Value of Farm Data in Farmland Rental Markets: Management vs. Signaling Value

Nathan D. DeLay*1, Nathanael M. Thompson2, James R. Mintert3, Todd Kuethe4, Jayson L. Lusk5

* Corresponding author.
1 Assistant Professor, Department of Agricultural and Resource Economics, Colorado State University, nathan.delay@colostate.edu
2 Associate Professor, Department of Agricultural Economics, Purdue University, thomp530@purdue.edu
3 Professor, Department of Agricultural Economics, Purdue University, jmintert@purdue.edu
4 Professor, Department of Agricultural Economics, Purdue University, tkuethe@purdue.edu
5 Vice President and Dean, Division of Agricultural Sciences and Natural Resources, Oklahoma State University, jayson.lusk@okstate.edu

Abstract

Precision farming data enhances agricultural productivity by informing site-specific resource management. These benefits may be capitalized into the underlying value of the farmland itself, raising rental rates. This paper uses a stated preference choice experiment to estimate farmers’ willingness-to-pay for farm data in farmland rental markets. Farmers are willing to pay a small premium to acquire data accrued by previous operators, depending on the field type and quality information provided by the landowner, as well as farmers’ use of precision agriculture technology. We find evidence that farm data confers both a "management value" and a "signaling value" to prospective tenants.

Keywords: farm data, precision agriculture, farmland values

JEL Codes: Q15, Q16
1. Introduction

The era of “big data” has brought attention to the value of farm data, which promises to increase crop yields while conserving inputs and minimizing environmental impacts (McFadden et al., 2023; Schimmelpfennig, 2016; Bongiovanni and Lowenberg-DeBoer, 2004). Data collection has accelerated in recent years thanks to advances in field sensors, high-resolution imagery, and decision support technologies. Between 2001 and 2016, geo-referenced yield mapping grew from less than 10% of corn acres to nearly 45% of corn acres while GPS guidance systems increased from 5% to almost 60% of corn acreage (Schimmelpfennig and Lowenberg-DeBoer, 2020).

Investment in ag-tech startups reached $52 billion in 2021, an 85% increase from the year prior. Of this, over $18 billion was invested in “upstream” firms—those developing farm management software, remote sensing, internet-of-things (IoT), robotics, and other technologies for use within the farm gate (AgFunder News, 2022). The success of farm data aggregation platforms such as Farmers Business Network (FBN) and Climate FieldView suggest data has value beyond the farm (Coble et al., 2018; Miller et al., 2018; Woodard, 2016).

Data collection in agriculture has implications for farmland markets. By revealing the underlying productivity of a field, historical data can correct information problems in farmland rental markets where the landlords and tenants have different knowledge about the underlying quality of the farmland (Seifert et al., 2021). Specifically, tenant-operators may use farm data they collect to renegotiate a farmland lease agreement. According to the Agricultural Resource Management Survey (ARMS), 5% of corn farm operators shared yield monitor data with their landlord for this purpose in 2016, up from 3% in 2010.¹

Landowners selling or renting out farmland may or may not transfer historical production and management data to the new occupant. Farm data provisions, or “data succession plans,” may
become a normal part of sale and lease arrangements in the future but for now, these practices are not common or expected (Griffin et al., 2016). Privacy concerns stemming from farm data’s ambiguous legal status, the absence of clear property rights to data, and the potential for data misappropriation likely deter these practices (DeLay et al., 2023; Coble et al., 2018).

High-quality data can benefit farm operators and owners alike by making inputs more productive, thus maximizing the net expected returns of the farmland asset (Berczi, 1981). If the accumulation of data increases farmland values, these gains could be realized through higher sale prices and cash rents, provided the landowner is willing to part with the accumulated data. But unless the data is shared prior to signing the lease or closing the sale, the data is in a black box until the new tenant takes possession of the field. A potential renter may simply perceive the landowner’s willingness to share data as a signal of farmland quality and offer to pay a premium. If so, farm data may be more akin to an educational credential in the labor market—signaling the worker’s value to prospective employers without revealing their actual productivity or educational performance—and less of an information clearinghouse like Car Fax or Kelley Blue Book in the used car market (Spence, 1974). Previous work on mandatory disclosure laws in real estate markets find that revealing even publicly available information can affect property values (Pope, 2008). A landowner will have little incentive to hand over data that reveals the farmland as a lemon, particularly if the lease is renewed yearly.

Therefore, farm data may confer value to tenants in two ways: one, as an input in profit enhancing management decisions—what we refer to as the “management effect”—and two, as a signal of farmland quality, or the “signaling effect.” The purpose of this paper is to test whether either or both sources of value might be capitalized into farmland prices through a stated preference choice experiment. We find evidence in support of both theories; farmers are willing to pay a
modest premium to rent farmland with historical farm data, subject to differences in soil heterogeneity, farmland quality information, and the precision agriculture engagement of the producer.

Existing research finds that farmland values are influenced both by productivity related characteristics and non-monetary amenities (Borchers, Ifft, and Kuethe, 2014; Nickerson, et al., 2012; Huanget al., 2006; Xu, Mittelhammer, and Barkley, 1993; Palmquist and Danielson, 1989; Chavas and Shumway, 1982). Hedonic pricing models have been used more recently to estimate the marginal economic value of various farmland characteristics, but are less useful in cases where the value of unobserved factors are of interest (Bigelow et al., 2020). In such cases, self-reported estimates or survey instruments may be best. Historical production practices can also be capitalized into land values, including organic certification (Fuller et al., 2021) and no-till crop production (Chen et al., 2022).

Others have studied the role of information in agricultural asset valuation. Using methods like those in this paper, Vestal et al. (2013) examine how providing animal genetic information influences stated and revealed preferences for breeding livestock. McCluskey (2000) shows that information problems in markets for credence goods such as organic produce can be overcome through repeat-purchasing or third-party monitoring. McCluskey’s (2000) results are applicable to other quality-differentiated markets like farmland. Because auctions in any given area are relatively infrequent, farmland rental markets are thin, leading to asymmetric information between landowners and tenants about the true quality of the land (Seifert et al., 2021; Bigelow et al., 2016; Cotteleer et al., 2008; Allen & Lueck, 1992). The high cost of acquiring information may prevent renters from overcoming these asymmetries (Pope, 2008). Kuethe and Bigelow (2018) find that
tenant-farmers can benefit from these information differences when landlords live far from the rented field by negotiating lower rental rates.

Due to its novelty, the role of farm data as a contributor to farmland value has yet to be explored. This paper contributes to the farmland literature by considering farm data as a potentially important attribute in land prices, and to the digital agriculture literature by testing a fundamental claim about farm data—that its ability to improve on-farm decision making is the source of its value to producers. Using a modified choice experiment, we elicit farmers’ willingness-to-pay (WTP) for two types of precision farming data—zone soil test data and yield monitor data—in a cash rent farmland auction scenario. We find that offering to transfer historical zone soil test data and yield monitor data to the new tenant raises farmland bids by about $4 and $5 per acre, respectively, about a 2% rental premium. WTP for farm data varies based on the characteristics of the field from which data is collected and data type. Within-field soil heterogeneity raises the value of zone soil test data, but reduces the desirability of yield monitor data. Revealing the true productivity of the farmland attenuates farmers’ WTP for zone soil data, suggesting that the collection of precision soil data may provide a quality signal in thin farmland markets where information is obscured.

