed 25. Aug. 2020]]. https://doi.org/10.25318/3210015301-eng

Statistics Canada. (2020c). *Table 32-10-0403-01 farms classified by farm type* [[Online; accessed 25. Aug. 2020]]. https://doi.org/10.25318/3210040301-eng

Statistics Canada. (2020d). *Table 32-10-0433-01 farms classified by operating arrangement* [[Online; accessed 25. Aug. 2020]]. https://doi.org/10.25318/3210043301-eng

Wilcox, D. (2012). *The crop rotation break interval effect in manitoba* (tech. rep.). Manitoba Agricultural Services Corporation. Portage la Prairie, MB.

Williams, C., King, J., Ross, S., Olson, M., Hoy, C., & Lopetinsky, K. (2014). Effects of three pulse crops on subsequent barley, canola, and wheat. *Agronomy Journal*, *106*(2), 343–350.

Wright, A. (1990). Yield effect of pulses on subsequent cereal crops in the northern prairies. *Canadian Journal of Plant Science*, *70*(4), 1023–1032.

Wu, J., Adams, R. M., Kling, C. L., & Tanaka, K. (2004). From microlevel decisions to landscape changes: An assessment of agricultural conservation policies. *American Journal of Agricultural Economics*, *86*(1), 26–41.

Zentner, R. P., Wall, D. D., Nagy, C. N., Smith, E. G., Young, D. L., Miller, P. R., Campbell, C. A., McConkey, B. G., Brandt, S. A., Lafond, G. P., et al. (2002). Economics of crop diversification and soil tillage opportunities in the canadian prairies. *Agronomy Journal*, *94*(2), 216–230.

A Appendix

A.1 Distribution of Maximum of Gumbel Distributed Variables

The cumulative distribution function (CDF) for the Gumbel distributed random variable Z with location parameter μ and scale parameter α is $\Pr(Z \leq z) = e^{-e^{-(z-\mu)/\alpha}}$. I have assumed that $\nu_1, ..., \nu_J$ are independent and identically distributed (i.i.d.) Gumbel$(-\alpha\gamma, \alpha)$.[33] I derive the expectation of the maximum of the choice-specific value function, beginning by writing the log probability that l is the optimal choice:

$$\begin{aligned}
\ln Pr\left(j^* = \underset{j\in\mathcal{J}}{\text{argmax}}\{v(j,x)+\nu_j\}\right) &= \ln \prod_{j\in\mathcal{J}} Pr\left(v(j,x)+\nu_j \leq v(j^*,x)+\nu_{j^*}\right) \\
&= \sum_{j\in\mathcal{J}} \ln Pr\left(\nu_j \leq v(j^*,x)+\nu_{j^*} - v(j,x)\right) \\
&= -\sum_{j\in\mathcal{J}} \exp\left(-\frac{v(j^*,x)+\nu_{j^*} - v(j,x)+\alpha\gamma}{\alpha}\right) \\
&= -\exp\left(-\frac{v(j^*,x)+\nu_{j^*}+\alpha\gamma}{\alpha}\right)\sum_{j\in\mathcal{J}}\exp(v(j,x)/\alpha) \\
&= -\exp\left(-\frac{v(j^*,x)+\nu_{j^*}+\alpha\gamma}{\alpha}\right)\exp\left(\ln\sum_{j\in\mathcal{J}}\exp(v(j,x)/\alpha)\right) \\
&= -\exp\left(-\left(v(j^*,x)+v_{j^*}+\alpha\gamma-\alpha\ln\sum_{j\in\mathcal{J}}\exp(v(j,x)/\alpha)\right)/\alpha\right),
\end{aligned}$$

where the first equality follows from the i.i.d. assumption, the third substitutes the cumulative distribution function, and the rest of the steps are straightforward algebra. Comparing the last line with the Gumbel cumulative distribution function above, we see that

$$\max_{j\in\mathcal{J}}\{v(j^*,x)+v_{j^*}\} \sim \text{Gumbel}\left(-\alpha\gamma+\alpha\ln\sum_{j\in\mathcal{J}}\exp(v(j,x)/\alpha)),\alpha\right), \tag{23}$$

which directly implies (10).

