The Basis Effects of Failures to Converge

by

Berna Karali, Kevin McNew, and Walter Thurman

Suggested citation format:

Karali, B., Kevin McNew, and Walter Thurman. 2010. “The Basis Effects of Failures to Converge.” Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO. [http://www.farmdoc.illinois.edu/nccc134].

The Basis Effects of Failures to Converge

Berna Karali

Kevin McNew

Walter N. Thurman

Paper presented at NCCC-134 Conference on
Applied Commodity Price Analysis, Forecasting, and Market Risk Management
St. Louis, Missouri April 19-20, 2010

Berna Karali is an Assistant Professor in the Department of Agricultural Economics at the University of Georgia (bkarali@uga.edu); Kevin McNew is the President of Cash Grain Bids, Inc. in Bozeman, Montana (kmcnew@cashgrainbids.com); Walter Thurman is a William Neal Reynolds Professor of Agricultural and Resource Economics at North Carolina State University (wally_thurman@ncsu.edu).

The Basis Effects of Failures to Converge

Abstract: *We study the spatial patterns to wheat basis (spot price minus futures price) for wheat contracts between 2005 and 2009. Restricting our attention to a single delivery market—Toledo, Ohio— and to cash markets within 100 miles of Toledo, we measure the co-movement of basis at inland markets with basis at Toledo. We examine the degree to which that co-movement is disrupted for contracts that failed to converge at delivery time.*

Keywords: wheat futures, convergence, basis

Introduction

Recent failures of convergence between delivery-month futures prices and spot prices, primarily in wheat markets, have received considerable attention. Evidence of concern at high levels is provided by Congress's authorization of an *ad hoc* committee to investigate the matter. For documentation and a review of the issues related to nonconvergence, see Irwin et al. (2008 and March 2009).[1]

The convergence of futures and spot prices during the delivery period of a contract is fundamental to the efficient workings of futures markets. Specifically, the forecasting value of futures prices prior to delivery is based on market participants' expectations that the spot and futures prices will converge. The fact that wheat futures contracts have in recent years sometimes only come within a dollar per bushel of spot prices is, therefore, cause for concern.

We contribute to an understanding of the non-convergence phenomenon by analyzing the effects of episodes of non-convergence on spatial basis patterns, with attention paid to both delivery and non-delivery markets. Since only a small subset of markets comprises delivery points, most market participants are subject to basis conditions at non-delivery points. As pointed out by Irwin et al., basis levels in non-delivery markets may be influenced by lack of convergence at delivery points, but no empirical research has explored the issue. We address the question: if delivery markets fail to converge, is the spot-futures discrepancy at the delivery point spatially amplified or damped?

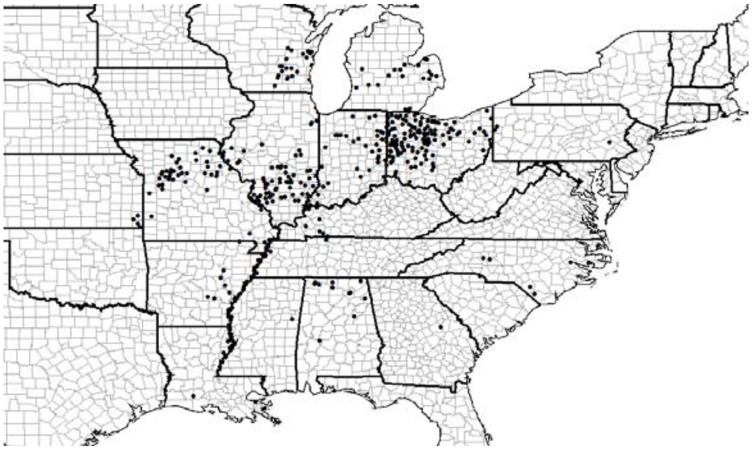
Our analysis exploits a unique and proprietary data set from Cash Grain Bids Inc. The data comprise daily spot bid prices from 380 wheat buyers scattered throughout the Midwest and

---

[1] Irwin et al. relate convergence failures to the slope of the delivery-time profile of futures prices. When the price spread between successive contracts rises close to the cost of storage, delivery-month convergence failures are more likely to occur. They also argue that recent changes in the CBOT corn and soybean contracts have ameliorated the problem in those markets, but that more fundamental changes in the delivery terms of wheat contracts are required. For an analysis of earlier episodes of non-convergence, see Pirrong, Haddock, and Kormendi (1993).