Our results suggest that historically collected farm data may become a part of farmland rental markets in the future. However, the long-term effects of farm data on crop production and agricultural land values will depend on emerging property rights arrangements governing farm data, which remains unclear.
2. A Model of Farm Data Valuation

Successful application of precision agriculture (PA) technologies depends on the availability of high-quality intra-field information (Bullock et al., 2009; Tenkorang and Lowenberg-DeBoer, 2008; Bullock and Lowenberg-DeBoer, 2007; Bullock, Lowenberg-DeBoer, and Swinton, 2002; Bullock and Bullock, 2000; Bullock et al., 1998). Miller et al. (2018) categorize PA tools that generate large amounts of farm data (e.g. yield monitors, precision soil mapping, crop sensors, weather stations, etc.) as “information intensive.” For our purposes, farm data refers to parcel-level information collected on a site-specific basis as opposed to other farm-level data such as financial records. By informing site-specific management, data-intensive technologies may reduce input costs for a given level of output or extend the production frontier for a given mix of inputs. Given the investment costs and time required for collection, information intensive technologies involve a lag time between collection and action. Farm tenants may pay a premium for land that comes with detailed historical data which they can incorporate into their production processes immediately.

As an economic good, farm data has several notable characteristics that affect its usefulness in farmland markets. First, farm data is non-rival. Multiple parties can access a farm’s data without affecting any one party’s ability to do so, meaning the total value of a single farm data record may be large (Jones and Tonetti, 2020; Varian, 2018). Farm data’s non-rivalry leads to positive network externalities (Miller et al., 2018). By pooling data across multiple farms, models can be better calibrated to soil types, commodities, and climatic zones, making the total data network more valuable. While these benefits flow beyond the farm gate, the costs of collecting data and the risk of data misappropriation are primarily borne by the farm operator (Coble et al., 2018). Thus, farm data may be under-supplied in the market, limiting its ability to add value to farmland.
While farm data is clearly non-rival, it is not obviously a public good which requires non-excludability. Various cloud-based platforms have emerged to store and analyze farm data, each offering varying degrees of excludability according to their individual data use agreements (Ferrell, 2014). In this context, farm data may be better characterized as a club good where access is limited to members of the data network or those the farmer elects to grant access to, including landlords in a tenancy relationship (Birner et al., 2021; Griffin et al., 2016).

To model how farmers value data, we consider a representative farm operator \(i\) looking to expand their production by leasing an additional farm field \(j\) on a cash rent basis. Farmland in the operator’s area varies in quality. High-quality farmland—or highly productive land with high expected yields—occurs with probability \(p\) while low-quality, low yielding farmland occurs with probability \(1 - p\). We assume that low- and high-quality are exhaustive of quality types and that all farmers have the same prior knowledge of the distribution of high- and low-quality farmland in the area. The true quality type of the farmland is known only by the property owner who may or may not reveal it to prospective tenants (e.g. by reporting average historical yield). Regardless of whether the landlord reveals the field’s true quality type, several observable characteristics such as soil types, drainage information, and location are made available to potential renters. These attributes are generally related to farmland quality but do not perfectly reveal the field’s true quality type.

Assume that the previous tenant has collected detailed historical data (e.g. geo-referenced yield monitor records) for field \(j\). Assuming the landlord has a legal ownership claim to this data regardless of how it was collected, they can again choose to share or not share this information with new tenants. Farm data may provide information to the new farm operator that improves seed
selection, nutrient management, pesticide application, and other production decisions that would be otherwise unknown to a new occupant.

**Case I: Perfect Information**

If the landowner reveals the farmland’s true quality type, operator $i$ offers to pay $r_{ij} \geq 0$ per acre to maximize their utility from leasing field $j$ according to the following optimization problem:

$$\max_{r_{ij}} U_{ij} = \alpha_i + \beta_i X_j + \gamma_i^H Q_j + \gamma_i^L (1 - Q_j) + \delta_i D_j - r_{ij}, \quad [1]$$

where $\alpha_i$ is the baseline amount that that producer $i$ is willing to pay for farmland in their area, regardless of the farmland’s individual characteristics (similar to a fixed-effect); $X_j$ is a vector of observable field characteristics such as soil types and proximity to market terminals; $Q_j$ is an indicator equal to one if the farm field is high-quality and zero if the field is low-quality; and $D_j$ is an indicator equal to one if the landlord agrees to transfer digital copies of the field’s data to the new operator. The coefficients contained in the vector $\beta_i$ represent the perceived value derived from observed farmland attributes which may vary across producers. The parameter $\gamma_i^H$ captures the utility of renting high-quality farmland to producer $i$ while $\gamma_i^L$, assumed to be non-positive, captures the disutility of low-quality farmland.\(^2\)

The subjective management value generated by possessing farm data is $\delta_i$. This term can be thought of as the expected change in operating profit per acre contributed by using historical data. If the producer does not expect data to increase the profitability of renting a given field, $\delta_i$ will equal zero. This will be the case if data does not boost productivity—either by shifting the production frontier out for a given mix of inputs or reducing input usage for a given level of output.
Note that the landlord’s willingness to provide data will not be interpreted as a signal of farm quality as they, in this case, have already divulged the field’s true quality type. Therefore, the term $\delta_i$ will only capture data’s perceived value as a management tool.

**Case II: Imperfect Information**

Now suppose, as is normally the case, the landlord does not reveal field $j$’s true quality but may still agree to transfer historical farm data associated with the field. With imperfect information, the producer now maximizes their expected utility of renting field $j$, given by:

$$\max_{r_{ij}} E U_{ij} = \alpha_i + \beta_i X_j + [\gamma_i^H P_i + \gamma_i^L (1 - P_i)] \cdot D_j + \gamma_i^H p + \gamma_i^L (1 - p) \cdot (1 - D_j) + \delta_i D_j - r_{ij},$$

where $P_i$ is the subjective and conditional probability that the field is high-quality, given the presence of farm data, that is $P_i = \text{Prob}(Q_i = 1|D_j = 1)$. We assume that $P_i$ is greater than or equal to the unconditional probability of encountering high-quality farmland $p$. This assumption is the manifestation of the signaling value of data; offering farm data may increase the believed likelihood that land is high-quality, but cannot diminish it.

Hence, farm data influences the utility derived from renting farmland in two ways: one, by changing the perception of the field’s underlying quality, and two, by altering its productive capacity. If data is promised, i.e. $D_j = 1$, equation (2) reduces to:

$$\max_{r_{ij}} E U_{ij} = \alpha_i + \beta_i X_j + \gamma_i^H P_i + \gamma_i^L (1 - P_i) + \delta_i - r_{ij}. \quad [3]$$
If field $j$ does not include historical farm data, equation (2) will be:

$$\max_{r_{ij}} EU_{ij} = \alpha_i + \beta_i X_j + \gamma_i^H p + \gamma_i^L (1 - p) - r_{ij}. \quad [4]$$

The farmer’s maximum willingness-to-pay (WTP) for farmland with and without digital farm data is found by setting (3) and (4) to be greater than or equal to zero and solving for $r_{ij}$.

$$WTP_{ij} = r_{ij} = \begin{cases} \alpha_i + \beta_i X_j + \gamma_i^H p_i + \gamma_i^L (1 - p_i) + \delta_i & \text{if } D_j = 1 \\ \alpha_i + \beta_i X_j + \gamma_i^H p + \gamma_i^L (1 - p) & \text{if } D_j = 0 \end{cases} \quad [5]$$

The difference in maximum WTP between farmland with and without farm data, but where all other relevant characteristics are identical is expressed as:

$$\Delta WTP_{ij} = (\gamma_i^H - \gamma_i^L)(P_i - p) + \delta_i \geq 0. \quad [6]$$

