MIS Quarterly

RESEARCH ARTICLE

THE EFFECTS OF DIGITAL TRADING PLATFORMS ON COMMODITY PRICES IN AGRICULTURAL SUPPLY CHAINS

Rajiv Banker
Fox School of Business, Temple University, Philadelphia, PA 19122, U.S.A. {banker@temple.edu}

Sabyasachi Mitra
College of Management, Georgia Institute of Technology, Atlanta, GA 30308, U.S.A. {saby.mitra@mgt.gatech.edu}

V. Sambamurthy
Eli Broad College of Business, Michigan State University, East Lansing, MI 48824, U.S.A. {sambamurthy@bus.msu.edu}

Appendix

The Effects of Digital Trading Platforms on Commodity Prices in Agricultural Supply Chains

We performed two additional robustness checks for the results reported in Table 4. First, we incorporated additional dummy variables to capture buyer bean type and seller bean type fixed effects. Second, we incorporated Heckman selection models to correct for some of the potential biases that arise from the unbalanced panel data used in our analysis.

Additional Notation

BB_{bk}	Dummy variable for each bean type (k) and buyer (b) combination (k = 1,b = 1 base case)
BS_{sk}	Dummy variable for each bean type (k) and seller (s) combination (k = 1, s = 1 base case)
TV_{mk}	Overall coffee export volume during month m (kilo tons)
DF_{m}	Fraction of coffee export volume traded on the digital platform during month m
Z_{bd}, Z_{sd}	Indicator variable set to 1 if buyer b (seller s) had a digital transaction on date d

Buyer-Product and Seller-Product Dummy Variables

Note that the results in Table 4 incorporate buyer (Model A) and seller (Model B) random effect specifications. The models also incorporate dummy variables for each coffee bean type (Arabica Parchment, Arabica Cherry, Robusta Parchment, and Robusta Cherry). However, it is also possible that buyers and sellers have special expertise in certain types of coffee beans. For example, an exporter (buyer in the digital platform) may supply to international coffee roasters that have a preference for specific bean types. Likewise, a planter may have expertise

in growing specific types of coffee beans. To incorporate such expertise differences, we included dummy variables for each bean type buyer (Model A) and bean type seller (Model B) combination (BB_{bk} and BS_{sk} variables described above). Obviously, STATA automatically drops those dummy variables that do not have at least two corresponding digital platform transactions. The dummy variables do not change the results significantly and the hypotheses are supported in the modified analysis.

Unbalanced Panel Data and Selection Bias

In our dataset, participants had an average of 10 transactions during the 20-month period of the study, indicating that each participant transacted on only 2.5 percent of the trading days. It is likely that the selection of trading days on the digital platform is entirely random and based solely on whether the participant has coffee lots to buy or sell on a specific day. If the selection of digital transaction dates is indeed random and uncorrelated with the error term, the parameter estimates in Table 4 are consistent (Wooldridge 2001).

However, consider the following situation where selection of transaction dates on the digital platform may not be entirely random. Participants have a choice between trading on the physical and digital platforms. For historical reasons and past familiarity with the physical platform, they may choose to transact on the digital platform only if the price difference between the digital and physical platforms is above a certain threshold or reservation value. This is analogous to the incidental truncation problem described in Wooldridge (2001, p. 578) for estimating the wage equation using a panel of individuals where wage data is only observed during months when the individual is working (that is, the wage is above an unobserved reservation wage for the individual). In such situations, selection for inclusion in the panel may be correlated with the error term and the estimation of parameters using only observed data is likely to be biased (Wooldridge 2001).