Eastern portions of the United States (see figure 1) from January 2005 to September 2009. Using a panel regression model we measure the spatial basis surface (the futures-spot differential as a function of distance to delivery points) and examine perturbations in that surface during non-convergence episodes. We also examine the extent to which convergence problems at delivery points have lead to weaker basis levels at non-delivery markets. Because non-convergence problems have been most pronounced for wheat contracts, we restrict attention to wheat.

Figure 1. Wheat Cash Markets Used for the Analysis (n=380)
Panel includes data from 2005 to 2008



Data

Cash Grain Bids Inc (CGBI) collects grain bid prices from over 3,500 U.S. grain buyers every trading day. Prices are collected on corn, soybeans, five classes of wheat, and minor grain and oilseed crops. Along with spot bids for immediate delivery, month-by-month forward prices are collected up to a year in advance. The prices collected from grain buyers is referred to as a "posted bid price," a common metric used in the industry. It is reported by the grain buyer after the futures market has closed at the end of the day and represents the price they are willing to pay for grain delivered that meets normal grading standards. No premiums or deductions for quality or moisture are reported in these prices.

The data used in this study comprise a subset of data from CGBI for Soft Red Winter Wheat—Wheat (SRW)—in the Eastern United States. These markets represent a wide array of merchants and users including wheat millers, export terminals, local coops, as well as delivery terminals for the underlying CBT wheat futures contract. Each market is geo-coded for the delivery location of the grain providing an exact reference for calculating distance. The markets are displayed in figure 1.

Data from the cash grain markets are matched to daily settlement prices of CBT soft red winter wheat futures contracts. There are five delivery months in each year: March, May, July, September, and December. We collect data over the life of each of these five contracts during the five years of 2005 through 2009. Thus, we have daily data on 25 different wheat contracts and contemporaneous cash prices.

The widely noted failures to converge in wheat futures contracts are illustrated in figure 2. Plotted there are the average values of basis (cash price minus futures price), over the last 20 days of trading, for each of the 25 contracts and for each of three delivery locations in Toledo, Chicago, and St. Louis. Each point plotted shows by how much the spot price at a delivery location deviated from the futures price in the 90 days prior to delivery—a time when fundamental notions of market efficiency predict that spot prices should be converging to futures prices. The convergence behavior across contracts was similar across the three delivery locations and spot prices were significantly below futures prices at the time of expiration for several contracts. In 2006 and 2007, basis the 90 days prior to expiration in Toledo was as large as 50 cents/bushel, spot below futures. Failures to converge continued into 2008, reaching a peak of almost \$1.50, again spot below futures.

The convergence behavior is similar across the three delivery points listed. We will focus in what follows on basis relationships relative to the Toledo delivery location. Figure 3 is a map displaying Toledo and all the markets with straight-line distance to Toledo of 100 miles or less. The void to the northeast of Toledo is Lake Erie.

Empirics

We consider two versions of a spatial basis regression model. One estimates fixed effects for each market. A second is much more parsimonious, representing markets by their distances from the delivery location.

A Basis Model with Market Fixed Effects – the Effects of Distance to Market

We first consider a spatial basis model of the following form:

$$(1)\ b_{it} = \phi_i + \beta b_{Dt} + \varepsilon_{it},$$

for i = 1, ..., n and t = 1, ..., T. The variable b_{it} is the basis (spot – futures) on day t in market i and b_{Dt} is the basis at the delivery point relevant to market i on day t. The basis model contains a fixed effect for each market.

The basis basis model captures both temporal and spatial variation in basis. Spatial aspects are captured by the fixed effects, φ_i, which should vary in equilibrium with distance from the delivery market. Temporal aspects are captured by β, which measures the extent to which basis at markets different from the delivery point vary with basis movements at the delivery point. While, in principle, the degree to which basis away from the delivery point moves with basis at the delivery point might vary with distance from the delivery point, equation (1) forces that response to be the same across all markets. That restriction will be loosened in subsequent specifications. Note, further, that (1) makes no claim that basis movements at the delivery location are predetermined with respect to variations away from the delivery point. We view OLS estimates of the parameters of (1) as reflecting partial correlations and not unidirectional causality.

