A Bayesian Analysis of Dual Trader Informativeness in Futures Markets

Sugato Chakravarty*
Purdue University
West Lafayette, IN 47907
Phone: (765) 494-6427
Fax: (765) 494-0869
Email: sugato@purdue.edu

Kai Li
University of British Columbia
Vancouver, Canada V6T 1Z2
Phone: (604) 822-8353
Fax: (604) 822-4695
Email: kai.li@commerce.ubc.ca

This Version: May 2001

* Corresponding author. We thank N.R. Prabhala, and seminar participants at Purdue University for helpful comments. Joshua Slive provided valuable editorial help. Li acknowledges financial support from the Social Sciences and Humanities Research Council of Canada. An earlier version of the paper was circulated under the title “An Analysis of Own Account Trading by Dual Traders in Futures Markets: A Bayesian Approach.” As usual, we are to blame for any remaining mistakes.
A Bayesian Analysis of Dual Trader Informativeness in Futures Markets

Abstract

We employ a conditional event study methodology to determine if dual traders in futures markets are informed traders. Underpinning this approach is the assumption that a dual trader’s decision to trade on her own account is not random, but is endogenously determined by her expectations of profit related to this decision. Therefore, the appropriate empirical framework to analyze the issue is a simultaneous equations model where the two endogenous variables are (1) the binary decision of own account trading (or not), and (2) the trading profit resulting from her own account trading. Our test of dual trader informativeness lies in estimating the correlation between the error terms of the two equations in our model, where one error term proxies for a dual trader’s unobserved private information and the other captures her abnormal profit. Upon estimating the model using the Bayesian approach, we find no correlation between the proxy for private information and a dual trader’s abnormal profit. This result is robust and does not appear to be driven by data, model specifications and/or estimation techniques. We conclude that dual traders seem not to possess any private information, in spite of their personal trades being positively correlated with profits. We also find that dual traders are heterogeneous in terms of trade-related characteristics.

Keywords: endogeneity, heterogeneity, informed trader, private information, Markov chain Monte Carlo, simultaneous equations

JEL Classification: G20, G28, C11, C15, C35
1. Introduction

Beginning with the seminal work of Kyle (1985), Glosten and Milgrom (1985), and Easley and O’Hara (1987), the examination of how informed traders can impact asset prices and market liquidity has been central to the theoretical literature on market microstructure. However, the implications emerging from this body of research are difficult to test, due to the simple fact that informed traders are unobservable and the empirical literature has had to resort to inferring informed trading through indirect means such as through trade sizes (Barclay and Warner (1993)), through trade sizes and trader types (Chakravarty (2001)), and through the timing of trades (Lee, Mucklow and Ready (1993)).

In the current paper, we adopt a new technique to shed light on the informativeness of an important class of traders in the futures markets: dual traders. Dual trading is an age-old custom in futures markets whereby some floor traders are allowed to trade both for themselves and for their customers.1 Over the past ten years a debate has raged regarding whether dual traders should enjoy such a privilege (Chakravarty and Li (2001)). Central to the debate is the issue of whether dual traders are informed. Unfortunately, the extant theoretical and empirical literature provides relatively little assistance on this matter. Most of the theoretical literature on dual trading, for example, has formalized the intuition of dual traders piggybacking on the information inherent in customer trades for personal profit (Grossman (1989), Roell (1990), Fishman and Longstaff (1992), Chakravarty (1994), and Sarkar (1995)). The empirical literature

---

1 Dual trading, however, is not just restricted to the futures markets. In equity markets, for example, the market makers or specialists are also dual traders in the sense that they can execute a customer order by matching a buyer with a seller while taking a residual portion of the order on personal account.
on dual trading is large and diverse but is not very helpful on this point either.2

Thus, while the theoretical models assume dual trader informativeness and use it as a springboard to study issues related to market liquidity and trading profitability, we take a step back and ask the more fundamental question of whether these dual traders are informed to begin with. To do so, we employ the conditional event study methodology of Prabhala (1997) and Li (1998), which does not assume that the personal trading decision of a dual trader is a random event. Rather, it assumes that the decision to trade on own account by a dual trader is based on her expectations of the abnormal profit she makes through the transaction. Thus, the own account trading decision of a dual trader is endogenously determined with respect to her trading profit. The conditional event study explicitly accounts for this endogeneity.

Specifically, we jointly examine a dual trader's own account trading decision and her profitability by employing a simultaneous equations model with a binary endogenous variable—the decision of own account trading—and a trading profit variable. In the own account trading equation, the error term captures the dual trader's (unobservable) private information. In the profit equation, the error term captures her abnormal profit. A significant correlation between the error term in the own account trading decision equation and the error term in the profit equation would indicate that the dual trader possesses (unobservable) private information that leads to abnormal profit. Note that an important feature of our modeling

---
2 The empirical literature on dual trading can be broadly classified into three themes. The first focuses on the liquidity effects of various dual trading restrictions imposed on the futures markets (Smith and Whaley (1994), Chang, Locke and Mann (1994), Chang and Locke (1996), and Locke, Sarkar and Wu (1999)). The second theme of the empirical literature examines the microstructure of futures markets under competitive market making (Manaster and Mann (1996), and Ferguson and Mann (2000)). The third theme, represented by Chakravarty and Li (2001), examines the timing and determinants of own account trading by dual traders in futures markets.
framework is that we isolate the abnormal trading profit associated with a dual trader’s personal trades from her overall trading profit and correlate the abnormal profit with the unobserved private information (if any) of the dual trader. This is, arguably, a more appropriate test of dual trader informativeness in contrast to the extant literature which has inferred dual trader informativeness from the overall trading profits associated with their personal trades (Fishman and Longstaff (1992)).

We estimate our system of equations using Bayesian techniques known as the Gibbs sampler and data augmentation. The Bayesian approach allows us to incorporate both parameter uncertainty and model uncertainty in a consistent manner, and, provides us with more accurate parameter estimates than Heckman’s (1976, 1979) two-step estimation technique. Moreover, the Bayesian estimates are likelihood-based with nice finite sample properties.

Using audit trail transaction data compiled by the Commodity Futures Trading Commission (CFTC), we estimate our conditional event study model on each of the 101 most active dual traders in the data. We find no support for dual trader informativeness—i.e., there is no significant correlation between a dual trader’s abnormal trading profit and her unobservable private information. Our results compliment the findings in Chakravarty and Li (2001) who conclude that the own account trading decision of a dual trader is solely driven by

---

3 Gibbs sampling is a simulation tool for obtaining marginal distributions from a non-normalized joint density (Casella and George (1992), and Gelfand and Smith (1990)). Data augmentation is a scheme to augment the observed data in order to simplify the likelihood/posterior (Tanner and Wong (1987)). Both techniques are special cases of the Markov chain Monte Carlo approach (MCMC, see Chib and Greenberg (1996) for a survey). Two recent market microstructure papers adopting the Bayesian approach are Hasbrouck (1999) and Ball and Chordia (1999).