The first term in equation (6) is the probability weighted change in farmland quality when farm data is made available which, by $P_i \geq p$ and $\gamma_i^H > \gamma_i^L$, must be non-negative. We refer to this as the “signaling value” of farm data. It depends on the prospective tenant’s subjective belief about the conditional probability $P_i$, and the utility premium for high-quality farmland. If the offer to transfer digital farm data from the old tenant does not alter the perceived quality of the farmland, then the difference $P_i - p$ will be zero. That is, farm data only creates a signaling effect if $P_i$ is strictly greater than $p$. 
The second term in (6), $\delta_i$, represents the “management value” of farm data to the producer. It too will vary across producers based on their experience with site-specific farming practices (e.g. variable-rate seed and fertilizer application) but may also depend on the observed characteristics of the field itself. For example, a high degree of within-field variability in soils or topography may increase the perceived value of yield monitor data showing the productive histories of each zone. Site-specific management is less necessary for a uniform field, making any historical farm data less valuable. If farmland characteristics influence the value of farm data, we could express (6) as:

$$\Delta WTP_{ij} = (P_l - p)(\gamma^H_l - \gamma^L_l) + \delta_i(X_j) \geq 0,$$  \hspace{1cm} [7]$$

where $\delta_i$ is now a function of producer- and field-specific factors.

To test the presence and strength of these effects, we develop a stated preference choice experiment using a novel survey design, described in the following section.

3. Survey Design & Data

To elicit farmer preferences and willingness-to-pay (WTP) for farm data, we utilize a form of conjoint analysis, in which respondents are asked to indicate preferences for products that vary along a set of characteristics. Use of choice experiments (or choice-based conjoint analysis) has become ubiquitous in the applied economics literature (e.g., de Bekker-Grob et al., 2012; Caputo and Scarpa, 2022; Louviere, Hensher, and Swait, 2000), in part because of their ability to accurately reflect observed market shares (e.g., Chang, Lusk, and Norwood, 2009; Brooks and Lusk, 2020). However, unlike markets for products like food, where consumers choose between many competing brands, farmers rarely make discrete choices between rental properties, but rather
bid to buy farmland or negotiate with farmland owners to rent. As such, we return to the older conjoint-based methods that ask respondents to rate or rank different profiles (e.g., see discussion in Holmes and Adamowicz, 2003) but instead, following approaches such as that in Ding, Grewal, and Liechty (2005) or Vestal et al. (2013) ask respondents to indicate their monetary values for each item. While this approach might result in hypothetical bias, distorting the overall mean amount a farmer indicates they are WTP for farmland, we are confident that the approach yields more accurate estimates of marginal WTP for one attribute vs. another (e.g., Haghani et al., 2021; Lusk and Schroeder, 2004), which is the focus of this paper.

A mail-in survey with a follow-up post card was sent to 5,000 crop producers in the Corn Belt region during the summer of 2021. A sample of 1,000 farm addresses from each of Indiana, Illinois, Iowa, Missouri, and Ohio was purchased from Farm Journal magazine. The target population was commercial-scale farms with a combined 500 acres or more of corn and soybeans. The sample was stratified to include 2,500 mid-size commercial farms (operations with 500-999 acres of cropland) and 2,500 large commercial farms (operations with 1,000 or more acres of cropland). Mid- and large-size farm operations were chosen because these farms are the most likely to rent farmland and bid in farmland auctions—the premise for our survey experiment. Addresses with these characteristics were randomly drawn from Farm Journal’s list frame of subscribers and provided to the researchers. All personally identifiable information was strictly anonymized. Each survey recipient received $2 cash to encourage participation and a follow-up post card was sent as a reminder to complete the survey. Five-hundred twenty-five surveys were returned, of which 430 were usable for a response rate of 8.6%.

Within the survey, we simulated a farmland rental auction using the aforementioned conjoint-based approach. The survey prompt asked participants to imagine they were a commercial
producer looking to lease additional farmland on a cash rent basis to expand their operation. Respondents were presented with six 80-acre farm fields available for cash rent in their area, each with a different combination of field attributes. Following Vestal et al. (2013), participants were instructed to write in the maximum amount they would pay (per acre) to lease each field. Each field description included an aerial photo of the field (standardized for all fields to not influence farmers’ stated bids), the accompanying soil map, and several field attributes listed in text. The field options were formatted to approximate actual farmland auction flyers that would be familiar to commercial producers (see Figure A1 in the supplementary appendix for an example of the auction flyer and the full survey instrument.).

Field attributes were selected based on our theoretical model to elicit willingness-to-pay for farm data under different information scenarios and for different field types. See Table 1 for a complete listing of attributes and their respective experimental levels. The first attribute is the composition of soil types within the field. Each field option presented survey takers with one of two color-coded soil maps: a highly-varied profile with nine heterogeneous soil types, or a relatively uniform profile consisting of two similarly colored soil types. The degree of soil variability will likely influence how valuable farm data is in improving the management of the field. The second attribute, tile drainage, can significantly improve the profitability of farmland and raise bids for land (Schnitkey et al., 2022). Fields were either listed as having pattern tile drainage installed, or no drainage information was provided, implying the field is un-drained.

[[Insert Table 1 about here]]
To separate the signaling effect of farm data from its management value, we include a measure of field quality in some field descriptions, but not others. This allows us to estimate WTP for farm data when the underlying quality of the field is revealed by the landowner (as in the perfect information case above) and when quality is unknown (as in the imperfect information case). However, this presents a challenge to the researchers because soil productivity, and soil productivity rating systems vary geographically. The state of Illinois, for example, uses its own soil productivity index that ranges from a low of 100 to a high of 147 while Iowa uses the Corn Suitability Rating on a scale from 5 to 100. States without their own rating systems may defer to the USDA’s National Commodity Crop Productivity Index (NCCPI) which ranks crop-specific soil productivity from 0 to 1. Even controlling for the state, farmers will interpret these ratings relative to their own location. A producer in south-eastern Indiana will find farmland with a NCCPI rating of 0.5 highly attractive while a producer in west-central Indiana is accustomed to farmland with NCCPI ratings above 0.8.

To provide an objective measure of soil quality and to be consistent with our theoretical framework, we opt for a simple approach using a hypothetical “field-level soil productivity rating.” Where provided, fields are either described as having a “high” or “low” soil productivity rating. If no quality information is provided for the field, the productivity rating is not listed, representing the “imperfect information” scenario. However, we acknowledge that this artificial rating may not perfectly represent the information cases described in our theoretical framework.