To implement (1) empirically requires a specification of delivery location, a universe of markets relevant to the delivery location, and a set of time series observations from which to estimate. In the following, Toledo is the delivery location and we include as outlying markets all those markets within 100 miles of Toledo (see figure 3). Because we are interested in basis behavior close to contract expiration, we estimate (1) using all basis observations within 90 days of expiration.

Figure 4 displays results from model (1) estimated as 25 OLS regressions, one for each wheat contract that expired during the 2005-2009 period. The purpose of figure 4 is to investigate how the market fixed effects relate to distance from Toledo.

Each of the 25 regressions is a linear basis regression, with dummy variables representing each of the 130-170 markets with observations for the contract. (The number of reporting markets varies some with contract.) Because no column of ones is included in the regression, the fixed effects can be interpreted as the market-specific intercepts in the basis regressions. Because Toledo basis is the regressor, the market fixed effects reflect the number of cents that basis at the location is expected to lie over the basis at Toledo, when basis at Toledo is zero. The intercept should reflect, in equilibrium, the costs of transporting grain from the location to Toledo.

To assess the reasonableness of this interpretation, the intercepts for each location are averaged across the 25 contracts. But before averaging, the market fixed effects are mean-corrected by contract. That is, each fixed effects is expressed as a deviation from the mean fixed effect for that contract. This is done to remove the effects of non-convergent contracts.

Once the mean-corrected fixed effects are calculated, averages across contracts are calculated for each market. This leaves 174 estimated intercepts: the number of markets that are both within 100 miles of Toledo and have at least one recorded basis observation over the 90 days before expiration of at least one of the 25 contracts. The panel of fixed effects is unbalanced; one market is included in only one of the 25 regressions. To avoid problems resulting from markets that are observed only for atypical contracts, figure 4 displays the 2005-2009 averages of fixed effects for the 125 markets that have observations for nine or more contracts. Each of the points in figure 4 represents the average over the 2005-2009 period of the intercepts for a single market.

Figure 4 displays a clear downward relationship: the farther away a market location is from Toledo, the lower is basis expected to be. The estimated slope indicates that the cost of transporting grain to the delivery point is \$0.22 per bushel per 100 miles. There is quite a bit of unexplained variation (R2 = 0.30) reflecting at least in part the fact that straight-line distance is only a proxy for transportation cost—costs of fuel and other factors drive transportation cost as well. We take from this investigation that the model is at least reasonably applicable to the spatial basis data.

A Basis Model with Distance from Delivery

We next turn to results from a model that replaces market fixed effects with distance from Toledo and an interaction variable—the product of distance to Toledo and basis at Toledo. The model takes the form:

$$(2)\ b_{it} = \alpha + \beta b_{Dt} + \gamma d_i + \delta d_i b_{Dt} + \varepsilon_{it},$$

where d_i is the distance from market i to Toledo and all other variables are defined as in (1). Distance (d_i) is measured in 100s of miles and so varies between 0 and 1 in the estimation sample. We derived confidence in the consistency between specifications (1) and (2) from the fact that estimates of β are quite similar between the two models.

Specification (2) allows the relationship between delivery-location basis and market-i basis to depend upon distance:

$$(3)\ \frac{\partial b_{it}}{\partial b_{Dt}} = \beta + \delta d_i\ .$$

The results of estimating specification (2) separately for each of the 25 futures contracts are reported in Table 1.

Table 1 groups the regression results by contract delivery month, and then by year. One thing to note is that goodness-of-fit varies considerably across contracts and contract months. Coefficients of determination for May contracts are lowest, ranging from 0.10 to 0.23 with an average of 0.15. Those for December are the largest, ranging from 0.16 to 0.68 with an average of 0.46.