4 The data provide detailed information about trade time, price, quantity, trade direction (buyer or seller), the contract and the trader’s identification and have been used within the CFTC for regulation and/or enforcement purposes.
her inventory control. We also find strong evidence to suggest that dual traders are distinct (heterogeneous) in their trade-related characteristics. Overall, our results indicate that dual traders are uninformed, which has important policy implications on the suggested restrictions on their personal trading that Congress is currently contemplating.

Our paper is related to Chakravarty and Li (2001) who examine the determinants of dual traders’ personal trades. The main result of that paper is that most dual traders trade on own account solely due to inventory rebalancing reasons. In contrast, the current paper focuses on the informativeness of dual traders through estimation of the statistical correlation between the private information of dual traders and their abnormal trading profits.

The plan for the rest of the paper is as follows. Section 2 describes the empirical model used to examine dual traders’ own account trading and provides relevant details on Bayesian estimation. Section 3 introduces the data and defines the explanatory variables. Section 4 reports our findings. We conclude in Section 5.

2. Methodology

2.1 The Model

Our conditional event study models a dual trader’s decision to trade on her own account and the effect of this decision on her profit in two separate equations. The (own account) trading decision is discrete and is modeled using a probit regression. The endogeneity of the own account trading decision (if any) is reflected in the model by the explicit inclusion of a covariance term between the error of the probit equation, interpreted as the dual trader’s unobservable private information, and the error of the profit equation, interpreted as the
abnormal profit. If this covariance is positive, we infer that the dual trader possesses significant private information that leads to abnormal trading profit.

Consider $I_{i,t}^{*}$ to be the unobservable latent variable representing the added utility of dual trader $i$ when she chooses own account trading over trading on behalf of her customers at time $t$. Let $I_{i,t}^{*} = E(I_{i,t}^{*}) + e_{i,t;it}$. Here we assume that $I_{i,t}^{*}$ is normally distributed with mean $E(I_{i,t}^{*})$ representing the market’s expectations of dual trader $i$’s utility increase through own account trading. For tractability, we assume a linear structure for the market’s expectations. This implies that $E(I_{i,t}^{*}) = X_{i,1t} \beta_{1,i}$, where $X_{i,1t}$ is an $n \times k_1$ matrix of observable trader and market characteristics and $\beta_{1,i}$ is a vector of $k_1$ parameters. In our paper, the information of dual trader $i$ consists of an unobservable component, her private information which is captured by the error term $e_{i,t;it}$, and a second component that can be proxied for by her customer trading in the recent past. We use a variable in $X_{i,1t}$ to capture this short-term order-flow based information component. Thus, we characterize dual trader $i$’s decision on own account trading over a 5-minute bracket $t$ in a dual trading day as

$$I_{i,t}^{*} = X_{i,1t} \beta_{1,i} + e_{i,t;it}$$  \hspace{1cm} (1)

where $I_{i,t} = 1$, if $I_{i,t}^{*} > 0$, that is, dual trader $i$ chooses to execute some own account trades during time interval $t$, and $I_{i,t} = 0$, if $I_{i,t}^{*} \leq 0$, that is, dual trader $i$ chooses to execute trades on behalf of her customer during time interval $t$.

In our second equation we examine the effect of dual trader $i$’s own account trading decision on her personal trading profit,

$$\Pi_{i,t} = I_{i,t} Y_{i,t} + X_{2,1t} \beta_{2,i} + e_{2,1t},$$  \hspace{1cm} (2)

where $I_{i,t}$ is the binary choice variable on own account trading by dual trader $i$ over time interval
$t$, $X_{2,it}$ is an $n \times k_2$ matrix of observable trader and market characteristics, $\beta_{2,i}$ is a vector of $k_2$ parameters and $e_{2,it}$ is the error term capturing the abnormal profit in time bracket $t$. $\Pi_{i,t}$ is the trading profit for dual trader $i$, up to and including time bracket $t$, computed as in Fishman and Longstaff (1992). Specifically, the trading profit of dual trader $i$ in time bracket $t$ on day $d$ is obtained as

$$\Pi_{i,d} = \text{Buy Volume}_{i,d} \times (\text{Settlement Price}_{d} - \text{Purchase Price}_{i,d})$$

$$+ \text{Sell Volume}_{i,d} \times (\text{Sale Price}_{i,d} - \text{Settlement Price}_{d}) .$$  \hspace{1cm} (3)

For the trading profit up to and including time bracket $t$, $\Pi_{i,t}$, we simply cumulate $\Pi_{i,d}$ from the beginning of a trading day $d$ to time bracket $t$.

In summary, equation (1) models the own account trading decision of a dual trader as a function of her unobservable private information, the customer order-flow based information, and relevant exogenous variables to be discussed in Section 3.2. Equation (2) models the trading profit of the same dual trader as a function of her own account trading choice variable ($I_{i,t}$) and relevant exogenous variables (also discussed in Section 3.2). One purpose of equation (2) is to determine whether dual traders make profits from own account trading. This can be analyzed through the sign and statistical significance of the coefficient associated with the binary choice variable $I_{i,t}$.

Under the single equation approach in which we assume no simultaneity, the error terms in equations (1) and (2) follow an independent and identical univariate normal distribution. Under the simultaneous equations approach, the error terms in equations (1) and (2) are postulated to have the following distributional characteristic. Specifically, \( \begin{pmatrix} e_{1,it} \\ e_{2,it} \end{pmatrix} \) follows
an independent and identical BVN\((0, \Sigma)\), where BVN denotes a bivariate normal distribution with mean zero and the variance-covariance matrix \(\Sigma = \begin{pmatrix} 1 & \sigma_{e_1e_2} \\ \sigma_{e_1e_2} & \sigma_{e_2e_2} \end{pmatrix}\). Note that in \(\Sigma\), \(\text{Var}(e_{1,i}) = 1\) because \(I_{i,t}\) is only observed as a binary variable.

A priori, the single equation model and the simultaneous equations model are both of interest. Statistically, the difference between the two approaches is that the former model sets \(\sigma_{e_1e_2} = 0\) while the latter leaves the covariance term \(\sigma_{e_1e_2}\) unconstrained. Economically, the key issue is whether dual trader \(i\) possesses unobservable private information that is systematically related to her abnormal trading profit, after controlling for observables such as inventory effects, liquidity, trading skills, etc. In other words, the question is whether dual trader \(i\)'s own account trading decision is exogenous to her personal trading profit.

To test for \(H_0: \sigma_{e_1e_2} = 0\) versus \(H_1: \sigma_{e_1e_2} \neq 0\), we compute the Bayes factor \((BF_{01})\) between the two models. The Bayes factor is the Bayesian version of the likelihood ratio test, which is obtained as the ratio of data densities under the model with zero covariance \((H_0)\) and under the model with non-zero covariance \((H_1)\), respectively.\(^5\) Noting that \(H_1\) nests \(H_0\), we employ the Savage-Dickey density ratio of Verdinelli and Wasserman (1995) to simplify the computation of the Bayes factor given by

\[
BF_{01} = \frac{f(\sigma_{e_1e_2} = 0 \mid y)}{f(\sigma_{e_1e_2} = 0)},
\]

where \(f(\sigma_{e_1e_2} \mid y) = \int \int f(\delta, \sigma_{e_1e_2}, \sigma_{e_2e_2} \mid y) d\sigma_{e_2e_2} d\delta\), \(f(\sigma_{e_1e_2}) = \int \int f(\delta, \sigma_{e_1e_2}, \sigma_{e_2e_2}) d\sigma_{e_2e_2} d\delta\), \(\delta = (\beta_1', \gamma, \beta_2')\), and \(y\) represents the data.