Two common types of precision farming data were chosen as attributes of the field: zone soil test data and yield monitor data—each provided for the past five years if offered by the landowner. Zone soil testing draws samples from designated management areas within the field that share certain soil and crop characteristics. Unlike grid-based soil sampling—often associated
with field-level management—zone soil data is used to inform precision farming techniques such as variable rate seed and chemical applications. Yield monitoring, which emerged in the early 1990s, has become the most popular precision agriculture practice in U.S. row crop production (Schimmelpfennig, 2016). A yield monitor automatically collects GPS-referenced crop yields as the combine moves through the field harvesting grain. If tracked over multiple years, yield monitor data can reveal persistently high- and low-performing areas of the field, which can be managed accordingly. Like zone soil data, historical yield monitor data is a useful input in precision farming management decisions. If a field option did not include one of these types of data, it was simply not mentioned in the field description. It was made clear to the survey taker that any soil data or yield monitor data offered by the landowner would be provided after signing the lease and not beforehand.8

With four attributes at two levels each (soil variability, drainage, zone soil data, and yield monitor data) and one with three levels (soil quality), the full factorial design would require respondents to evaluate $2^4 \times 3 = 48$ field options. Instead, we select a sub-set that allows us to estimate all main effects and the interaction effects between both farm data types and soil quality as well as between farm data and soil variability. We selected among the full factorials to maximize D-efficiency given the constraint that we wanted a design with 18 total farm field options, blocked into three sets of six fields each (i.e., each respondent evaluated six different fields with varying attributes). The ultimate design had a D-efficiency score of 90%.9

In addition to the WTP exercise, farmer participants were asked a series of questions regarding their farm’s demographics, production practices, land tenure, and existing farm data usage. Summary statistics for the 430 usable survey responses are shown in Table 2. On average, farms bid about $218 per acre for farmland in our choice experiment, with a minimum bid of $10
per acre and a maximum of $500 per acre. Farms in our sample tend to be larger, partly by construction of the list frame, with 65% of sampled farms having 1,000 acres or more of cropland in production and 8% with 5,000 acres or more. The age distribution and educational attainment of farmers in our sample is generally consistent with the broader farming population (USDA NASS, 2017).

Most farms in our sample (81%) lease some farmland from at least one landowner, and 45% rent 50% or more of the cropland they operate, confirming that our sample is representative of commercial-scale operations with experience in farmland rental markets. Farms that rent cropland were asked to select their average rental payment across all current lease agreements from a list of binned options in $50 increments (see question 9 in the appended survey). Three-quarters of surveyed farms pay between $150 and $300 per acre, with 17% renting farmland below $150 per acre and 8% paying over $300 per acre, consistent with average cash rental rates in the targeted states.

Figure 1a shows the distribution of bids across all 2,580 evaluated fields, while Figure 1b plots the distributions of participants’ bids in the farmland auction experiment against their reported rental rates. The graph confirms that stated preferences for farmland are largely shaped by prevailing rental rates in the farmers’ local market. One exception appears among farmers paying over $300 per acre for cropland. These farms tend to under-bid relative to what they currently pay.
Farmers were asked to indicate their current precision farming and data collection practices. Consistent with larger operations, our sample is significantly more technologically advanced than crop producers overall. Yield monitors, generally considered an entry-level precision farming technology, are used by over 90% of surveyed farms. GPS guidance is used by 86% of farms in our sample compared to 33% of corn farms nationwide as of 2016 (USDA ARMS, 2016). Rates of variable rate technology (VRT) and aerial drone adoption are similarly higher than national estimates. The collection of soil data (via grid or zone sampling) and yield monitor data is performed by 84% and 85% of participating farms, respectively, and nearly 60% use a farm data software product hosted by an agricultural technology provider (ATP). Data sharing with agronomists, input suppliers, ATPs, and other farm service providers is common (87% of farms in our sample report doing so), though sharing data with a landlord is less common at 37% of farms.

4. Empirical Methods

Following from our theoretical framework and incorporating the survey design, we estimate producers’ WTP for the farmland attributes described above with the following regression equation estimated via OLS:

\[
WTP_{ij} = \alpha_i + \beta_1 Var_j + \beta_2 Drain_j + \gamma_1 HQ_j + \gamma_2 LQ_j + \delta_1 SoilData_j + \delta_2 YMDdata_j + \varphi_1 (SoilData_j \times Var_j) + \varphi_2 (SoilData_j \times HQ_j) + \varphi_3 (SoilData_j \times LQ_j) + \omega_1 (YMDData_j \times Var_j) + \omega_2 (YMDData_j \times HQ_j) + \omega_3 (YMDData_j \times LQ_j) + \varepsilon_{ij}, \tag{8}
\]
where $WTP_{ij}$ is the stated maximum amount per acre that producer $i$ would pay to rent farm field $j$. The producer fixed-effect $\alpha_i$ acts as a baseline amount, controlling for any regional heterogeneity in farmland rental markets, as well as unobservable operator-specific characteristics such as farmland preferences, experience with rental auctions, and risk-aversion that affect the average rate a producer would likely bid. $Var_j$ is a dummy variable equal to one if the soil contour map is highly varied and zero if the soil map is uniform. The presence of tile drainage is represented by the indicator $Drain_j$. High- and low-quality fields are represented by the dummy variables $HQ_j$ and $LQ_j$, respectively. Note that because the soil productivity rating can be missing (i.e. the auction flyer does not reveal the underlying quality of the field) $HQ_j$ and $LQ_j$ do not necessarily sum to one, and the coefficients $\gamma_1$ and $\gamma_2$ are interpreted as differentials in WTP relative to a “no field quality information” baseline. The random error $\varepsilon_{ij}$ captures any random variation in bids not explained by the farmland’s attributes or producer characteristics.

$SoilData_j$ and $YMData_j$ equal one if field $j$ advertises zone soil data or yield monitor data included with the field. In isolation, the coefficients $\delta_1$ and $\delta_2$ represent farmers’ WTP for each type of data on homogeneous farmland and when the field’s true quality type is kept hidden. To test for differences in WTP for farm data in the presence of revealed farmland quality, we interact both $SoilData_j$ and $YMData_j$ with the high- and low-soil productivity variables. Similarly, we estimate the value of farm data for heterogeneous soil types by interacting the farm data indicators with the variable soils dummy. The coefficients associated with these interaction terms allow us to test for the “signaling value” and “management value” of each farm data type.

To see this, suppose the field’s soil productivity is left un-reported, i.e. $HQ_j = LQ_j = 0$, and either or both forms of digital farm records are promised, i.e. $SoilData_j = 1$ or $YMData_j = 1$. In this imperfect information case, both the “signaling” and “management” values of farm data...
are subsumed by the parameters $\delta_1$ and $\delta_2$, as in equation (5) above for $D_j = 0$. Now suppose the field’s soil productivity rating is revealed to be high by $HQ_j = 1$. Agreeing to transfer farm data does not signal any additional information about the farmland’s underlying quality in this case as it is already known to be a highly productive field (as in equation (5) for $D_j = 1$). If the farm data variables enter the model significantly, and remain so when $HQ_j = 1$, i.e. $\hat{\delta}_1 + \hat{\phi}_2 > 0$ and $\hat{\delta}_2 + \hat{\omega}_2 > 0$, then data is still considered valuable under perfect information and this would be evidence for the management value of farm data.$^{10}$ The same logic applies when the field is revealed to have low soil productivity ($LQ_j = 1$) and the relevant tests are $\hat{\delta}_1 + \hat{\phi}_3 > 0$ and $\hat{\delta}_2 + \hat{\omega}_3 > 0$.

Including the interaction terms between farm data and soil variability allows a further test of the management value theory of farm data. Presumably, soil sample data and yield monitor data are made more actionable when soil types are heterogeneous as these data sources reveal within-field differences in productivity and soil needs. Positive and significant estimates of $\phi_1$ and $\omega_1$ would lend further support for farm data as a valuable management tool.

The above specification assumes that WTP for farm data is constant across producers. However, as alluded to in our theoretical framework, the value of farm data is likely heterogeneous depending on farmers’ engagement with precision agriculture. To explore this, we estimate the WTP for farm data separately by the types of precision agriculture technologies (PATs) used by the operation. PAT adoption tends to take place sequentially—farmers first adopt basic PATs, to which they add more advanced technologies over time (Miller et al., 2017; Khanna, 2001). A producer’s stage in this process can be assessed by looking at the bundle of PATs adopted at a specific point in time.