Estimates of the sensitivity of market basis to Toledo basis can be evaluated at any distance between 0 miles to Toledo ($d_i = 0$) and 100 miles to Toledo ($d_i = 1$). Table 1 reports the derivative evaluated at both extremes. In the column labeled "Basis at Toledo" one can see variability across contracts in the estimated coefficient with averages across years of 0.62 (March), 0.66 (May), 0.62 (July), 0.74 (September), and 1.05 (December). The grand mean across all contracts and years is 0.74, implying that a one cent increase in basis at Toledo predicts a 0.74 cent increase at a location very nearby.

The estimated direct effect of distance (γ) is to reduce basis in most of the 25 regressions. The exceptions include two regressions in which the coefficient is not statistically significantly different from zero, but also include two anomalously positive and significant effects in the July 2008 and 2009 contracts. The grand mean of the γ estimates is -26.29, implying that a location 100 miles away from Toledo has a basis (and cash price) that is 26.29 cents lower than in

Toledo. This number is reasonable to the extent that actual transportation costs are in the range of 26.29 cents per bushel per hundred miles.

The interaction coefficient (δ) captures the effect on basis co-movement from a market being located farther away from Toledo. A negative estimate for δ implies that basis farther away from Toledo is less sensitive to basis movement in Toledo—see equation (3). The grand mean of the estimates of δ is -0.12, implying less sensitivity to Toledo basis movements the farther away one is from Toledo. But there is considerable variability in the estimates of δ. Some are anomalously positive, notably in the July 2008 and 2009 regressions, which also contain positive estimates for γ. For the most part, estimates of δ are negative.

Overall, the estimates most consistent with prior expectations as to the signs and sizes of coefficients are December contracts. Coefficients of determination are the highest for the December contracts, basis declines at a reasonable rate with distance (i.e., on the order of shipping costs), and the sensitivity of market basis to Toledo basis declines with distance but is positive at all distances, and for all contracts.

The Spatial Propagation of Failures to Converge

In order to assess the possible effects of convergence failures on inland basis, one should examine the performance of these regressions for the contracts with the largest convergence problems. Reference to figure 2 identifies the 2008 contracts as the most troublesome, especially May, July, September, and December. The labels for those four contracts are shaded in table 1. For the most part, these contracts do not show any disconnection between basis at the delivery location (where convergence) was failing and basis at inland locations. The estimates of β and β + δ are not out of line with those from other contracts, leading to the tentative conclusion that cash prices continue to be propagated spatially through cash prices in the system, even during periods of non-convergence between cash and futures prices at delivery locations. Put differently, these results suggest that the disconnection between cash and futures at delivery locations is a phenomenon that has more to do with the market for futures contracts—and the delivery options for those contracts—than it does with the physical shipping and marketing of grain.

Discussion and Future Research

The fact that estimates of β, and β+δ, are substantially less than one for most contracts is evidence of some non-instantaneousness in the connection between inland locations and Toledo. If grain transport happened instantly, but at a cost, then the basis at a non-delivery location would always reflect the basis at the delivery location adjusted for transport costs. Estimates of β and β+δ less than one could reflect the fact that price differences take time to arbitrage away—suggesting that the spatial regressions should incorporate dynamic responses. We intend to pursue such specifications.

Estimates of β and $\beta+\delta$ less than one, even in the long run, could reflect a delivery location (Toledo) that is removed from the main trade flow of wheat. Grain flow through Toledo has, in fact, diminished in recent years. Future work will involve extending the geographic scope of markets considered, possibly considering the wheat marketing network that flows through, and is calibrated to, New Orleans.

Other modeling extensions will involve incorporating the contract-specific basis regressions into a system, from within which restrictions across contracts can be tested and imposed. Measures of nonconvergence will be incorporated explicitly into the spatial basis models. We also plan on taking into account non-distance factors that influence transportation costs, like fuel costs and barge shipping rates.

We view the modeling reported in this paper as a first attempt to systematically model the complex relationships among the markets represented by the Cash Grain Bids data and to uncover the implications of nonconvergence on non-delivery markets. While real markets are decidedly more complex than textbook spatial basis models (for an influential contribution, see Bressler and King [1970]), we are optimistic that application of the no-arbitrage principles that underlie such markets will be useful in understanding the effects of episodes of failures to converge and, possibly, their causes.

References

Bressler, Raymond G. Jr., and Richard A. King. *Markets, Prices, and Interregional Trade*. New York: John Wiley, 1970.