\(^5\) See Kass and Raftery (1995) for a survey on Bayes factors.
According to Kass and Raftery (1995), there exists decisive evidence from the data against $H_1$ when $BF_{01}$ exceeds 100. In practice, unless the data evidence overwhelmingly supports one particular formulation, for inference purposes we can average out model uncertainty by pooling posterior densities under $H_0$ and $H_1$ respectively, as suggested by Poirier (1995, pp. 604-605). Specifically, from the definition of the Bayes factor $BF_{01}$, the posterior probability that the single equation model holds true equals $\frac{BF_{01}}{1 + BF_{01}}$, and the posterior probability that the simultaneous equations model holds true equals $\frac{1}{1 + BF_{01}}$. Then the pooled posterior point estimate of any parameter is obtained as the weighted average of the corresponding posterior point estimates under the single equation model and the simultaneous equations model, using the two weights above.

2.2 Estimation Using the Markov Chain Monte Carlo (MCMC) Method

Under our simultaneous equations framework there is a nontrivial covariance structure ($H_1: \sigma_{el_e} \neq 0$) between the error terms in equations (1) and (2). Due to the non-linearity in the likelihood function (caused by the binary own account trading choice variable $I$ and the covariance $\sigma_{el_e}$), full information maximum likelihood estimation is generally avoided in favor of the less efficient but computationally simpler estimation procedures such as the two-step algorithm developed by Heckman (1976, 1979). We discuss the Heckman approach in Section 2.3. In the current paper, we follow the method developed in Li (1998) to conduct a finite sample likelihood-based analysis of our empirical model in equations (1) and (2), using a combination of Gibbs sampling and data augmentation.

In a standard 2 $\times$ 2 variance-covariance matrix $\Sigma$ with four elements, there are three
unique elements that need to be estimated, as the two off-diagonal elements are identical.  In our case, since equation (1) is a probit, unity is imposed on the first diagonal element for identification, leaving only two free parameters $\sigma_{e_{1e2}}, \sigma_{e_{2e2}}$ (the off-diagonal element and the second diagonal element) to be estimated.  This creates complications in the estimation procedure and requires us to reparameterize $\Sigma$ and to estimate the two free parameters separately. 

Decomposing the joint bivariate normal distribution for $\begin{pmatrix} e_{1,t} \\ e_{2,t} \end{pmatrix}$ in equations (1) and (2) into the product of the marginal distribution for $e_{1,t}$ and the conditional distribution $e_{2,t} | e_{1,t}$, we obtain

\begin{align}
I_{i,t}^* &= X_{i,t}\beta_{i,t} + e_{1,t},
\end{align}

\begin{align}
\Pi_{i,t} &= I_{i,t}\gamma_{i} + X_{2,it}\beta_{i2} + e_{1,t}\sigma_{e_{1e2}} + v_{i,t},
\end{align}

where $e_{1,t} = I_{i,t} - X_{1,typi1}, \sigma^2 = \sigma_{e_{2e2}} - \sigma_{e_{1e2}}^2$, and $v_{i,t} \sim N(0,\sigma^2), e_{1,t} \sim N(0,1)$ are independent. Conditional on the data and the regression parameter $\delta = (\beta_1', \gamma, \beta_2')'$ ($e_{1,t}, e_{2,t}$ are given), drawing $\sigma_{e_{1e2}}, \sigma^2$ is like drawing from the posterior distribution of the univariate regression of $e_{2,t}$ on $e_{1,t}$,

\begin{align}
e_{2,t} &= e_{1,t}\sigma_{e_{1e2}} + v_{i,t},
\end{align}

$ v_{i,t} \sim N(0,\sigma^2)$.

---

6 The complication arises because when we have three free parameters to estimate in a 2x2 variance-covariance matrix $\Sigma$, from Zellner (1971, pp. 224-227) we know that, conditional on the data and the regression parameter vector, $\delta$, the inverse of the variance-covariance matrix $\Sigma^{-1}$ (with its three free parameters), follows a Wishart distribution; while conditional on the data and $\Sigma$, $\delta$ follows a multivariate normal distribution. But in our current setting, due to the probit equation (1), the variance of the error term in equation (1) is fixed at 1, which reduces the number of free parameters from three to two and precludes us from using the standard results above.
From here on, we focus on the reparameterized variance-covariance matrix
\[
\Sigma = \begin{pmatrix} 1 & \sigma_{e1e2} \\ \sigma_{e1e2} & \sigma_{e1e2}^2 + \sigma^2 \end{pmatrix}.
\]

We assume the following prior distribution
\[
f(\delta, \sigma_{e1e2}, \sigma^{-2}) \propto f(\delta) \cdot f(\sigma_{e1e2}) \cdot f(\sigma^{-2}),
\]
where
\[
\begin{align*}
  f(\delta) & \sim MVN(\delta_0, \Psi_0^{-1}), \\
  f(\sigma_{e1e2}) & \sim N(r_0, h_0^{-1}), \\
  f(\sigma^{-2}) & \sim G\left(\frac{v_0}{2}, \frac{c_0}{2}\right),
\end{align*}
\]
and MVN denotes a multivariate normal distribution, N denotes a univariate normal distribution, and G denotes a Gamma distribution (Poirier, 1995, pp. 98-99).

Throughout the paper, we choose the following prior to report our final estimation results,
\[
\delta_0 = 0, \Psi_0^{-1} = 10^8 \cdot I_p, r_0 = 0, h_0 = 2, v_0 = 4, c_0 = 1,
\]
where \(p (= k_1 + 1 + k_2)\) is the dimension of the regression parameter \(\delta\), and \(I_p\) denotes an identity matrix of rank \(p\). The set of priors chosen has a fairly flat distribution on \(\delta\) centered at a vector of zeros, and the prior mean for the variance-covariance matrix, \(\Sigma\), is an identity matrix.

The Bayesian estimation proceeds as follows. First, we augment the observed data \(I\) with the unobservable data (i.e., the incremental utility of dual trader \(i\) associated with her own account trading decision). This implies generating the latent incremental utility variable \(I^*_i\), based on our observation of dual trader \(i\)'s own account trading decision \(I\). When the augmented data are generated consistently within the structure of the model, the distribution of the augmented data converges asymptotically to the distribution of the observed data. We then use the likelihood of both the observed data and the augmented data as a proxy for the
likelihood of the observed data. Conditional on the observed and augmented data, approximate posteriors for the model parameters may be obtained using standard simulation methods. Next, we use the Gibbs sampler to integrate out the uncertainty introduced by the involvement of unobserved data, thereby obtaining posteriors conditional only on the observed data (the actual choice of own account trading made by dual trader $i$). We then iterate between the data augmentation and the Gibbs sampler steps. Our Bayesian estimates are obtained as sample averages of these Gibbs draws. The operations discussed above are collectively referred to as the Markov chain Monte Carlo (MCMC) method.