The survey asks producers if they currently use three “standard” PATs: yield monitoring (YM), GPS guidance/auto-steer (GPS), and variable rate technology (VRT), and two “advanced”
PATs: unmanned aerial vehicles (UAV)/drones and agricultural data software products. We sort farms into one of five groups based on their responses: 1) farms that use no PATs; 2) farms using a single standard technology (YM, GPS, or VRT only); 3) farms with two standard PATs (YM + GPS, YM + VRT, or GPS + VRT), 4) farms adopting all three standard PATs (YM + GPS + VRT); and 5) farms using one or more advanced PAT (UAV, ag-data software, or both).11

We create dummy variables based on this classification, then interact them with the farm data field attributes and estimate the main-effects model as:

\[
WTP_{ij} = \alpha_i + \beta_1 Vr_{ij} + \beta_2 Drain_{ij} + \gamma_1 HQ_{ij} + \gamma_2 LQ_{ij} + \delta_1 SoilData_{ij} + \delta_2 YMData_{ij} + \sum_{k=1}^{4} \lambda_k (SoilData_{ij} \times PAT_{ik}) + \sum_{k=1}^{4} \psi_k (YMData_{ij} \times PAT_{ik}) + \epsilon_{ij}, \tag{9}
\]

where \(PAT_{ik}\) is a dummy variable representing producer \(i\)’s PAT bundle ranging from 1 (use of a single standard PAT type) to 4 (use of “advanced” technologies). In this specification, the parameters \(\delta_1\) and \(\delta_2\) represent the WTP for farm data among non-adopters while \(\lambda_k\) and \(\psi_k\) are the average differences in WTP between non-adopters and those further along the technology adoption curve.

Given the limited number of field options participants were asked to evaluate and the relatively large number of fixed-effects to be estimated, the above specification may be under-powered. However, the D-efficiency experimental design provides a sufficient number of combinations of survey attributes to allow for identification of the main and interaction effects of interest with maximum precision. For example, each survey included all possible combination of each type of data with the soil variability and soil quality attributes.
We estimate equations (8) and (9) with Stata’s “xtreg” command to accommodate fixed-effects at the farmer level. Though a Hausman test favored the use of fixed-effects over random-effects, estimating the above equations using a random-effects model produces virtually identical results. We also estimated a pooled OLS specification, controlling for observed farm-specific characteristics. Again, the results of this approach are highly similar to our main fixed-effects models. A complete discussion of both alternatives can be found in the supplementary appendix.

A survey distribution error led to an uneven number of survey versions returned and an uneven geographic distribution of farms in the sample. For example, version 1 of the survey, which was sent mostly to farms in Indiana and Ohio, makes up 38% of the sample, while version 3, made up mostly of Illinois and Missouri farms, makes up 27%. This may bias our WTP estimates due to geographic differences in farmland markets, precision agriculture engagement, or familiarity with farm data. We correct for this using a simple non-parametric bootstrapping procedure, in which we randomly sample 100 surveys, without replacement, of each version out of the 430 usable total and estimate our regression equations. We repeat this procedure 100 times, taking the simple averages of the coefficients and their standard deviations to bootstrap the standard errors.

5. Results

Our main estimation results are shown in Table 3. For comparison, we first present a specification with only the main effects, omitting interaction effects. The main effect coefficients represent farmers’ average WTP for each field attribute (see Figure 2 for a visual depiction). For zone soil test data and yield monitor data, these are the average WTP for each type of farming data, across all farm field types and soil quality information sets. On average, farmers are willing to pay an
additional $3.92 (95%CI [3.26, 4.70]) per acre for farmland that comes with historical zone soil test data. The stated premium for farmland with historical yield monitor data is higher, averaging $5.15 (95%CI [4.37, 5.88]) per acre. Computed at the sample’s mean bid, these WTP values represent farmland rental premiums of 1.8% and 2.4%, respectively. Our main effects specification suggests that landowners who offer farm data can command a positive, albeit modest, rental premium.

The marginal impacts of the other field characteristics, shown in Figure 2, generally conform to expectation. Relative to a no-information baseline, farmland with low soil productivity (low quality) is worth about $21.29 (95%CI [-22.66, -19.92]) less per acre in farmland auctions. By contrast, bids for high-quality farmland are $30.47 (95%CI [28.78, 31.16]) higher on average than under imperfect information. Soil maps with heterogeneous soil types (i.e. “varied soils”) reduce stated bids for farmland by $10.19 (95%CI [-11.70, -8.68]) per acre. Tile drainage however, enjoys a WTP of $14.49 (95%CI [13.04, 15.94]) per acre, an increase of about 7% relative to un-drained farmland.

The main effects model shows that farmers ascribe some value to farm data in farmland transactions, but it cannot tell us where this value comes from. The interacted model in Table 3 disaggregates the main effects into a signaling effect (the influence of farm data on farmers’ perception of the field’s productivity) and a management effect (driven by the potential for data to increase future profitability). To see this, we analyze the cumulative effects of farm data for each
combination of field characteristics, starting with zone soil test data. See Figure 3 for a visual
depiction of the net WTP values.

The coefficient on zone soil data equals $3.58—close to the main effect reported in the un-
interacted model. Interpreted independently, this is farmers’ WTP for soil data if the field’s soils
are homogeneous (uniform soil map) and the true quality of the field is unknown (soil productivity
is neither revealed to be “high” nor “low”). Assuming soil quality remains unknown to the farmer,
their marginal WTP for zone soil test data rises by $4.81 to $8.39 per acre if the field is
heterogeneous (varied soil map). This increase—equal to a 4% rental premium—indicates that
historical precision soil data conveys a non-trivial management value to prospective operators.14

Revealing the true quality of the field has a significant impact on farmers’ WTP for zone
soil test data, which falls by $3.06 for low-productivity fields and by $4.73 for high-productivity
fields. In both revealed-information scenarios, the net WTP for soil data becomes statistically
indistinguishable from zero for fields with a uniform soil profile. That is, if farmers have full
information about land quality and there is no within-field heterogeneity, offering historical zone
soil sample data does not affect the rental price of the field. Owners of heterogeneous fields that
transfer soil data to the new tenant may still command a rental premium, even if the true quality
of the field is disclosed to prospective operators. The marginal WTP for zone soil data in such a case equals $5.33 per acre for un-productive farmland vs. $3.66 for highly productive farmland.

Moving on to yield monitor data, we again find that farmers’ WTP for this type of farm data depends on the characteristics of the field from which it was collected. For historical yield monitor data offered with homogenous farmland under imperfect information (no soil productivity information is given), farmers bid an additional $5.73 on average to acquire it. Unlike zone soil data, the marginal WTP for yield monitor data shrinks by $3.66 to an estimated $2.07 per acre for fields with varied soil types, though the effect remains statistically significant. This WTP for yield monitor data is worth about 1% of the average rental bid.