Irwin, S.H., P. Garcia, D.L. Good, and E.L. Kunda. "Poor Convergence Performance of CBOT Corn, Soybean, and Wheat Futures Contracts: Causes and Solutions," Marketing Outlook Research Project Report 2009-02, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, March 2009. Available online: http://www.farmdoc.uiuc.edu/marketing/morr/morr_09-02/MORR_09-02.pdf.

Irwin, S.H., P. Garcia, D.L. Good, and E.L. Kunda. "Recent Convergence Performance of CBOT Corn, Soybean, and Wheat Futures Contracts." Choices, 2nd Quarter, 23 (2008): 16-21. Available online: http://www.choicesmagazine.org/magazine/pdf/article_26.pdf.

Pirrong, S.C., D. Haddock, and R. Kormendi. *Grain Futures Contracts: An Economic Appraisal*. Boston, MA: Kluwer Academic Publishers, 1993.

Figure 2: Wheat Basis at Delivery Locations
Average over the 20 days prior to expiration

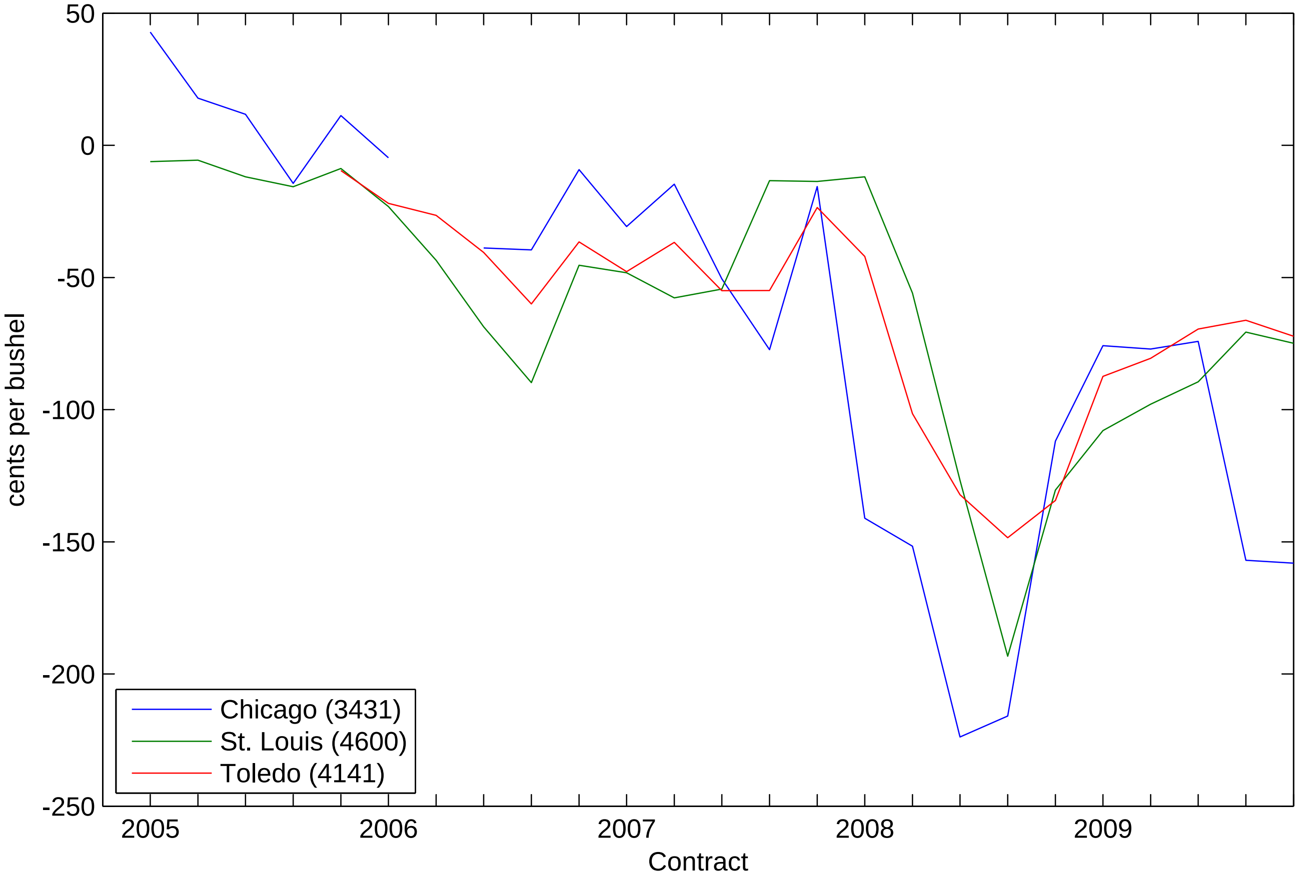


Figure 3: Markets within 100 miles of Toledo, Ohio

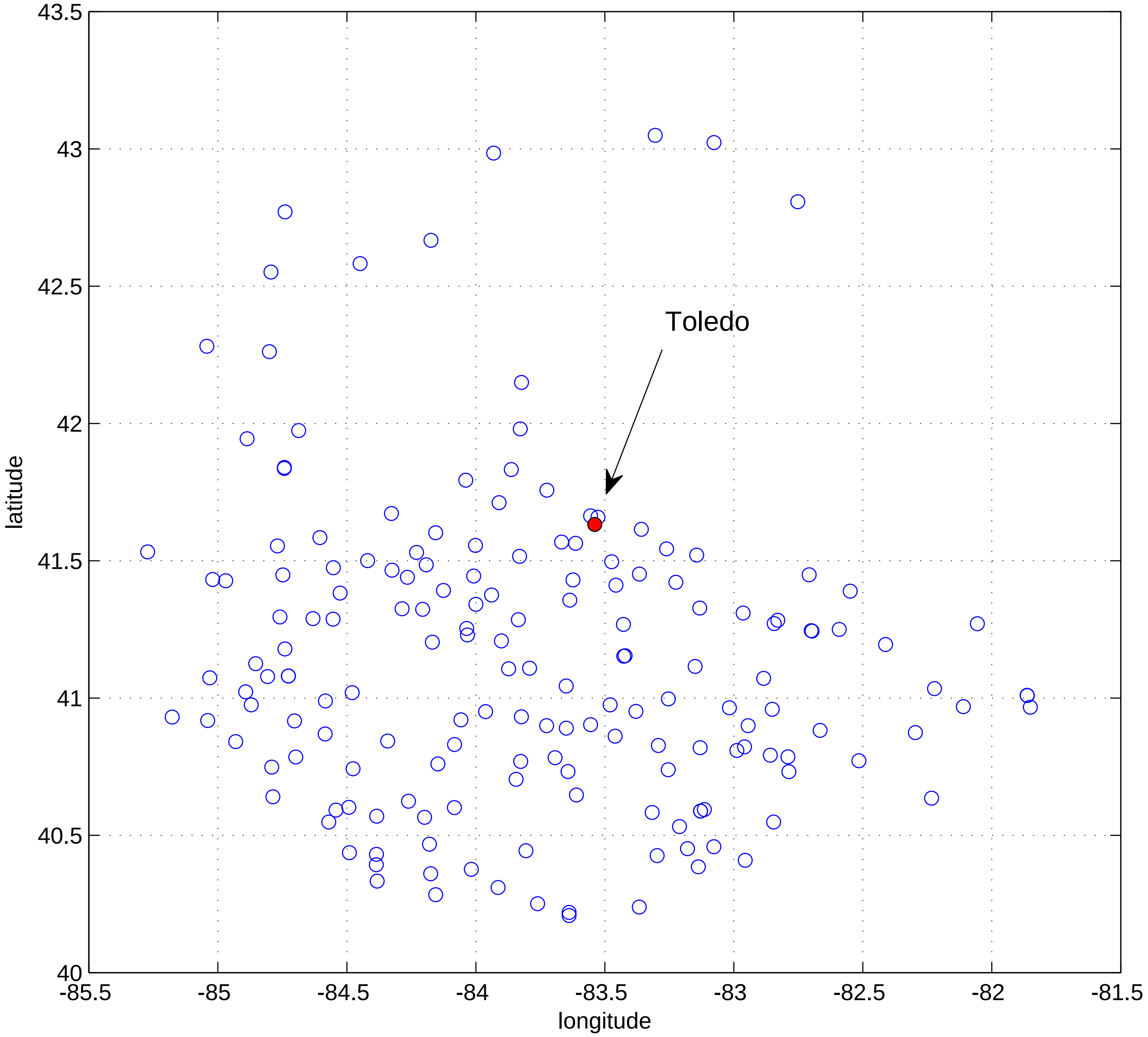


Figure 4: Basis regression intercepts by distance from Toledo
125 markets within 100 miles of Toledo