### 2.3 The Heckman Two-Step Estimation Method

Our simultaneous equations framework in equations (1) and (2) is well known in the econometrics literature on limited dependent variables (Maddala (1983)). A slightly different model specification that does not include the endogenous dummy variable in the second equation has been extensively applied in conditional event studies in finance (see Prabhala (1997) for a survey). Prabhala argues that when the endogenous event dummy variable is not included in the announcement effect equation (i.e., a different version of our equation (2)), consistent estimation may be achieved through a simple two-step procedure (Heckman (1976, 1979)). Below, we show why this approach is inappropriate in estimating our model, given by equations (1) and (2).

Following the bivariate normal assumption, the conditional mean of the error term $e_{2,it}$ can be shown in Heckman (1976) as

$$E(e_{2,it} \mid I_{i,t} = 1, X_{2,it}) = E(e_{2,it} \mid e_{1,it} > -X_{1,it} \beta_{1,j}, X_{2,it}) = \sigma \phi \frac{\phi(X_{1,it} \beta_{1,j})}{\Phi(X_{1,it} \beta_{1,j})}. \quad (9)$$
\[
E(e_{2,it} \mid I_{i,t} = 0, X_{2,it}) = E(e_{2,it} \mid e_{1,it} \leq -X_{1,it}^T \beta_{1,i}, X_{2,it}) = -\sigma_{e_{1,t}} \frac{\phi(X_{1,it}^T \hat{\beta}_{1,i})}{1 - \Phi(X_{1,it}^T \hat{\beta}_{1,i})},
\]

(10)

where \(\phi(X_{1,it}^T \beta_{1,i})\) and \(\Phi(X_{1,it}^T \beta_{1,i})\) are, respectively, the standard normal density function and standard normal cumulative distribution function evaluated at \(X_{1,it}^T \beta_{1,i}\). Based on the above observation, Heckman (1976, 1979) develops the so-called two-step estimation method. In the first step, the probit model in equation (1) is estimated by the maximum likelihood method. This step provides a consistent estimate \(\hat{\beta}_{1,i}\) of \(\beta_{1,i}\). Then \(\hat{\beta}_{1,i}\) is used to obtain estimates of \(\hat{\beta}_{2,i}\) and \(\hat{\sigma}_{e_{1,t}}\).

The Heckman two-step approach provides a convenient way of obtaining consistent point estimates, in most cases of simultaneous equations models with limited dependent variables, but it is inappropriate in our particular model formulation. Specifically, our simultaneous equations model of equations (1) and (2) is different because the second equation in the system also contains the endogenous dummy variable \(I_{i,t}\). Note that in equation (11), the dummy variable \(I_{i,t}\) is a function of \(X_{1,it}^T \hat{\beta}_{1,i}\) as well as the two added regressors \(\frac{\phi(X_{1,it}^T \hat{\beta}_{1,i})}{\Phi(X_{1,it}^T \hat{\beta}_{1,i})}\) and \(\frac{\phi(X_{1,it}^T \hat{\beta}_{1,i})}{1 - \Phi(X_{1,it}^T \hat{\beta}_{1,i})}\), which causes multicollinearity among regressors. Given that the data are not able to distinguish between the dummy variable and the two added regressors, the estimates of...
\(\gamma_i\) and \(\sigma_{e,2}\) would have large standard errors that would make them unreliable.

In contrast, using MCMC, we do not need to introduce any additional regressors in equation (2), and our estimate of the covariance term \(\sigma_{e,2}\) is obtained as part of the variance-covariance matrix. By construction, the Bayesian approach does not suffer from multicollinearity.

3. Data and Explanatory Variables

3.1 The Data

Our data consist of audit trail transaction records of the 101 most active dual traders in eight futures contracts traded at the Chicago Mercantile Exchange (CME) during the first six months of 1992 and is the same as in Chakravarty and Li (2001). Overall, there are over two million records that provide a detailed look at the complete trading history of all floor traders in eight different futures pits. We supplement the above data with the daily settlement price data for each of the contracts over the sample period in order to calculate the traders’ personal trading profits.

Our definitions of a dual trading day, dual traders, locals and brokers follow Locke, Sarkar and Wu (1999), and Chakravarty and Li (2001). Unique to this data, each record also specifies the trade direction and a classification of the customer types for each side of a trade. There are four customer type indicators (CTI), labeled 1 through 4. The CTI 1 trades are market making trades for personal accounts (39% of the volume); CTI 2 trades are trades executed for

---

7 Locals are floor traders trading for their own accounts and are viewed as important suppliers of liquidity in the futures markets (see also Working (1967), Silber (1984), and Smidt (1985)).
the account of the trader's clearing member (6.2% of the volume); CTI 3 trades are trades executed for the account of any other exchange member (5.7% of the volume); and CTI 4 trades are the trades of outside customers (49.1% of the volume). Other details of the data are provided in Chakravarty and Li (2001).

3.2 The Explanatory Variables

The dependent variable which captures a dual trader's own account trading decision in equation (1) is \textit{TRADE DUMMY}, \( I \), where \( I \) equals 1 if the dual trader trades on own account in time bracket \( t \), and 0 otherwise. The dependent variable which captures a dual trader's own account trading profit in equation (2) is \textit{PROFIT}, \( \Pi \), computed as her cumulative own account trading profit from the beginning of a trading day up to time bracket \( t \) within a day.

Our choice of a parsimonious set of exogenous variables, determining the trading decision of dual traders and their personal trading profit, is based on the existing literature.\(^8\) The set of explanatory variables can be broadly classified into variables capturing market liquidity, information, trading momentum, contract risk, inventory effects, trading skills and timing of trades. These are

- \textit{LagNLOCAL} — the number of pure locals (sole own account traders) over the prior 5-minute bracket. This is a proxy for market liquidity.

- \textit{LagFRACTI4} — dual trader \( i \)'s customer trading (i.e., brokering) volume as a fraction of her total trading volume in the 5-minute bracket prior to the current time bracket \( t \). This proxies for a dual trader's transient, customer order-flow based information.

---

- **LagVOLCTI1** — dual trader $i$'s own account trading volume in the prior 5-minute time bracket. This proxies for a momentum effect.

- **LagVOLATILITY** — the maximum of the buy-price and sell-price standard deviations in the previous 5-minute time bracket. This is a proxy for contract risk.

- **LagINVENTORY** — for dual trader $i$, computed as her CTI 1 buy trades minus CTI 1 sell trades, cumulated from the beginning of a trading day to time bracket $t-1$, assuming that dual trader $i$ only has full control of her CTI 1 trades.

- **SKILL** — for dual trader $i$, computed as the difference between her volume-weighted mean buy price (sell price) and the volume-weighted mean buy price (sell price) for all other floor traders during each 5-minute bracket.

- **HOT** — a trade timing dummy, which equals 1 if the 5-minute bracket in which dual trader $i$ trades on her own account (CTI 1 trades only) belongs to the first two hours or the final half-hour of a trading day, and 0 otherwise.