As suggested by Wooldridge (p. 582), we utilize the two-step Heckman selection procedure to correct for the impact of this selection bias on parameter estimates. The first stage of the Heckman requires a completely balanced panel of explanatory variables that predict the existence of a transaction on a specific date. We identify two factors that affect the existence of a digital transaction on a specific date First, the volume of coffee traded on the digital platform per month varies between 1 percent and 6.5 percent of total export volume, indicating that there are months in the year when participants find it beneficial to trade on the digital platform. Second, the likelihood of a transaction is higher during times when the overall coffee trading volume is high. Thus, in the first stage of the Heckman model, we use DF_m (fraction of overall coffee export volume traded on the digital platform in month m) and TV_m (total export volume of coffee in month m) as explanatory variables to predict the existence of a digital transaction on a specific date. We constructed a completely balanced panel of the TV_m , DF_m, and Z_{bd} (Z_{sd}) variables for every buyer (seller) and for every date in the period of the study. The Heckman procedure estimates a logistic regression with Z_{bd} (Z_{sd}) as the outcome variable, and TV_m and DF_m as independent variables. In the second stage of the Heckman procedure, the Mills ratio calculated from the logistic regression in the first stage is included as an independent variable, along with the other variables in Models A and B in the paper (Wooldridge 2001). If a participant had more than one transaction on the same day, we spread the transactions within the same month to construct the daily panel data.

The results of the analysis appear in Table A1. As expected, the first stage of the Heckman analysis shows that transactions are more likely to occur during the months when the fraction of coffee traded online is high (the coefficient of the DF_m variable is positive and significant at the 1 percent level). In the second stage, the coefficient for the Mills ratio from stage 1 of the Heckman procedure is not significant indicating that selection bias does not affect the parameter estimates. The coefficients of the focal variables in the models remain almost unchanged from Table 4 and support hypotheses H1 through H4.

Table A1. Heckman Selection Models with Buyer and Seller Random Effects

Heckman Second Stage				
Model Variables	**Explanation**	**Pred Sign**	**Buyer Random Effects**	**Seller Random Effects**
Intercept	Constant term	+/-	0.11549 (0.17916)	0.11472 (0.16842)
R_g	Raw grade (R_g = 1) indicator	+	0.02837 (0.00714)***	0.03221 (0.00803)***
P_g	Premium grade (P_g = 1) indicator	-	-0.01918 (0.00485)***	-0.01696 (0.00425)***
CV_g	Coefficient of variation of price	-	-0.5219 (0.17296)***	-0.4758 (0.15072)***
S	Sell (S = 1) transaction indicator	+	0.01219 (0.00451)***	0.00718 (0.0042)*
CB	Click and Book indicator	+	0.00205 (0.00643)	-0.00245 (0.00624)
OB	Order Book Management indicator	+/-	-0.01184 (0.01284)	-0.00945 (0.01234)
ST/BT	Seller/Buyer is a trader	+	-0.00196 (0.00369)	-0.00192 (0.00381)
SE/BE	Number of seller/ buyer transactions	+/-	-0.00002 (0.00006)	-0.00007 (0.00005)
Q	Lot Size (Tons)	+/-	-0.00001 (0.00027)	0.00001 (0.00027)
V_g	Market Size for coffee grade (KTons)	+/-	0.01535 (0.01391)	0.0076 (0.01263)
M_x	Month dummy variables (x = 2 – 12)	+/-	Included	Included
B_k	Bean Type dummy variables (k = 2 – 4)	+/-	Included	Included
Mills Ratio	Mills ratio from Heckman Stage 1	+/-	-0.03362 (0.07791)	-0.03099 (0.06964)
R^2			24.5%	22.9%
Wald χ^2			177.9***	118.68***
Heckman First Stage				
Intercept	Constant Term in Heckman First stage		-2.6171 (0.06709)***	-2.68975 (0.06215)***
TV_m	Total Export Volume in month m (KTons)		0.000005 (0.000003)*	0.000004 (0.000003)
DF_m	Fraction of Coffee Traded Online in month m		10.14443 (0.89756)***	10.45842 (0.90624)***

Notes: Results from Heckman selection models for Buyer (A) and Seller (B) random effects are shown. Significance is shown at the ***(1%), **(5%), and *(10%) levels based on a two-tailed test. Standard errors are in parenthesis.