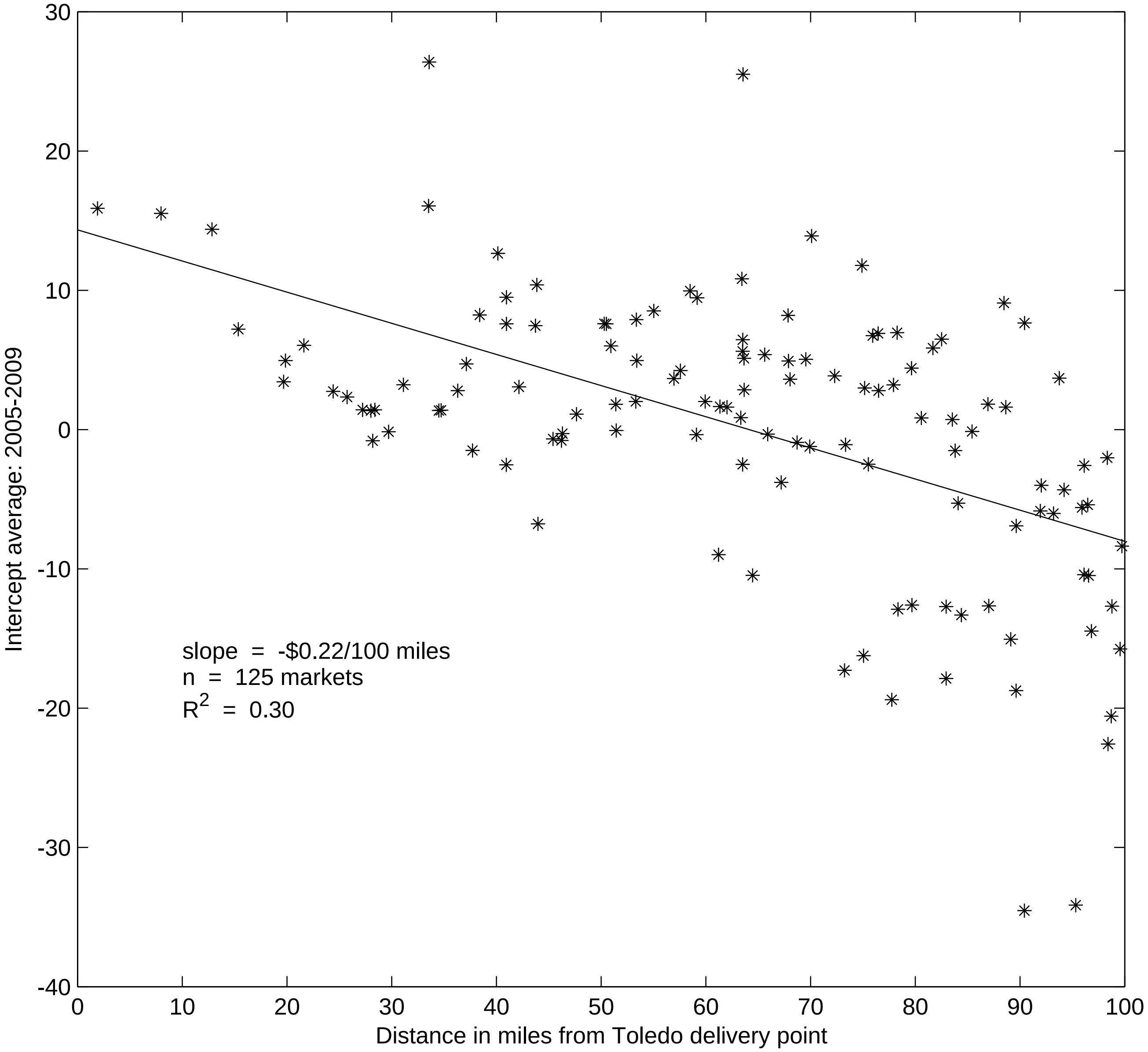


Table 1: Basis Regressions -- Toledo (markets closer than 100 miles)