In summary, the variables explaining a dual trader's own account trading decision, TRADE DUMMY, in equation (1), are lagNLOCAL, lagFRACTI4, lagVOLCTI1, lagVOLATILITY, lagINVENTORY, lagSKILL, and HOT. The corresponding variables explaining the dual trader's own account trading profit, PROFIT, in equation (2), are TRADE DUMMY, VOLATILITY, SKILL, and HOT. By construction, PROFIT has a high serial correlation. We therefore include a lagged profit variable, lagPROFIT, on the right hand side of equation (2) to ameliorate the problem.

Equations (1) and (2) are first estimated as two independent equations (the naïve model) and then simultaneously, recognizing the possible correlation in the error terms. The model
comparison results (as captured by the Bayes factor, $BF_{01}$, in Table 3.3) indicate that neither model formulation is predominantly favored by the data. To take into account model uncertainty, the final estimates are obtained by pooling the estimates from the two models, with the pooling weights obtained endogenously within the estimation process (see Section 2.1). We repeat this exercise for each of the 101 active dual traders in our sample.

4. Dual Trader Informativeness

There are 101 sets of coefficient estimates, one for each dual trader in our sample. We first present, in Tables 1 and 2, the fractions of positive and negative coefficients, as well as the fractions of positive and statistically significant (at the 5% level) and negative and statistically significant (at the 5% level) coefficient estimates in equations (1) and (2), respectively.

From Table 1, we find lagFRACTI4 is negative and significant while lagINVENTORY is overwhelmingly positive and significant in all eight contracts. The remaining coefficients in equation (1) display varying degrees of significance, both within the same contract and across the different contracts. The important variable in equation (2) is TRADE DUMMY. From Table 2, we find that the corresponding coefficient is overwhelmingly positive in all contracts. It appears that dual traders in general profit from the act of own account trading. Whether or not these trades (and resultant profits) are private information related, remains to be seen.

The results for the median dual trader in each of the eight contracts are provided in Tables 3.1, 3.2 and 3.3, to give a sense of the magnitude of individual coefficients for a representative dual trader in each of the eight contracts. The median dual trader in each contract is the trader whose number of dual trading days is the median of the dual trading days
of all selected dual traders in that contract. The estimation results of all other dual traders are
omitted for brevity, but are available from us on request. Specifically, Table 3.1 reports the
single equation (naïve model) estimates, with the corresponding posterior standard deviations
in parentheses, of equations (1) and (2), while Table 3.2 reports the simultaneous equations
estimates.\(^9\) Table 3.3 reports the pooled estimates (of the naïve and simultaneous equations
models), where the pooling weights are derived from the Bayes factor, \(BF_{01}\), reported in panel C
of Table 3.3.

An examination of Tables 3.1, 3.2 and 3.3 reveals that the final coefficient estimates
obtained from pooling (Table 3.3) are more precise than the corresponding estimates in either of
Tables 3.1 and 3.2. This is intuitive because by averaging across the two contending models, we
take into account model uncertainty and the resulting pooled estimates have smaller posterior
standard deviations.

Interestingly, the Bayes factors in panel C of Table 3.3 indicate that for each dual trader,
the values hover around 1, far below the 100 necessary to reject the null of no covariance. In
results not reported for brevity, we verify that this is, in fact, true for all dual traders in our
sample. Thus, while dual traders make profits through their own account trades, these profits
are not the result of their unobservable private information.

It should be noted that Fishman and Longstaff (1992) show that, for the floor traders in
their soybean futures contract, the average profit on dual trading days is significantly higher
than that on own account trading days. This result has been used by some as evidence that dual

\(^9\) To investigate the goodness-of-fit of our simultaneous equations model, we compute the Bayes factors
comparing the simultaneous equations model with a model that only contains an intercept term for dual
traders in all eight contracts. The resulting Bayes factors are well over 100 (not reported), indicating that
our current model formulation provides a good fit of the data.
traders are informed traders. While we too find positive correlation between overall dual trader profits and their own account trading decision, there appears to be no correlation between abnormal trading profits and private information of a dual trader. This, we argue, is a more appropriate test of dual trader informativeness (or lack thereof).\textsuperscript{10}

We also estimate our simultaneous equations model using the Heckman two-step method (not reported, but available on request). The estimation results for the probit model in equation (1) are almost identical across different estimation methods, which is not surprising given that our Bayesian analyses do not indicate strong simultaneity between equations (1) and (2). The Heckman two-step method, however, gives us very different results for equation (2). Most noticeably, the coefficients associated with TRADE DUMMY ($l$) and the added regressors (see discussion in Section 2.3) tend to have much larger standard errors than those of the corresponding Bayesian estimates, and the values of these two coefficients ($\gamma$, and $\sigma_{e_1c_2}$) also tend to differ in sign, which are typical symptoms of multicollinearity. This supports our argument made in Section 2.3 about the inappropriateness of using Heckman’s two-step method to estimate our simultaneous equations model.

In Figures 1 through 8, we plot the posterior distributions of the parameters corresponding to the seven explanatory variables in equation (1) (excluding the intercept term) and TRADE DUMMY in equation (2), for the median dual trader in each contract. As before, the number at the end of each contract in each graph identifies the specific median dual trader in that contract whose posterior estimates are provided in Tables 3.1 to 3.3. From these figures,

\textsuperscript{10} We also conduct various sensitivity analyses to ensure that the peculiarities of sample selection and/or the estimation procedure do not drive our results. In all cases, our main conclusions appear robust to the various sample selection rules and prior specifications.
it is clear that the posterior distributions of the parameters are well dispersed without significant overlaps, indicating distinct posterior means and dispersions across the median dual traders. Overall, these graphs suggest heterogeneity across the dual traders.

To provide further support for dual trader heterogeneity, we conduct statistical tests on the difference in the location of the parameters across the median dual traders. Assuming that the posterior distributions of the parameters follow independent normal distributions, the standard procedure for comparing the means of two normal distributions is the two-sample t-test. If, however, the normality assumption is considered too strong, we can compare the medians of two posterior distributions using the nonparametric Wilcoxon test, and the medians of multiple posterior distributions using the nonparametric Kruskal-Wallis test.

Table 4 reports the test results. For the pair-wise comparison, the null hypothesis is that there is no difference in the means (t-test) or in the medians (Wilcoxon) of the posterior distributions of the parameters. There are eight median traders, one from each of the eight contracts, and thus twenty-eight unique pairs. We present the fraction of the corresponding test statistic with a p-value below 0.01 in Table 4. For the simultaneous comparison, the null hypothesis is that there is no difference in the medians of the posterior distributions of the parameters across all eight dual traders (Kruskal-Wallis). We present the p-value associated with the test statistic. As Table 4 indicates, regardless of the test employed, each parameter is distinctly different across the eight median dual traders, and thus provides further support of dual trader heterogeneity.

In summary, while there is evidence to support a direct and positive connection between own account trading by dual traders and their profits, we find no evidence to suggest that such
profits occur because of any (unobserved) private information that these traders may possess. Dual traders also appear heterogeneous in their observed characteristics.

