this divergence is zero when beliefs are symmetric around a 50% probability[5] and generally very small when prices are in the \$0.20-\$0.80 range.[6] Fourth, the sign of the deviation between prices and beliefs varies with the assumed utility function. Fifth, increasingly disperse beliefs yield a larger gap between prices and mean beliefs. And sixth, the extent of the divergence between prices and mean beliefs depends on the specific assumptions adopted about the utility function of traders, and the distribution of beliefs (particularly when beliefs are close to zero or one). Manski's model consistently delivers the most extreme results.

We now turn to trying to extract some empirical evidence on the most relevant parameters, from field data.

4. Field Evidence on the Distribution of Beliefs

We begin with a very simple, but salient, example. Throughout 2003 and 2004, Tradesports ran a prediction market in a security that paid \$1 if President Bush were re-elected, and nothing otherwise. The price of this security on Election eve was \$0.55. At the same time, pre-election polls suggested that 62 percent of the population thought that President Bush was more likely to win than John Kerry.[7]

The prediction market price and the poll result place restrictions on the distribution of beliefs (assuming that beliefs among the public are representative of traders). For different utility and belief distribution functional forms, we can derive the belief distribution that matches these two facts and examine how the observed prediction

---

[5] For a formal proof, see theorem 2 in Gjerstad (2004).

[6] This generalization may not hold for particularly disperse beliefs, and especially when beliefs are bimodal. For instance, Manski considers maximally disperse distributions so as to establish bounds on mean beliefs implied by a price. Appendix A expands on this analysis.

[7] We draw this number from various polls. Specifically, the following proportions thought Bush more likely to win in the final pre-election poll: CBS/NYT: 60% (n=920 adults polled 10/28-10/30); Gallup/CNN/USA Today: 61% (n=1013 adults polled 10/14-10/16); ABC: 62% (n=3617 adults polled 10/27-10/30); Marist: 60% (n=1300 registered voters polled 10/31); Pew: 64% (n=2804 registered voters polled 10/27-10/30); Princeton: 64% (n=1117 registered voters polled 10/27-10/39); Fox: 61% (n=1000 likely voters polled 10/17-10/18). Allocating non-respondents 50-50 (instead of dropping them) yields proportions predicting Bush that are usually around 2 percentage points lower. The Gallup question is roughly representative, asking: "Regardless of whom you support, and trying to be as objective as possible, who do you think will win the (presidential) election in November (2004) – John Kerry or George W. Bush?" The only real divergence was Fox, who asked: "Imagine you were given \$100 dollars to place a bet on the outcome of the upcoming (2004) presidential election. Which candidate – (George W.) Bush or (John) Kerry--would you put your money on to win this November?"

market price relates to central moments of the implied belief distribution. (In the absence of contrary evidence, we continue to assume that wealth is orthogonal to beliefs.)

If the perceived probability of a Bush victory is q, then the poll result and the prediction market price respectively imply that:

$$\int_{0.5}^{1} f(q)dq = 0.62 \qquad \textit{[62\% thought Bush more likely to win]}$$

$$\int_{0}^{0.55} X(q) f_{\mu,\sigma}(q)dq = \int_{0.55}^{1} X(q) f_{\mu,\sigma}(q)dq \qquad \textit{[\$0.55 was a prediction market equilibrium]}$$

In Table 2, we examine various two-parameter functions for the distribution of beliefs, and a range of utility functions, and solve for the implied mean belief. We report the mean of the implied distribution of beliefs (and display the parameters underlying these two-parameter distributions in parentheses).

Recalling that the market price of this security was \$0.55, Table 2 shows that this price is a good approximation to the mean belief under any of the specific assumptions that we examined.[8] Note that this occurs despite the fact that in some cases the belief distributions needed to reconcile the market price and poll results are highly asymmetric. Naturally this robustness partly reflects the tendency hinted at earlier that prices close to \$0.50 are typically fairly accurate.[9]

Deriving distributions of beliefs from two data points and a distributional assumption may not be particularly satisfying, so we would like a setting where we observe beliefs directly. Unfortunately data surveying expectations about the likelihood of specific events for which there are prediction or other financial market prices is rather rare.

For this reason we turn to two rather unique datasets. The first was provided to us by Probability Football, an advertising-supported free contest that requires players to

[8] Note from Table 1 that with CARA utility, risk aversion does not affect the shape of the demand function, just its slope. The same is true of quadratic utility and y^{max}. Since aggregate supply is zero, multiplying all trader's demand by a constant does not affect market prices, we report results for only one parameter value for CARA and quadratic utility. Likewise, HARA utility and CRRA utility yield betting functions of the same shape (again, allowing for a difference in slope), so we do not report results separately for HARA.

[9] Interestingly, the table also shows that – for a given set of assumptions – market prices can also reveal the dispersion of beliefs in the population. That said, unlike the inferences about the mean beliefs, inferences about the underlying dispersion of beliefs are quite sensitive to the specific assumptions adopted.

estimate the probability of victory in every NFL game in a season.[10] Including the pre-season and playoffs, this yields 259 games in the 2000 and 2001 seasons and 267 in 2002 and 2003. On average we observe the probability assessments of 1320 players in each game, for a total sample size of 1.4 million observations. Contestants are scored using a quadratic scoring rule; they receive $100 - 400(w - q)^2$, points where w is an indicator variable for whether the team wins and q is the stated probability assessment. Truthfully reporting probabilities yields the greatest expected points, a fact that is explicitly explained to contestants.

The top three players receive cash prizes. While these rank-order incentives potentially provide an incentive to add variance to one's true beliefs, it turns out that given the number of games in a season, this incentive is small. For instance, in 2003, two mock entrants to this contest that simply used prices from Tradesports and the Sports Exchange (a sports-oriented play-money prediction market run by NewsFutures.com) as their probabilities placed seventh and ninth out of almost 2,000 entrants.[11] We simulated strategies that took these prediction market prices and added or subtracted noise, finding that adding or subtracting 1 percentage point to the market price yielded the highest probability of winning a prize. Even so, quite a few players appear to believe that more variance is optimal; and about 40 percent of players report zero or one for at least ten percent of their games.[12] This is a losing strategy, despite comprising 40 percent of players, they account for only 5 percent of those who make the top 5 percent. Since we are interested in learning about the distribution of beliefs, we drop all probability reports from these players. (Qualitatively this doesn't much affect our results.)

Figure 5 reports the distribution of all probability for games in which prediction market prices are close to \$0.33. Even after cleaning the data, there are still mass points at zero and one and some clustering at focal numbers. Beyond this, the distribution appears roughly normal.

Figure 6 examines how this distribution of beliefs varies with the (real-money) prediction market price. In general, prices closely approximate the mean or median of

[10] Levitt (2004) analyzes a related sample from a different source. The advantage of our data is that they also include a measure of participants' beliefs. We are grateful to Brian Galebach for sharing these data.
[11] Servan-Schreiber, Wolfers, Pennock and Galebach (2004) used the data collected from this to compare the predictive power of real and play-money markets, finding that they were roughly equal.
[12] Probability Football has added strategy advice to its website that makes this point to players.