The negative coefficient on the interaction between yield monitor data and varied soils is a surprising result given the potential for yield monitor data to enhance the management of heterogeneous soils. Instead, our results suggest that yield monitor data is most valuable when paired with a low-variability soil map. The interaction between yield monitor data and low-productivity soils is positive, though not statistically significant, suggesting that WTP for yield monitor data is not different between low-productivity fields and fields where the true soil quality is unknown. However, if the farmland’s soil productivity rating is high, farmers’ WTP for yield monitor data increases by $4.45. The amplifying effect of yield monitor data accompanying high-quality farmland suggests a complementary relationship between this type of field data and high soil quality. Given the soil productivity index is an average for the entire field, yield monitor data could reveal where the most productive areas are within the field—adding to its value. Alternatively, yield monitor data may simply serve as confirmation that the field is indeed high-productivity as advertised.15
Results of the WTP by PAT adoption model (equation (9) above) are reported in Table 4. In general, producers using multiple precision technologies value zone soil test data more than non-adopters, but differences between groups beyond two PATs are insignificant. Conversely, WTP for yield monitor data does not bear an obvious relationship to PAT usage. Differences in WTP between non-adopters and adopters of advanced PAT bundles are small, with the exception of farms using only a single basic PAT who bid on average $3.76 less per acre for farmland with yield monitor data than those with no precision agriculture engagement. However, this difference is not statistically significant.

\[
[[\text{Insert Table 4 about here}]]
\]

We perform several robustness checks which can be found in the supplemental appendix to this paper. First, we use alternative sampling procedures to address the unevenness of the survey distribution. These include post-stratification weighting based on the origin and version of the surveys received, as well as several alternative sample balancing and bootstrapping techniques. Specifically, we augment our main sampling procedure to require that bootstrap samples also be balanced by farm size and state-of-origin to correct for any correlation between farm characteristics and the survey version each participant received, which may bias our WTP estimates. We also test the sensitivity of our results to the choice of sample size within each bootstrap sample. In virtually all cases, the main results presented in this paper are robust to changes in sampling design.

We also test various alternative specifications. First, to reduce the potential for hypothetical bias, which can exaggerate WTP values, we restrict our estimation sample to farms that currently
rent farmland. Results from this restricted model are largely comparable to our main results. Second, we explore the effects of clustering our standard errors at different thresholds, including the state and county levels. Ultimately, we conclude that these approaches are not appropriate for our data given the limited geographic scope.

Finally, an abiding assumption throughout this paper is that fixed-effects allow us to identify WTP for field attributes by controlling for unobserved farmer characteristics. The most important of these unobserved variables are the local farmland rental market conditions in the farmer’s area and their individual preferences for farmland. Our estimates may be under-powered however, given that we have over 400 farms and only six observations per farm. For comparison, we test two flexible alternatives to our primary fixed-effects approach. In the first, we retain the panel structure of our data but use a random-effects model as opposed to a fixed-effects model, while in the second, we use a pooled OLS model with relevant farm- and farmer-specific control variables. Both approaches generate results that are consistent with the fixed-effects results presented here. See the supplementary appendix for a detailed explanation of these and all other robustness checks.

6. Discussion & Conclusions

Our results confirm that producers generally attribute some value to historically collected farm data when in the market to lease farmland. Farmers state a modest average WTP—about $4 and $5 per acre for 5 years-worth of zone soil test data and yield monitor data, respectively—to acquire farm data collected by a previous tenant. For the typical field, this works out to bid premiums of about 2%. Traditional soil sampling costs between $8 and $16 per acre depending on the grid size with zone soil sampling and analysis services being considerably more expensive. The WTP we
identify is half or less than half the cost of a tenant-operator performing the task themselves. However, soil sampling is typically performed once every three years, implying a cost of $2.67 to $5.33 per acre on an annualized basis. Looked at in this way, the average WTP for data detected in this paper is comparable to the cost of traditional soil sampling. Our findings imply that tenant-operators view the availability of soil sampling test results as a substitute for paying to conduct the analysis independently.

WTP for farm data is subject to the amount of soil quality information provided by the landowner and the variability of soils within the field, each with differing effects according to the type of farm data. Revealing farmland quality information affects the value of data types differently. WTP for zone soil test data diminishes greatly when the field’s productivity rating is displayed in the farmland description. This attenuation is particularly large for fields with uniform soil profiles, among which soil test data has a net zero effect on farmers’ stated bids. Conversely, revealing the farmland to be of low productivity does not statistically change farmers’ WTP for yield monitor data, but raises bids for high-productivity farmland by about $4.45 per acre.

A notable result is the opposing marginal effects of historical farm data when soil types are varied. We show that farmers will pay $4.81 more per acre to acquire zone soil test data that accompanies a heterogeneous field than they are for the same data if the field is uniform. For yield monitor data, this pattern is reversed. While farmers express a positive WTP for yield monitor data among both soil types, uniform fields command a significantly higher premium than varied fields ($5.73 per acre vs. $2.07 per acre).

Ultimately, these results provide evidence that farm data demonstrates both a management value and signaling value to producers. The clearest evidence of a signaling value comes from the interaction between zone soil test data and quality information. Farmers value soil data most when
they have imperfect information about the true productive capacity of the field, suggesting that zone soil data imparts some signal of farmland quality. Support for the management value theory is found in that producers are willing to pay more to acquire zone soil test data with heterogeneous fields. By contrast, yield monitor data does not appear to carry the same signal of quality and appears to have less of a management value to producers when soils are variable.

A potential explanation for these contrary results may be that different types of farm data impart distinct information to potential operators. Producers may treat zone soil data as complementing variable soils—allowing them to identify and manage the multiple soil types within the field, while yield monitor data may reveal the true variability underlying a soil map that appears uniform at first glance. In other words, producers may simply perceive varied soil maps as the “complete picture” while uniform maps are seen as limited, incomplete, or outdated. In such a scenario, yield monitor data may substitute for variable soil maps.

Complicating the signaling vs. management value story is the result that farmers’ precision agriculture engagement has a positive influence on WTP for zone soil test data but does not affect WTP for yield monitor data. This implies that zone soil test data is more informative to existing precision farming users than historical yield monitor data. Yield outcomes are subject to the influence of the operator and PAT bundle users may perceive this data as less valuable from a management perspective. Farmers using at least two basic precision technologies are willing to pay between $5 and $8 more per acre for zone soil test data than farms using a single technology or less, but among PAT adopters, there is no difference in WTP between adopters of basic bundles and more advanced bundles. In fact, the point estimates of WTP for zone soil data falls slightly as farms adopt more advanced technology packages; active PAT users may discount other producer’s data, preferring to collect their own.
Single-technology users may have adopted precision technologies passively—failing to integrate a broader precision farming system. For example, a farm may own a yield monitor that came standard with their latest combine harvester purchase, but they do not actively use the technology to generate yield maps and inform variable applications. For these passive PAT adopters, zone soil test data is no more valuable than it is to non-adopters. However, the absence of an increase in WTP for data as farms use increasingly complex PATs suggests that the relationship between farm data usage and value in farmland markets is dichotomous—farmers actively using PATs value historical zone soil test data but among active users, this value does not differ significantly as intensity of PAT usage changes.

Another noteworthy result is the difference in bids between low- and high-productivity fields. It suggests that the utility farmers derive from high-quality farmland is asymmetrically larger than the disutility of operating low-quality farmland. It also shows that owners of un-productive farmland have a nearly $20 per acre incentive to keep that information hidden. In a real-world farmland auction, the owner of a low-productivity field may know the NCCPI rating for soils within the field or its actual production history (APH) used to determine their crop insurance premiums but has an incentive to suppress that information ex-ante, or reveal only the minimum information required by the auctioneer.

Historical farm data could become a commonly recognized attribute of farmland that is transferrable when farmland changes hands. Not unlike the “mineral estate,” farm data could come to occupy a new “digital estate,” legally separable from the “surface estate” from which it is collected (DeLay et al., 2023). This paper shows that this new digital asset could be capitalized into land values when offered as part of farm real estate transactions. However, such a scenario depends on the establishment of clear property rights to farm data. This could come in the form of
data-use clauses being incorporated into farmland lease contracts. But widespread adoption appears to be far off. Only 6% of producers have ever had a lease that specified ownership over data collected from a rented field according to a nationwide survey (Purdue University-CME Group Ag Economy Barometer, 2021). Data interoperability poses an additional challenge. Different data systems that do not communicate make the transfer of data from previous tenant to landowner to new tenant even more difficult.