5. Conclusions

Using detailed audit trail transaction data compiled by the CFTC and a conditional event study methodology, we investigate if dual traders are informed traders. Our study goes significantly beyond existing research. We recognize and account for the potential endogeneity between the own account trading decision of a dual trader and her trading profit. Our methodology also allows us to isolate the abnormal trading profits associated with dual traders’ personal trades and to compute their correlation with the unobserved private information (if any) of the dual traders. This, we argue, is a more appropriate test of dual trader informativeness, in contrast to the extant literature that has inferred dual trader informativeness from total trading profits associated with their personal trades.

The estimation of our model is performed for each dual trader in our sample, using the MCMC method. Our parameter estimates are obtained as weighted averages of the corresponding parameter estimates of two models: one of which recognizes the correlation while the other one (the naïve approach) does not. The weights themselves are determined within the estimation process.

We find that dual traders do not appear to possess any significant private information. We also uncover strong evidence that dual traders are heterogeneous in terms of trade-related characteristics. Overall, our results have important policy implications since they raise doubts about the notion that dual traders are informed traders.
References


Chakravarty, Sugato, and Kai Li, 2001, An examination of own account trading by dual traders in futures markets, working paper, Purdue University.


Ferguson, Mike, and Steven C. Mann, 2000, Execution costs and their intraday variation in futures markets, forthcoming in Journal of Business.


Heckman, James, 1979, The sample selection bias as a specification error, *Econometrica* 47, 153-162.

Heckman, James, 1976, The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models, *Annals of Economic and Social Measurement* 5, 475-492.


Table 1: Bayesian Estimates of the Dual Trader's Own Account Trading Equation

This table provides an overview of the Bayesian estimates of equation (1) across our sample of 101 dual traders in eight futures contracts. For each futures contract, the first row gives the fractions of positive (+ve) and negative (-ve) coefficient estimates; the second row gives the fractions of significantly positive and negative coefficient estimates, at the 5% significant level. The estimation results are obtained by pooling the single equation model and the simultaneous equations model. The dependent variable in equation (1) is \( \text{TRADE DUMMY} \) (\( I_t \)), which equals 1 if dual trader \( i \) trades on her own account in time bracket \( t \), 0 otherwise.

\[ L_{\text{NLOCAL}} \] is the number of pure locals (sole own account traders) in the 5-minute bracket prior to the current time bracket \( t \).

\[ L_{\text{FRACTI4}} \] is dual trader \( i \)'s customer-trading volume as a fraction of her total trading volume in time bracket \( t-1 \).

\[ L_{\text{VOLCTI1}} \] is dual trader \( i \)'s own account trading volume in time bracket \( t-1 \).

\[ L_{\text{VOLATILITY}} \], in time bracket \( t-1 \), is obtained as the maximum of the buy-price and sell-price standard deviations.

\[ L_{\text{INVENTORY}} \], is computed as dual trader \( i \)'s CTI 1 buy trades minus CTI 1 sell trades, cumulated from the beginning of a trading day to time bracket \( t-1 \).

\[ L_{\text{SKILL}} \] is a proxy to capture dual trader \( i \)'s trading skill up to time bracket \( t-1 \).

\[ \text{HOT} \] is a trade timing dummy that equals 1 if the 5-minute bracket in which dual trader \( i \) trades on own account belongs to the first two and final half-hour trading periods of a trading day, and 0 otherwise.

<table>
<thead>
<tr>
<th></th>
<th>Pit</th>
<th>Live</th>
<th>Hogs</th>
<th>Pork Bellies</th>
<th>Lumber</th>
<th>Canadian Dollar</th>
<th>S&amp;P 400</th>
</tr>
</thead>
<tbody>
<tr>
<td>% positive (Intercept)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% positive (LagNLOCAL)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% positive (LagFRACTI4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% positive (LagVOLCTI1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% positive (LagVOLATILITY)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% positive (LagINVENTORY)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% positive (LagSKILL)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% positive (HOT)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% positive (Intercept)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% positive (LagNLOCAL)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% positive (LagFRACTI4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% positive (LagVOLCTI1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% positive (LagVOLATILITY)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% positive (LagINVENTORY)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% positive (LagSKILL)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% positive (HOT)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2: Bayesian Estimates of the Dual Trader’s Profit Equation

This table provides an overview of the Bayesian estimates of equation (2) across our sample of 101 dual traders in eight futures contracts. For each futures contract, the table gives the fractions of positive (+ve) and negative (-ve) coefficient estimates; the second row gives the fractions of significantly positive and negative coefficient estimates, at the 5% significant level. The estimation results are obtained by pooling the single equation model and the simultaneous equations model. The dependent variable in equation (2) is \( \text{PROFIT} \), which is computed by cumulating dual trader \( i \)'s personal trading profit from the beginning of a trading day up to time bracket \( t \). The \( \text{TRADE DUMMY} \) equals 1 if dual trader \( i \) trades on her own account in time bracket \( t \), 0 otherwise. \( \text{VOLATILITY} \) is obtained as the maximum of the buy-price and sell-price standard deviations. \( \text{SKILL} \) is a proxy to capture dual trader \( i \)'s trading skill up to time bracket \( t \). \( \text{HOT} \) is a trade timing dummy that equals 1 if the 5-minute bracket in which dual trader \( i \) trades on own account belongs to the first two and final half-hour trading periods of a trading day, and 0 otherwise. \( \text{LAGPROFIT} \) is the trading profit up to and including time bracket \( t-1 \). The \( \text{INTERCEPT} \) represents the constant term in the model.
Table 3.1: Bayesian Estimates for the Median Dual Trader under the Single Equation Model

This table reports the single equation estimates for the median dual trader in each of our eight futures contracts with the corresponding posterior standard deviations in parentheses. Note that the number after the futures contract in each column denotes the specific median dual trader in that contract whose posterior estimates are provided right beneath. The dependent variable in equation (1) is \( \text{TRADE DUMMY} \), which equals 1 if dual trader \( i \) trades on her own account in time bracket \( t \), 0 otherwise. The explanatory variables in equation (1) are as follows.

- \( \text{LagNLOCAL} \) is the number of pure locals (sole own account traders) in the 5-minute bracket prior to the current time bracket \( t \).
- \( \text{LagFRACTI4} \) is dual trader \( i \)'s customer-trading volume as a fraction of her total trading volume in time bracket \( t-1 \).
- \( \text{LagVOLCTI1} \) is dual trader \( i \)'s own account trading volume in time bracket \( t-1 \).
- \( \text{LagVOLATILITY} \), in time bracket \( t-1 \), is obtained as the maximum of the buy-price and sell-price standard deviations.
- \( \text{LagINVENTORY} \), is computed as dual trader \( i \)'s CTI 1 buy trades minus CTI 1 sell trades, cumulated from the beginning of a trading day to time bracket \( t-1 \).
- \( \text{LagSKILL} \) is a proxy to capture dual trader \( i \)'s trading skill up to time bracket \( t-1 \).
- \( \text{HOT} \) is a trade timing dummy that equals 1 if the 5-minute bracket in which dual trader \( i \) trades on own account belongs to the first two and final half-hour trading periods of a trading day, and 0 otherwise.