A.2 Derivation of Hotz Miller Equation

I begin with the probability of the agent choosing j^* conditional on the agent knowing the current draws of ν_j $\forall j \in \mathcal{J}$ (i.e. they are in the information set $\mathcal{I}$):

[33] The mean of a Gumbel distributed variable is $\mu + \alpha\gamma$ where γ is Euler's gamma.

$$\begin{aligned}
Pr(\text{choose } j^*|\nu_j \in \mathcal{I}\ \forall j \in \mathcal{J}) &= Pr(v(j,x) + \nu_j \leq v(j^*,x) + \nu_{j*})\ \forall j \in \mathcal{J}/j^* \\
&= \prod_{j\in\mathcal{J}/j^*} Pr(v(j,x) + \nu_j \leq v(j^*,x) + \nu_{j^*}) \\
&= \prod_{j\in\mathcal{J}/j^*} Pr(\nu_j \leq v(j^*,x) - v(j,x) + \nu_{j^*}) \\
&= \prod_{j\in\mathcal{J}/j^*} e^{-\exp\left(-(v(j^*,x)-v(j,x)+\nu_{j*}+\alpha\gamma)/\alpha\right)},
\end{aligned}$$

where I employ the i.i.d. assumption in the second equality and the Gumbel CDF in the last. Now, the unconditional probability of the agent choosing j^* is given by integrating over the probability distribution:

$$Pr(\text{choose } j^*) = \int_{-\infty}^{\infty} \left(\prod_{j\in\mathcal{J}/j^*} e^{-\exp\left(-\frac{v(j^*,x)-v(j^*,x)+\nu_{j*}+\alpha\gamma}{\alpha}\right)} \right) f(\nu_{j^*}) d\nu_{j^*}.$$

Define $z = \frac{\nu_{j*}+\alpha\gamma}{\alpha}$ so that:

$$\begin{aligned}
Pr(\text{choose } j^*) &= \int_{-\infty}^{\infty} \left(\prod_{j\in\mathcal{J}/j^*} e^{-\exp\left(-\frac{v(j^*,x)-v(j,x)}{\alpha}-z\right)} \right) e^{-z} e^{-\exp(-z)} dz \\
&= \int_{-\infty}^{\infty} \left(\prod_{j\in\mathcal{J}} e^{-\exp\left(-\frac{v(j^*,x)-v(j,x)}{\alpha}-z\right)} \right) e^{-z} dz, \\
&= \int_{-\infty}^{\infty} \left(e^{-\exp(-z)\sum_j \exp\left(-\frac{v(j^*,x)-v(j,x)}{\alpha}\right)} \right) e^{-z} dz,
\end{aligned}$$

where in the second equality I include j^* in the product, noting that $\exp(-(v(j^*,x)-v(j^*,x))/\alpha-z) = \exp(-z)$.

Finally, a change of variables $u = \exp(-z)$ allows us to evaluate the integral and simplify:

$$\begin{aligned}
Pr(\text{choose} j^*) &= -\int_{\infty}^{0} \left(e^{-u\sum_j \exp\left(-\frac{v(j^*,x)-v(j,x)}{\alpha}\right)} \right) du \\
&= \int_{0}^{\infty} \left(e^{-u\sum_l e^{-(v(j^*,x)-v(j,x))/\alpha}} \right) du \\
&= \frac{-e^{-u\sum_j e^{-(v(j^*,x)-v(j,x))/\alpha}}}{\sum_j e^{-(v(j^*,x)-v(j,x))/\alpha}} \Bigg|_0^{\infty} \\
&= \frac{1}{\sum_j e^{-(v(j^*,x)-v(j,x))/\alpha}} \\
&= \frac{e^{v(j^*,x)/\alpha}}{\sum_j e^{v(j,x)/\alpha}}
\end{aligned}$$

Taking the ratio of choice probabilities for two different choices, we can perform the Hotz-Miller inversion to obtain (11).