Farm data capitalization may also have a dynamic component. The management value of farm data likely depreciates over time as the new tenant’s data and experience crowds out any historical data inherited from the previous operator. Conversely, the signaling value of data, by imparting information of profitability in every future period, could raise farmland values permanently. These dynamics may be better examined in the context of farmland sales values as opposed to rental markets, which is beyond the scope of our study.

These challenges, combined with the newness of precision farming and the hypothetical nature of the experiment, may explain the modest WTP to acquire farm data found in our analysis. And yet we find that in many cases, farmers do express an interest in farmland with historical precision data. As data collection on farms becomes more common, and if absentee landownership gives way to large data-driven institutional investors, farmers may become conditioned to the scenario presented in our experiment. This increased awareness raises their WTP, which in turn increases the likelihood of landowners offering farm data to new tenants or buyers.

This analysis has several shortcomings that limit our ability to understand these forces. First, out of a desire to keep the experiment simple for farmers, we did not vary the length of time farm data has been collected from the field in question. Instead, we standardized data history to five years. Varying the data history would allow us to test how collection length affects the
desirability of data. Second, other field attributes, such as the soil map and whether pattern tile drainage has been installed, may themselves provide a signal of underlying soil quality in the absence of explicit soil productivity information, or alter the perceived value of farm data. We do not incorporate these interactions in the experimental design as our focus is on the quality-signaling potential of farm data. Relatedly, the profitability of precision farming likely rises when multiple data sources are layered together. Farmers, particularly those using precision farming practices, may pay more for any one type of data if combined with complementary data. Our current experimental design does not allow us to explore these data-on-data interaction effects. Finally, an overriding limitation of this study is its hypothetical nature based on stated preferences, which may limit the external validity of our results. Opportunities to estimate the value of data using revealed preferences methods may grow as farm data collection becomes more common. Ultimately, the extent to which farm data is incorporated in farmland markets will evolve in response to emerging uses by producers and landowners alike.

Acknowledgements

The authors thank the two anonymous reviewers for their helpful comments and suggestions on an earlier version of this paper. The authors would also like to thank the Purdue Center for Commercial Agriculture Advisory Council for their helpful feedback on an early version of the survey instrument and Sara Zahn for her assistance with the graphic design and layout of the survey. Funding for this research was provided by a Purdue University AgSEED grant and the Purdue Center for Commercial Agriculture.
References


<table>
<thead>
<tr>
<th>ATTRIBUTE</th>
<th>LEVELS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soil map profile</td>
<td>Varied: 9 heterogeneous soil types shown w/ high contrast between types</td>
</tr>
<tr>
<td></td>
<td>Uniform: 2 similar soil types w/ low contrast</td>
</tr>
<tr>
<td>Drainage</td>
<td>Pattern tile drainage installed</td>
</tr>
<tr>
<td></td>
<td>No tile drainage installed</td>
</tr>
<tr>
<td>Field-level soil productivity rating</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>No information provided</td>
</tr>
<tr>
<td>Soil data&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Zone soil test data provided for the past 5 years</td>
</tr>
<tr>
<td></td>
<td>No soil sample data available</td>
</tr>
<tr>
<td>Yield monitor data&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Yield monitor data provided for the past 5 years</td>
</tr>
<tr>
<td></td>
<td>No yield monitor data available</td>
</tr>
</tbody>
</table>

<sup>a</sup> Farm data provided by the landlord upon signing the lease.
### Table 2. Farm Data Willingness to Pay Survey - Summary Statistics

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Farmland bid ($ per acre)</td>
<td>218.03</td>
<td>68.51</td>
<td>10</td>
<td>500</td>
</tr>
<tr>
<td><strong>Farm size</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1,000 acres</td>
<td>0.35</td>
<td>0.48</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1,000-1,999 acres</td>
<td>0.30</td>
<td>0.46</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2,000-4,999 acres</td>
<td>0.27</td>
<td>0.44</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5,000+ acres</td>
<td>0.08</td>
<td>0.26</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Operator demographics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;35 years</td>
<td>0.04</td>
<td>0.19</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>35-49 years</td>
<td>0.17</td>
<td>0.38</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>50-64 years</td>
<td>0.41</td>
<td>0.49</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>65+ years</td>
<td>0.38</td>
<td>0.49</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>College or higher&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.60</td>
<td>0.49</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Farmland lease practices</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Currently lease farmland</td>
<td>0.81</td>
<td>0.40</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Lease &gt;50% of farmland operated</td>
<td>0.45</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Precision agriculture adoption</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yield monitor</td>
<td></td>
<td></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>GPS guidance/autosteer</td>
<td>0.86</td>
<td>0.34</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Variable rate technology (VRT)</td>
<td>0.76</td>
<td>0.43</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Drone/UAV</td>
<td>0.33</td>
<td>0.47</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Ag-tech software</td>
<td>0.59</td>
<td>0.49</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Data collection and analysis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collect yield monitor data</td>
<td>0.85</td>
<td>0.36</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Collect soil sample data</td>
<td>0.84</td>
<td>0.36</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Collect drone or satellite imagery</td>
<td>0.28</td>
<td>0.45</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Share data w/ service provider</td>
<td>0.87</td>
<td>0.34</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Share data w/ landlord</td>
<td>0.37</td>
<td>0.48</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

*Notes: <sup>a</sup> Highest level of educational attainment among all full-time employees of the farm, including owner/operators.*
Table 3. Willingness-to-Pay for Farm Data in Farmland Markets

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Coef.</th>
<th>SE</th>
<th>Coef.</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Quality</td>
<td>-21.29***</td>
<td>(0.70)</td>
<td>-19.23***</td>
<td>(1.02)</td>
</tr>
<tr>
<td>High Quality</td>
<td>30.47***</td>
<td>(0.86)</td>
<td>30.05***</td>
<td>(1.09)</td>
</tr>
<tr>
<td>Varied Soils</td>
<td>-10.19***</td>
<td>(0.77)</td>
<td>-10.28***</td>
<td>(0.96)</td>
</tr>
<tr>
<td>Drainage</td>
<td>14.49***</td>
<td>(0.74)</td>
<td>14.78***</td>
<td>(0.82)</td>
</tr>
<tr>
<td>Zone Soil Data</td>
<td>3.92***</td>
<td>(0.38)</td>
<td>3.58***</td>
<td>(1.38)</td>
</tr>
<tr>
<td>Soil Data × Varied Soils</td>
<td></td>
<td></td>
<td>4.81***</td>
<td>(1.38)</td>
</tr>
<tr>
<td>Soil Data × Low Quality</td>
<td></td>
<td></td>
<td>-3.06**</td>
<td>(1.27)</td>
</tr>
<tr>
<td>Soil Data × High Quality</td>
<td></td>
<td></td>
<td>-4.73***</td>
<td>(1.21)</td>
</tr>
<tr>
<td>Yield Monitor Data</td>
<td>5.14***</td>
<td>(0.36)</td>
<td>5.73***</td>
<td>(1.31)</td>
</tr>
<tr>
<td>YM Data × Varied Soils</td>
<td></td>
<td></td>
<td>-3.66***</td>
<td>(1.35)</td>
</tr>
<tr>
<td>YM Data × Low Quality</td>
<td></td>
<td></td>
<td>1.14</td>
<td>(1.35)</td>
</tr>
<tr>
<td>YM Data × High Quality</td>
<td></td>
<td></td>
<td>4.45***</td>
<td>(0.98)</td>
</tr>
<tr>
<td>Constant</td>
<td>208.38***</td>
<td>(1.76)</td>
<td>207.65***</td>
<td>(2.22)</td>
</tr>
</tbody>
</table>