The dependent variable in equation (2) is \( \text{PROFIT} \), which is computed by cumulating dual trader \( i \)'s personal trading profit from the beginning of a trading day up to time bracket \( t \). The explanatory variables in equation (2) are:

- \( \text{TRADE DUMMY} \)
- \( \text{VOLATILITY} \)
- \( \text{SKILL} \)
- \( \text{HOT} \)
- \( \text{LagPROFIT} \), which is the trading profit up to and including time bracket \( t-1 \).

Panel A. Probit Regression – Explaining the Own Account Trading Decision

<table>
<thead>
<tr>
<th>Futures Contract</th>
<th>Parameter Estimate</th>
<th>Parameter Estimate</th>
<th>Parameter Estimate</th>
<th>Parameter Estimate</th>
<th>Parameter Estimate</th>
<th>Parameter Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live Cattle</td>
<td>20</td>
<td>Hogs</td>
<td>08</td>
<td>Pork Bellies</td>
<td>09</td>
<td></td>
</tr>
</tbody>
</table>
| Panel B. Explaining the Trading Profit

<table>
<thead>
<tr>
<th>Futures Contract</th>
<th>INTERCEPT</th>
<th>TRADE DUMMY</th>
<th>VOLATILITY</th>
<th>SKILL</th>
<th>HOT</th>
<th>LagPROFIT</th>
</tr>
</thead>
</table>
| Panel C. The Variance-Covariance Matrix

\[
\Sigma 
\]

\[
\sigma 
\]

\[
\sigma_1^2 = 14.6967 (5.5672)
\]

\[
\sigma_2^2 = 25.8787 (5.8611)
\]

\[
\sigma_3^2 = 62.9523 (15.7127)
\]

\[
\sigma_4^2 = 17.0021 (8.237)
\]

\[
\sigma_{12} = 5.3215 (1.4427)
\]

\[
\sigma_{13} = 1.4514 (0.4829)
\]

\[
\sigma_{14} = 1.4041 (0.4515)
\]

\[
\sigma_{23} = 9.7105 (3.3166)
\]

\[
\sigma_{24} = 3.9369 (1.357)
\]

\[
\sigma_{34} = 1.9695 (0.7238)
\]
Table 3.2: Bayesian Estimates for the Median Dual Trader under the Simultaneous equations Model

This table reports the simultaneous equations estimates for the median dual trader in each of our eight futures contracts with the corresponding posterior standard deviations in parentheses. The number after the futures contract in each column denotes the specific median dual trader in that contract whose posterior estimates are provided right beneath. The dependent variable in equation (1) is \( \text{TRADE DUMMY} \) \((I)\), which equals 1 if dual trader \( i \) trades on her own account in time bracket \( t \), 0 otherwise. The explanatory variables in equation (1) are as follows.

- \( \text{LagNLOCAL} \) is the number of pure locals (sole own account traders) in the 5-minute bracket prior to the current time bracket \( t \).
- \( \text{LagFRACTI4} \) is dual trader \( i \)'s customer-trading volume as a fraction of her total trading volume in time bracket \( t-1 \).
- \( \text{LagVOLCTI1} \) is dual trader \( i \)'s own account trading volume in time bracket \( t-1 \).
- \( \text{LagVOLATILITY} \), in time bracket \( t-1 \) is obtained as the maximum of the buy-price and sell-price standard deviations.
- \( \text{LagINVENTORY} \), is computed as dual trader \( i \)'s CTI 1 buy trades minus CTI 1 sell trades, cumulated from the beginning of a trading day to time bracket \( t-1 \).
- \( \text{LagSKILL} \) is a proxy to capture dual trader \( i \)'s trading skill up to time bracket \( t-1 \).

The dependent variable in equation (2) is \( \text{PROFIT} \) \((\Pi)\), which is computed by cumulating dual trader \( i \)'s personal trading profit from the beginning of a trading day up to time bracket \( t \). The explanatory variables in equation (2) are:

- \( \text{TRADE DUMMY} \);
- \( \text{VOLATILITY} \);
- \( \text{SKILL} \);
- \( \text{HOT} \);
- \( \text{LagPROFIT} \), which is the trading profit up to and including time bracket \( t-1 \).

The variance-covariance matrix is as follows:

<table>
<thead>
<tr>
<th></th>
<th>INTERCEPT</th>
<th>TRADE DUMMY</th>
<th>VOLATILITY</th>
<th>SKILL</th>
<th>HOT</th>
<th>LAGPROFIT</th>
<th>LAGINVENTORY</th>
<th>LAGVOLATILITY</th>
<th>LAGVOLCTI1</th>
<th>LAGFRACTI4</th>
<th>LAGSKILL</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTERCEPT</td>
<td>0.00</td>
<td>-0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>TRADE DUMMY</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>VOLATILITY</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>SKILL</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>HOT</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>LAGPROFIT</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>LAGINVENTORY</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>LAGVOLATILITY</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>LAGVOLCTI1</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>LAGFRACTI4</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>LAGSKILL</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Panel C: The Variance-Covariance Matrix
Table 3.3: Pooled Bayesian Estimates for the Median Dual Trader in Each Futures Contract

This table reports the pooled estimates for the median dual trader in each of our eight futures contracts with the corresponding pooled posterior standard deviations in parentheses. Note that the number after the futures contract in each column denotes the specific median dual trader in that contract whose posterior estimates are provided right beneath. The pooled estimates are obtained as weighted averages of the corresponding parameter estimates from the single equation model (table 4.1) and the simultaneous equations model (table 4.2) respectively, based on the Bayes factor values presented below. According to Kass and Raftery (1995), there exists decisive evidence from the sample data against $H_1$ when $\text{BF}_{01}$ exceeds 100.