A.3 Price, Yield, and Cost Series

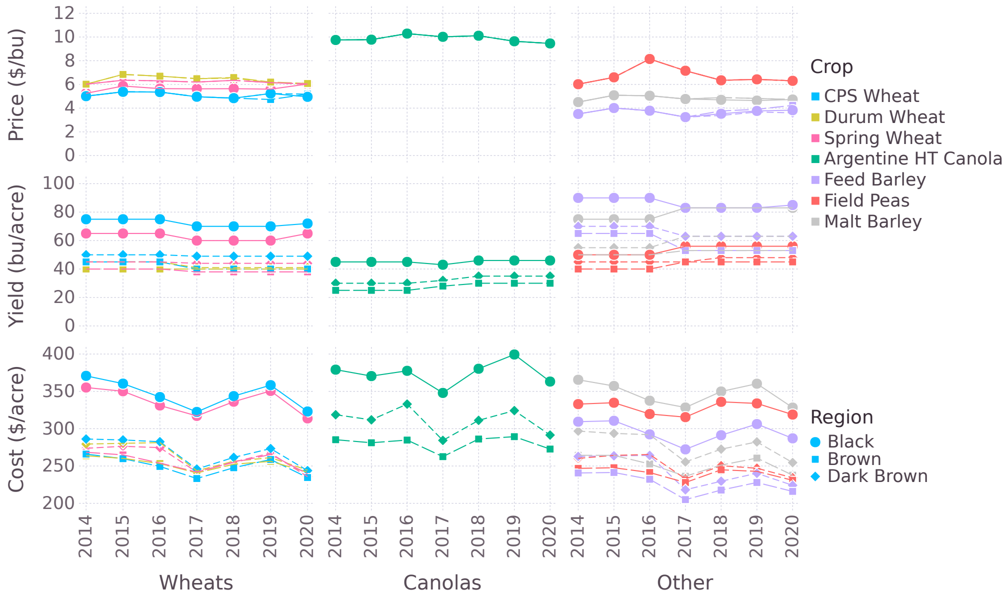


Figure 10: Price, yield, and cost time series from Agriprofit$ Cropping Alternatives. Prices and costs adjusted to 2014 dollars.