Producer Fixed-Effects          Yes   Yes
Observations                     2,580 2,580
Producers                       430   430
Bootstrap sample size           300   300
Bootstrap samples               100   100

R-squared                       0.53 (0.01) 0.54 (0.01)

Notes: *** p<0.01, ** p<0.05, * p<0.1. Coefficients, std. errors, and R-squared values estimated using non-parametric bootstrap sampling procedure. Each bootstrap sample contains 300 producers (surveys) drawn from the full sample of 430 without replacement. The sampling procedure forces an even number of each survey version (100 for each of 3 versions) for a total of 300 surveys per sample.
Table 4. Willingness-to-Pay for Farm Data by Precision Ag Technology Engagement

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Coef.</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Quality</td>
<td>-21.14***</td>
<td>(0.75)</td>
</tr>
<tr>
<td>High Quality</td>
<td>30.52***</td>
<td>(0.98)</td>
</tr>
<tr>
<td>Varied Soils</td>
<td>-10.29***</td>
<td>(0.71)</td>
</tr>
<tr>
<td>Drainage</td>
<td>14.46***</td>
<td>(0.94)</td>
</tr>
<tr>
<td>Zone Soil Data (No Prec Ag Tech)</td>
<td>-1.29</td>
<td>(1.70)</td>
</tr>
<tr>
<td>Soil Data × 1 Standard PAT(^a)</td>
<td>2.96</td>
<td>(2.14)</td>
</tr>
<tr>
<td>Soil Data × 2 Standard PATs</td>
<td>7.73***</td>
<td>(2.28)</td>
</tr>
<tr>
<td>Soil Data × 3 Standard PATs</td>
<td>6.16***</td>
<td>(2.11)</td>
</tr>
<tr>
<td>Soil Data × Advanced PAT(^b)</td>
<td>5.22***</td>
<td>(1.76)</td>
</tr>
<tr>
<td>Yield Monitor Data (No Prec Ag Tech)</td>
<td>4.67***</td>
<td>(1.72)</td>
</tr>
<tr>
<td>YM Data × 1 Standard PAT</td>
<td>-3.76</td>
<td>(2.33)</td>
</tr>
<tr>
<td>YM Data × 2 Standard PATs</td>
<td>0.55</td>
<td>(2.11)</td>
</tr>
<tr>
<td>YM Data × 3 Standard PATs</td>
<td>0.37</td>
<td>(2.09)</td>
</tr>
<tr>
<td>YM Data × Advanced PAT(^b)</td>
<td>0.60</td>
<td>(1.83)</td>
</tr>
<tr>
<td>Constant</td>
<td>208.43***</td>
<td>(2.01)</td>
</tr>
</tbody>
</table>

Producer Fixed-Effects: Yes
Observations: 2,580
Producers: 430
Bootstrap sample size: 300
Bootstrap samples: 100
R-squared: 0.54 (0.01)

Notes: *** p<0.01, ** p<0.05, * p<0.1. Coefficients, std. errors, and R-squared values estimated using non-parametric bootstrap sampling procedure. Each bootstrap sample contains 300 producers (surveys) drawn from the full sample of 430 without replacement. The sampling procedure forces an even number of each survey version (100 for each of 3 versions) for a total of 300 surveys per sample. \(^a\) Standard PATs include yield monitoring, guidance systems, and VRT. \(^b\) Advanced PAT adopters use UAV/drones and/or ag data software platforms.
Figure 1a. Distribution of Stated Farmland Bids

Figure 1b. Distribution of Bids by Current Rental Rate

Note: triangle=predicted bid based on midpoint of current rental rate group, bar=median bid.
Figure 2. Willingness-to-Pay for Farmland Attributes from Main Effects Model
Figure 3. Willingness-to-Pay for Digital Farm Data by Field Information from Interacted Model
While the proportion of all farms that share yield monitor data with their landlord is low, the share of farms that share data with their landlord among those that both rent farmland and have a yield monitor is about 14%.

Note that if $\gamma_i^L = 0$, there is no disutility associated with low-quality farmland per-se. Rather, the producer considers low-quality farmland the baseline against which high-quality land is compared. The parameter $\gamma_i^H$ in this case is the premium producer $i$ places on high-quality farmland.

This region was selected to target farms primarily commercial-scale corn and soybean operations with relatively similar production and agro-climatic conditions.

The survey was evaluated and approved through Purdue University’s Institutional Review Board (IRB) No. IRB-2021-727 which is available from the authors upon request.

Surveys returned without all six farmland bid experiments filled out were deemed non-usable. Eleven surveys were returned with some, but not all, experiments filled out. We exclude these surveys from our regression estimation, though including them does not materially affect our results.

We focus our analysis on the cash rent market which represents the vast majority of cropland lease agreements (Bigelow et al., 2016).

Our method of asking for bids (rather than choices) is preferrable for two reasons. One, as previously indicated, farmers that rent cropland—our primary population of interest—are accustomed to bidding for farmland through the kind of first price auction that our experiment simulates. This is the most direct way of eliciting farmers’ reservation values. Second, it reduces the number of options required to be presented to the survey taker because the rental price, stated by the farmer participant, is not included as an attribute of the experimental design. This allows us to include other field attributes and estimate their interacted effects.

In practice, sharing data with multiple prospective renters is not practical. Offering to transfer historical data from the previous tenant-operator to the new occupant upon signing the lease is the most realistic option. Data ownership questions certainly arise with type of arrangement (e.g. previous data collectors may have a legal claim to the data and prevent transfer to the new tenant). We abstract away from these issues for simplicity. See DeLay et al. (2023) for an examination of these property rights issues.
The D-efficiency, or D-error, criteria maximizes the efficiency of the parameter estimates required by the experimental design when a perfectly orthogonal design is not practical or achievable. This approach allows for a limited amount of correlation among attributes.

One could argue that by promising to transfer farm data with the field after the field’s quality is revealed, the landlord is signaling that the farmland is of higher quality than is implied by its soil productivity rating. This would be especially true if the true quality is low. In this case, prospective tenants may still regard farm data as a signal of quality beyond what is suggested in the auction flyer. However, the signaling portion of farm data’s value is significantly diminished relative to the management value when soil quality—even a poor quality—is known.

As an alternative, we used the number of precision agriculture technologies (out of the five possible) used by the farm which produces similar results to the specification discussed here.

Due to an error, the recipient list was ordered by zip code, then divided into thirds (one for each version) as opposed to randomly assigning a survey version to each individual address. As a result, the three survey versions were not evenly distributed across the five targeted states and an uneven number of survey versions were returned. In addition to the bootstrapping technique used in the main paper, we perform a battery of robustness checks for potential bias, all of which confirm the findings presented here. See the supplementary appendix for more details.

While the authors believe this procedure is necessary, the results are largely consistent with those estimated using the full sample (see Table A1 in the supplementary appendix).

The percent premium was computed for farmland with heterogeneous soil types (varied soil map) and no quality information provided (soil productivity rating not reported), which have an average bid of $207.13 per acre.

One possibility is that farmers simply interpret the absence of soil productivity information as representing average soil quality. However, the results of the interacted model are not consistent with this possibility. If true, farmers only value zone soil data when soil quality is average and are unwilling to pay for it when quality falls above or below the average.