### Panel A. Probit Regression – Explaining the Own Account Trading Decision

<table>
<thead>
<tr>
<th>Futures Contract</th>
<th>INTERCEPT</th>
<th>LagNLOCAL</th>
<th>LagFRACTI4</th>
<th>LagVOLCTI1</th>
<th>LagVOLATILITY</th>
<th>LagINVENTORY</th>
<th>LagSKILL</th>
<th>HOT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live Cattle</td>
<td>.9414 (.1126)</td>
<td>-.0689 (.0070)</td>
<td>-.7404 (.0631)</td>
<td>.0017 (.0094)</td>
<td>-.0062 (.0025)</td>
<td>.0321 (.0054)</td>
<td>-.00013 (.00003)</td>
<td>.0918 (.0517)</td>
</tr>
<tr>
<td>Hogs</td>
<td>1.3684 (.1461)</td>
<td>.0198 (.0122)</td>
<td>-1.120 (.0664)</td>
<td>.0060 (.0144)</td>
<td>-.0052 (.0039)</td>
<td>.0100 (.0056)</td>
<td>.00003 (.00003)</td>
<td>-.0933 (.0568)</td>
</tr>
<tr>
<td>Pork Bellies</td>
<td>.0781 (.0848)</td>
<td>.0572 (.0144)</td>
<td>-1.190 (.0638)</td>
<td>.0249 (.0125)</td>
<td>-.0016 (.0023)</td>
<td>.0433 (.0130)</td>
<td>.00004 (.00005)</td>
<td>.2121 (.0520)</td>
</tr>
<tr>
<td>Feeder Cattle</td>
<td>.0489 (.0870)</td>
<td>.0167 (.0133)</td>
<td>-.9460 (.0670)</td>
<td>-.0024 (.0101)</td>
<td>-.0002 (.0023)</td>
<td>.0253 (.0058)</td>
<td>.00002 (.00003)</td>
<td>.0545 (.0403)</td>
</tr>
<tr>
<td>Lumber</td>
<td>-.1184 (.0952)</td>
<td>.0035 (.0153)</td>
<td>-1.116 (.0728)</td>
<td>-.0405 (.0160)</td>
<td>.0042 (.0029)</td>
<td>.1221 (.0211)</td>
<td>-.00011 (.00006)</td>
<td>.4251 (.0591)</td>
</tr>
<tr>
<td>Canadian Dollar</td>
<td>.2909 (.0428)</td>
<td>.0194 (.0117)</td>
<td>-.8186 (.0405)</td>
<td>.0071 (.0025)</td>
<td>-.0011 (.0018)</td>
<td>.0123 (.0017)</td>
<td>.00022 (.00008)</td>
<td>-.1900 (.0327)</td>
</tr>
<tr>
<td>T-bill</td>
<td>.3002 (.0484)</td>
<td>-.0050 (.0123)</td>
<td>-.4959 (.0381)</td>
<td>.0009 (.0026)</td>
<td>.0013 (.0014)</td>
<td>.0081 (.0013)</td>
<td>.00004 (.00004)</td>
<td>.2710 (.0337)</td>
</tr>
<tr>
<td>S&amp;P 400</td>
<td>.0324 (.1197)</td>
<td>.0772 (.0367)</td>
<td>-1.120 (.1040)</td>
<td>-.0707 (.0360)</td>
<td>-.0033 (.0091)</td>
<td>.1402 (.0284)</td>
<td>-.0015 (.0005)</td>
<td>.1510 (.0701)</td>
</tr>
</tbody>
</table>

### Panel B. Explaining the Trading Profit

<table>
<thead>
<tr>
<th>Futures Contract</th>
<th>INTERCEPT</th>
<th>TRADE DUMMY</th>
<th>VOLATILITY</th>
<th>SKILL</th>
<th>HOT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live Cattle</td>
<td>10.253 (27.453)</td>
<td>-7.249 (12.985)</td>
<td>.1515 (.0086)</td>
<td>.0155 (.0086)</td>
<td>-34.717 (13.401)</td>
</tr>
<tr>
<td>Hogs</td>
<td>-23.480 (25.422)</td>
<td>29.920 (11.588)</td>
<td>.6327 (.6684)</td>
<td>.0031 (.0057)</td>
<td>8.754 (9.135)</td>
</tr>
<tr>
<td>Pork Bellies</td>
<td>19.823 (14.571)</td>
<td>7.162 (11.169)</td>
<td>-.2843 (.5011)</td>
<td>-.0241 (.124)</td>
<td>-4.674 (11.366)</td>
</tr>
<tr>
<td>Feeder Cattle</td>
<td>-.7322 (10.655)</td>
<td>15.489 (7.042)</td>
<td>.5736 (.3711)</td>
<td>.0127 (.0055)</td>
<td>-7.267 (6.596)</td>
</tr>
<tr>
<td>Lumber</td>
<td>12.015 (20.000)</td>
<td>33.752 (15.462)</td>
<td>.2473 (.7999)</td>
<td>-.0091 (.169)</td>
<td>4.411 (15.426)</td>
</tr>
<tr>
<td>Canadian Dollar</td>
<td>11.378 (11.887)</td>
<td>35.626 (11.905)</td>
<td>.7235 (.6941)</td>
<td>-.0345 (.319)</td>
<td>-29.308 (12.003)</td>
</tr>
<tr>
<td>T-bill</td>
<td>-46.651 (25.264)</td>
<td>30.417 (20.632)</td>
<td>1.857 (.8501)</td>
<td>-.0345 (.319)</td>
<td>58.085 (19.949)</td>
</tr>
<tr>
<td>S&amp;P 400</td>
<td>48.182 (24.596)</td>
<td>35.558 (22.243)</td>
<td>2.187 (2.688)</td>
<td>.1359 (.0247)</td>
<td>-6.512 (20.234)</td>
</tr>
</tbody>
</table>

### Panel C. The Variance-Covariance Matrix

<table>
<thead>
<tr>
<th>Futures Contract</th>
<th>$\sigma_{e1e2}$</th>
<th>$\sigma_{e2e2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live Cattle</td>
<td>.0015 (3557)</td>
<td>141386.83 (3403.63)</td>
</tr>
<tr>
<td>Hogs</td>
<td>.0204 (3524)</td>
<td>74514.76 (1680.63)</td>
</tr>
<tr>
<td>Pork Bellies</td>
<td>.0348 (3549)</td>
<td>97047.90 (2298.99)</td>
</tr>
<tr>
<td>Feeder Cattle</td>
<td>.0080 (3537)</td>
<td>53197.59 (1014.42)</td>
</tr>
<tr>
<td>Lumber</td>
<td>-.0033 (3525)</td>
<td>146971.75 (3896.59)</td>
</tr>
<tr>
<td>Canadian Dollar</td>
<td>-.0039 (3564)</td>
<td>258840.43 (4213.76)</td>
</tr>
<tr>
<td>T-bill</td>
<td>-.0062 (3554)</td>
<td>629400.55 (11110.70)</td>
</tr>
<tr>
<td>S&amp;P 400</td>
<td>.0077 (3555)</td>
<td>170019.24 (5857.32)</td>
</tr>
</tbody>
</table>

$BF_{01}$ values are also provided for each futures contract, with $BF_{01}$ values exceeding 100 indicating decisive evidence against the null hypothesis ($H_1$) of no own account trading.
Table 4: Tests of Heterogeneity across the Median Dual Traders

This table provides evidence on the heterogeneity across the median dual trader in each of our eight futures contracts. For pair-wise comparison, we employ both the standard two-sample t-test and a nonparametric procedure (Wilcoxon) to test that the posterior distributions of the parameters have the same location across two distinct traders. With eight traders, one from each of the eight contracts, we perform a total of twenty-eight (7×8/2) pair-wise comparisons. For each parameter, the number under T-test (Wilcoxon) provides the fraction of the corresponding test statistic with a p-value below .01 for the null hypothesis of equality of the means (medians) across each pair. For a simultaneous comparison of the medians of a given parameter, across all eight dual traders, we employ the nonparametric Kruskal-Wallis test. For each parameter, the number under Kruskal-Wallis provides the p-value associated with the test statistic for the null hypothesis of equality of the eight medians.

<table>
<thead>
<tr>
<th>1000</th>
<th>1000</th>
<th>1000</th>
<th>1000</th>
<th>1000</th>
<th>1000</th>
<th>1000</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LagNLOCAL</th>
<th>LagFRACTI4</th>
<th>LagVOLCTI1</th>
<th>LagVOLATILITY</th>
<th>LagINVENTORY</th>
<th>LagSKILL</th>
<th>HOT</th>
<th>TRADE DUMMY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>96.4%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

This table provides evidence on the heterogeneity across the median dual traders.