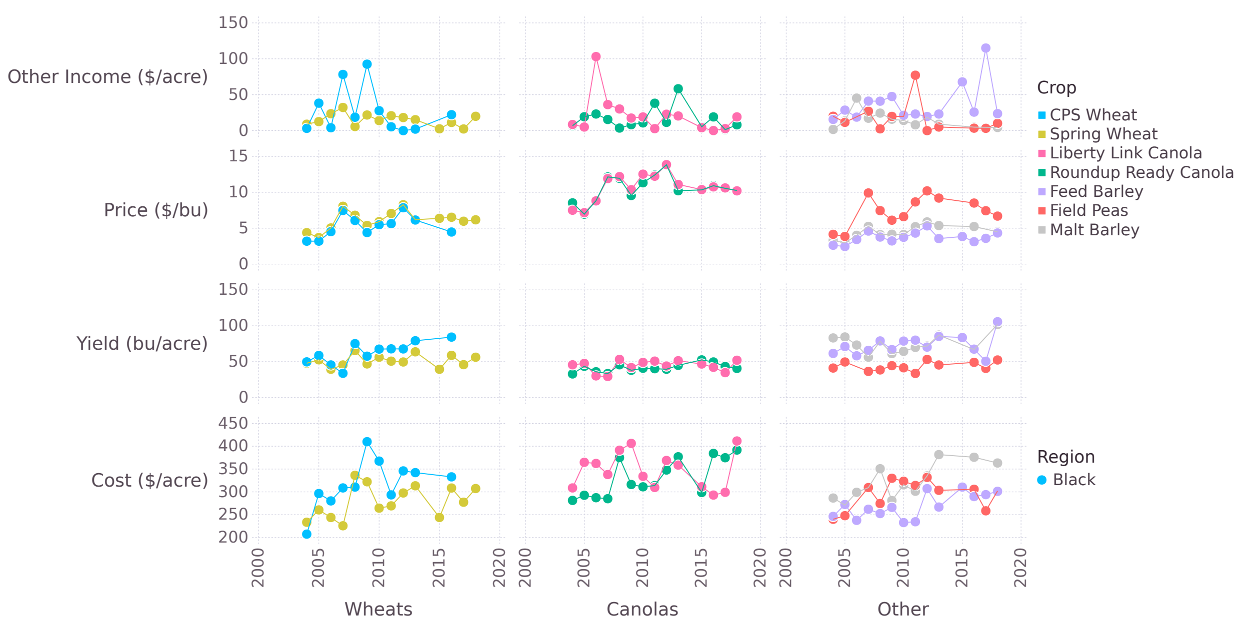


Figure 11: Price, yield, and cost time series from Agriprofit$ Cost and Return Benchmarks. Prices and costs adjusted to 2014 dollars.

A.4 Benchmark Yields

Assuming the average expected yield value $y_{rt}(j)$ represents an unbiased estimate of the population mean, we can express the population mean as the (group population-) weighted mean of field-state group averages (N denotes group size):

$$\begin{aligned} y_{rt}(j) &= \frac{1}{N_{rt}(j)} \sum_{i=1}^{N_{rt}(j)} y_{irt}(j) \\ &= \frac{1}{N_{rt}(j)} \sum_k \sum_i y_{irt}(j)|_{k_{it}=k} \\ &= \frac{1}{N_{rt}(j)} \sum_k N_{krt}(j) * \bar{y}_{krt}(j) \\ &= \frac{1}{N_{rt}(j)} \sum_k N_{krt}(j) * \gamma_k(j) * y^*_{rt}(j) \\ &= y^*_{rt}(j) * \frac{1}{N_{rt}(j)} \sum_k N_{krt}(j) * \gamma_k(j) \end{aligned}$$

where $N_{rt}(j) = \sum_k N_{krt}(j)$, $\bar{y}_{krt}(j)$ is the average yield of crop j by field state, region, and year, and $y^*_{rt}(j) = y^*(j, y_{rt}(j), N_{krt}(j))$ is the benchmark yield.

A.5 Model Variants

The classifications used for each model variant are described in Table 5.

Variant	**Region**	**Model**	**Crop Inventory / Yield Penalty**	**Cropping Alternatives**	**C&R Benchmarks**
Preferred	Black, Brown, Dark Brown	Wheat	Spring Wheat	Spring Wheat	Spring Wheat
		Canola	Canola	Argentine HT Canola	Liberty Link Canola, Roundup Ready Canola
		Other	Barley, Peas	Feed Barley, Malt Barley, Field Peas	Feed Barley, Malt Barley, Field Peas
All Crops	Black, Brown, Dark Brown	Wheat	Spring Wheat	Spring Wheat	Spring Wheat
		Canola	Canola	Argentine HT Canola	Liberty Link Canola, Roundup Ready Canola
		Other	*All Other Crops*	Feed Barley, Malt Barley, Field Peas	Feed Barley, Malt Barley, Field Peas
Top 3 Crops	Black, Dark Brown	Wheat	Spring Wheat	Spring Wheat	Spring Wheat
		Canola	Canola	Argentine HT Canola	Liberty Link Canola, Roundup Ready Canola
		Barley	*Barley*	*Feed Barley, Malt Barley*	*Feed Barley, Malt Barley*
	Brown	Wheat	Spring Wheat	Spring Wheat	Spring Wheat
		Canola	Canola	Argentine HT Canola	Liberty Link Canola, Roundup Ready Canola
		Peas	*Peas*	*Field Peas*	*Field Peas*
Top 3, All	Black, Dark Brown	Wheat	Spring Wheat	Spring Wheat	Spring Wheat
		Canola	Canola	Argentine HT Canola	Liberty Link Canola, Roundup Ready Canola
		Barley	*All Other Crops*	*Feed Barley, Malt Barley*	*Feed Barley, Malt Barley*
	Brown	Wheat	Spring Wheat	Spring Wheat	Spring Wheat
		Canola	Canola	Argentine HT Canola	Liberty Link Canola, Roundup Ready Canola
		Peas	*All Other Crops*	*Field Peas*	*Field Peas*
3rd Penalty	Black, Dark Brown	Wheat	Spring Wheat	Spring Wheat	Spring Wheat
		Canola	Canola	Argentine HT Canola	Liberty Link Canola, Roundup Ready Canola
		Other	Barley, Peas *I Barley*	Feed Barley, Malt Barley, Field Peas	Feed Barley, Malt Barley, Field Peas
	Brown	Wheat	Spring Wheat	Spring Wheat	Spring Wheat
		Canola	Canola	Argentine HT Canola	Liberty Link Canola, Roundup Ready Canola
		Other	Barley, Peas *I Peas*	Feed Barley, Malt Barley, Field Peas	Feed Barley, Malt Barley, Field Peas
Durum	*Brown, Dark Brown*	Wheat	Spring Wheat	Spring Wheat *Durum Wheat*	
		Canola	Canola	Argentine HT Canola	
		Other	Barley, Peas	Feed Barley, Malt Barley, Field Peas	