Samples include all observations less than 90 days from contract expiration

	Intercept (α)	Basis at Toledo ∂bi/∂bD|d=0 (β)	distance (100s of miles) (γ)	Basis*distance (δ)	∂bi/∂bD|d=100 (fn) (β + δ)	R2	n
March contracts							
2005	-3.3 ***	0.62 ***	-5.0 ***	0.77 ***	1.39	0.34	4,330
2006	-21.5 ***	0.19	-18.6 ***	-0.15	0.03	0.14	5,603
2007	4.3	1.10 ***	14.6	0.89 ***	2.00	0.46	5,002
2008	-3.8	0.22	-50.8 ***	0.24	0.45	0.21	4,816
2009	0.9	0.99 ***	-12.5	0.02	1.01	0.31	6,550
Means	**-4.7**	**0.62**	**-14.47**	**0.35**	**0.98**	**0.29**	
May contracts							
2005	-5.6 ***	0.73 ***	-19.6 ***	-0.81 ***	-0.08	0.23	3,869
2006	-25.2 ***	0.42 ***	-26.1 ***	-0.37 **	0.05	0.10	4,670
2007	-21.6	0.62 *	30.2	0.85	1.47	0.22	3,286
2008	-13.2 ***	0.96 ***	-55.0 ***	-0.78 ***	0.18	0.08	3,791
2009	-50.2 ***	0.56 ***	-33.1 *	-0.19	0.37	0.10	6,265
Means	**-23.2**	**0.66**	**-20.74**	**-0.26**	**0.40**	**0.15**	
July contracts							
2005	-9.3 ***	0.41 *	-21.3 ***	-0.53	-0.12	0.24	3,323
2006	-10.7 ***	0.83 ***	-26.3 ***	-0.32 ***	0.50	0.25	4,553
2007	-24.8 ***	0.55 ***	-41.9 ***	-0.43 ***	0.12	0.26	3,544
2008	-36.0 ***	0.71 ***	31.3 ***	0.59 ***	1.30	0.59	3,584
2009	-45.2 ***	0.61 ***	79.5 ***	0.91 ***	1.53	0.33	6,079
Means	**-25.2**	**0.62**	**4.25**	**0.04**	**0.668**	**0.34**	
September contracts							
2005	-12.7 ***	0.30 *	-8.6 *	0.07	0.37	0.14	2,488
2006	-56.4 ***	0.43 ***	-65.1 ***	-0.74 ***	-0.31	0.01	3,671
2007	1.4	1.16 ***	-59.5 ***	-0.67 ***	0.49	0.63	3,353
2008	-44.1 ***	0.73 ***	-103.2 ***	-0.40 ***	0.32	0.48	5,780
2009	0.7	1.09 ***	-61.4 ***	-0.65 ***	0.44	0.52	6,400
Means	**-22.2**	**0.74**	**-59.56**	**-0.48**	**0.263**	**0.36**	
December contracts							
2005	-19.0 ***	0.45 ***	-16.0 ***	-0.08	0.37	0.16	4,388
2006	8.4 ***	1.14 ***	-58.1 ***	-0.33 ***	0.80	0.67	4,316
2007	13.2 ***	1.69 ***	-57.8 ***	-0.37	1.33	0.49	4,228
2008	-13.4 ***	0.97 ***	-37.2 ***	-0.08 *	0.89	0.68	5,238
2009	1.1	0.98 ***	-45.5 ***	-0.43 ***	0.55	0.28	7,723
Means	**-1.9**	**1.05**	**-42.91**	**-0.26**	**0.788**	**0.46**	
Grand means	**-15.4**	**0.74**	**-26.69**	**-0.12**	**0.618**	**0.32**	

* = significant at the 5% level
** = significant at the 1% level
*** = significant at the 0.1% level

(fn) = Basis coefficient + Basis*distance coefficient'