Table 5: Variants of crop type classifications and their correspondences between model and datasets. Italics highlight variations from the preferred variants. Values are averaged to aggregate data from multiple crops.

A.6 Delta Method Derivation

The second stage error estimate and efficient weighting matrix are derived as follows. The first order condition for (22) is:

$$0 = \nabla_\theta \varepsilon(\hat{\theta}, \hat{\eta}) W \varepsilon(\hat{\theta}, \hat{\eta}).$$

A Taylor expansion of the last $\varepsilon(\hat{\theta}, \hat{\eta})$ term around η_0, followed by a second expansion of $\varepsilon(\hat{\theta}, \eta_0)$ around θ_0 gives:

$$\begin{aligned} 0 &= \nabla_\theta \varepsilon(\hat{\theta}, \hat{\eta}) W \left[\varepsilon(\hat{\theta}, \eta_0) + \nabla_\eta \varepsilon(\hat{\theta}, \eta_0)(\hat{\eta} - \eta_0) + O(\|\hat{\eta} - \eta_0\|^2)\right] \\ &= \nabla_\theta \varepsilon(\hat{\theta}, \hat{\eta}) W \left[\varepsilon(\theta_0, \eta_0) + \nabla_\theta \varepsilon(\theta_0, \eta_0)(\hat{\theta} - \theta_0) + \nabla_\eta \varepsilon(\hat{\theta}, \eta_0)(\hat{\eta} - \eta_0) + O(\|\hat{\eta} - \eta_0\|^2) + O(\|\hat{\theta} - \theta_0\|^2)\right]. \end{aligned}$$

Noting that $\varepsilon(\theta_0, \eta_0) = 0$ by definition and adding and subtracting $\nabla_\eta \varepsilon(\theta_0, \eta_0)(\hat{\eta} - \eta_0)$, we have:

$$0 = \nabla_\theta \varepsilon(\theta_0, \eta_0) W \left[\nabla_\theta \varepsilon(\theta_0, \eta_0)(\hat{\theta} - \theta_0) + \nabla_\eta \varepsilon(\theta_0, \eta_0)(\hat{\eta} - \eta_0)\right] + O(E),$$

where

$$E = \max\{\|\hat{\eta} - \eta_0\|^2, \|\hat{\theta} - \theta_0\|^2, \|\nabla_\theta \varepsilon(\theta_0, \eta_0) - \nabla_\theta \varepsilon(\hat{\theta}, \hat{\eta})\| \|\hat{\theta} - \theta_0\|, \|\nabla_\eta \varepsilon(\theta_0, \eta_0) - \nabla_\eta \varepsilon(\hat{\theta}, \eta_0)\| \|\hat{\eta} - \eta_0\|\}$$

Assuming ε is twice differentiable, $O(E) = O(\|\hat{\eta} - \eta_0\|^2)$. Additionally, assuming that the first stage estimates are asymptotically normal with covariance matrix Ω:

$$\sqrt{n}(\hat{\eta} - \eta_0) \to N(0, \Omega),$$

we have that $\sqrt{n} O(E) = O_p(1/\sqrt{n}) = o_p(1)$. Rearranging, we have:

$$\sqrt{n}(\hat{\theta} - \theta_0) = \left(\nabla_\theta \varepsilon_0 W \nabla_\theta \varepsilon_0'\right)^{-1} \left[\nabla_\theta \varepsilon_0 W \nabla_\eta \varepsilon_0 \sqrt{n}(\hat{\eta} - \eta_0)\right] + o_p(1)$$

where $\nabla_\eta \varepsilon_0$ is short-hand for $\nabla_\eta \varepsilon(\theta_0, \eta_0)$ and similarly for $\nabla_\theta \varepsilon_0$. Hence,

$$\sqrt{n}(\hat{\theta} - \theta_0) \to N(0, \Sigma),$$

where

$$\Sigma = \left(\nabla_\theta \varepsilon_0 W \nabla_\theta \varepsilon_0'\right)^{-1} \left[\nabla_\theta \varepsilon_0 W \nabla_\eta \varepsilon_0 \Omega \nabla_\eta \varepsilon_0' W \nabla_\theta \varepsilon_0'\right] \left(\nabla_\theta \varepsilon_0 W \nabla_\theta \varepsilon_0'\right)^{-1}$$

The efficient weight matrix is then

$$W = \left(\nabla_\eta \varepsilon_0 \Omega \nabla_\eta \varepsilon_0'\right)^{-1}.$$

A.7 First Step Estimates

Data	Variant	Region	Discount Factor		Error Scale Parameter	
			Step 1	Step 2	Step 1	Step 2
Cropping Alternatives	Preferred	Black	0.6 (0.5)	0.65 (0.09)	25 (14)	24.49 (0.09)
		Brown	0.8 (0.6)	0.8 (0.1)	19 (7)	15.5 (0.1)
		Dark Brown	0.3 (0.2)	0.33 (0.04)	25 (11)	24.2 (0.04)
	All Crops	Black	0.6 (0.5)	0.65 (0.09)	25 (14)	25.14 (0.09)
		Brown	0.7 (0.5)	0.7 (0.1)	21 (6)	17.1 (0.1)
		Dark Brown	0.3 (0.3)	0.32 (0.04)	28 (11)	27.73 (0.04)
	Top 3 Crops	Black	-0.1 (0.5)	-0.121 (0.005)	19 (13)	16.844 (0.005)
		Brown	2 (1)	2.3 (0.2)	14 (9)	11.3 (0.2)
		Dark Brown	-0.2 (0.6)	-0.19 (0.02)	13 (8)	14.41 (0.02)
	Top 3, All	Black	-0.1 (0.5)	-0.091 (0.003)	23 (14)	20.601 (0.003)
		Brown	2 (1)	1.5 (0.1)	16 (9)	15.9 (0.1)
		Dark Brown	0.4 (0.5)	0.36 (0.03)	17 (8)	17.14 (0.03)
	3rd Penalty	Black	0.7 (0.5)	0.65 (0.09)	24 (14)	24.2 (0.09)
		Brown	0.8 (0.6)	0.8 (0.1)	19 (7)	15.3 (0.1)
		Dark Brown	0.3 (0.2)	0.33 (0.04)	25 (11)	24.45 (0.04)
	Durum	Brown	0.3 (0.2)	0.27 (0.02)	23 (6)	19.4 (0.02)
		Dark Brown	0.4 (0.2)	0.38 (0.03)	26 (11)	24.48 (0.03)
C&R Benchmarks	Preferred	Black	0.6 (0.3)	0.598 (0.001)	14 (25)	12.975 (0.001)
	All Crops	Black	0.6 (0.3)	0.5997 (0.0005)	15 (27)	13.3713 (0.0005)
	Top 3 Crops	Black	0.6 (0.5)	0.637 (0.005)	23 (28)	22.383 (0.005)
	Top 3, All	Black	0.7 (0.5)	0.69 (0.03)	27 (31)	21.64 (0.03)
	3rd Penalty	Black	0.6 (0.3)	0.5934 (0.0009)	15 (25)	13.5578 (0.0009)

Table 6: First and second step estimation results for all variants considered. Standard errors in parentheses.