Demand for Crash Insurance, Intermediary Constraints, and Risk Premia in Financial Markets

Hui Chen Scott Joslin Sophie Ni*

August 4, 2015

Abstract

We propose a new measure for the variations in intermediary constraints by observing how financial intermediaries manage their tail risk exposures. Using a unique dataset on the trading activities in the market for deep out-of-the-money S&P 500 put options, we identify periods when shocks to intermediary constraints are likely to be the main driver of the variation in the net amount of trading between public investors and financial intermediaries, which enables us to infer variations in intermediary constraints from the quantities of trading. Besides its effects on option pricing, our measure of intermediary constraint shocks is a strong predictor of future returns for a wide range of financial assets, and it is connected to existing funding constraint measures based on changes in broker-dealer leverage.

*Chen: MIT Sloan and NBER (huichen@mit.edu). Joslin: USC Marshall (sjoslin@usc.edu). Ni: Hong Kong University of Science and Technology (sophieni@ust.hk). We thank Tobias Adrian, David Bates, Robert Battalio, Geert Bekaert, Harjoat Bhamra, Ken French, Nicolae Garleanu, Zhiguo He, Bryan Kelly, Lei Lian, Andrew Lo, Dmitriy Muravyev, Jun Pan, Lasse Pedersen, Steve Ross, Ken Singleton, Moto Yogo, Hao Zhou, and seminar participants at USC, MIT Sloan, HKUST, Notre Dame, Washington University in St. Louis, New York Fed, Boston Fed, Copenhagen Business School, Dartmouth, AFA, the Consortium of Systemic Risk Analytics Meeting, CICF, SITE, the OptionMetrics Research Conference, and the Risk Management Conference at Mont Tremblant for comments, and Ernest Liu for excellent research assistance. We also thank Paul Stephens at CBOE, Gary Katz and Boris Ilyevsky at ISE for providing information on the SPX and SPY option markets.

1 Introduction

In this paper, we present new evidence connecting financial intermediary constraints to asset prices. We propose to measure the tightness of intermediary constraints by observing how financial intermediaries manage their aggregate tail risk exposures. Using a unique dataset on the trading activities between public investors and financial intermediaries in the market for deep out-of-the-money put options on the S&P 500 index (henceforth referred to as DOTM SPX puts), which are effectively insurance against large market crashes, we identify periods when shocks to the intermediary constraints are likely to be the main driver of the variation in the equilibrium quantities of trading between public investors and financial intermediaries. This enables us to infer variations in intermediary constraints from the quantity of trading. We then show that a tightening of intermediary constraint according to our measure is associated with increasing expensiveness of DOTM puts, rising risk premia for a wide range of financial assets, deterioration in funding liquidity, as well as de-leveraging by broker-dealers.

Our main measure of the option market trading activity is the net amount of DOTM SPX puts that public investors in aggregate acquire each month (henceforth referred to as $PNBO$), which also reflects the net amount of the same options that broker-dealers and market-makers sell in that month. While it is well known that financial intermediaries are net sellers of these types of options during normal times, we find that $PNBO$ varies significantly over time and tends to turn negative during times of market distress.

We exploit the basic relation between the quantities of trading, measured by $PNBO$, and prices (expensiveness of SPX options), as measured by the variance premium. At higher frequency (e.g., at transaction level), a positive price-quantity relation is consistent with the presence of shocks to public investors' demand for crash insurance (demand shocks), while a negative price-quantity relation is consistent with the presence of shocks to financial intermediaries' capacity to provide crash insurance (supply shocks).[1] We

[1] We assume that public investors' demand curve is downward sloping, and financial intermediaries' supply curve is upward sloping. These assumptions are true when public investors and financial intermediaries are (effectively) risk-averse and cannot fully unload the inventory risks through dynamic hedging. Notice that supply shocks and demand shocks are not mutually exclusive. For example, a negative price-quantity

summarize the daily price-quantity relations into a monthly measure, and consider those months with negative price-quantity relations as the periods when supply shocks are the main driver of the net quantity of trading, $PNBO$. Thus, low $PNBO$ in a month with negative price-quantity relation is indicative of tight intermediary constraints.

In monthly data from January 1991 to December 2012, $PNBO$ is significantly negatively related to the option expensiveness, and this negative relation becomes stronger when jump risk in the market is higher. In daily data, the correlation between $PNBO$ and our measure of option expensiveness is negative in 159 out of 264 months. These results highlight the significant role that supply shocks play in the options market.

When interacted with the monthly indicator of negative price-quantity relation, $PNBO$ significantly predicts future market excess returns. In the full sample, a one-standard deviation decrease in $PNBO$ in a month with negative price-quantity relation is on average associated with a 4% increase in the subsequent 3-month market excess return. The adjusted R^2 of the return-forecasting regression is 15.4%. The predictive power of $PNBO$ is even stronger in the months when market jump risk is above the median level, and it is weaker and sometimes insignificant (when jump risk is low) in the months when the price-quantity relation turns positive. Besides stocks, a lower $PNBO$ also predicts higher future excess returns for high-yield corporate bonds, an aggregate hedge fund portfolio, a carry trade portfolio, and a commodity index, and it predicts lower future excess returns on long-term Treasuries and (pay-fix) SPX variance swap.

We also show that the predictability results survive an extensive list of robustness checks. They include different statistical methods for determining the significance of the predictive power, the sensitivity of the results to the financial crisis and to extreme observations, different ways to define option moneyness, and the sensitivity to constructing the quantity measure using end-of-period open interest instead of trading volume, etc.

The return predictability results above are consistent with the intermediary asset pricing theories, where the reduced risk-sharing capacity of the intermediaries raises the aggregate risk premium in the economy. Moreover, our results suggest that it is the combination of

relation does not rule out the presence of demand shocks.

a negative price-quantity relation and a lower public net-buying volume in the market for crash insurance that best identifies the tightening of intermediary constraints.

An alternative explanation of the predictability results is that PNBO is merely a proxy for standard macro/financial factors that simultaneously drive the aggregate risk premium as well as intermediary constraints. This is in contrast to the variation in intermediary constraints having a direct impact on the aggregate risk premium, which is a crucial distinction for intermediary asset pricing theories. If this alternative explanation is true, then the inclusion of the proper risk factors into the predictability regression should drive away the predictive power of $PNBO$. We find that the predictive power of $PNBO$ is unaffected by the inclusion of a long list of return predictors in the literature, including various price ratios, the consumption-wealth ratio, the variance risk premium, the default spread, the term spread, and a variety of tail risk measures. While these results do not lead to the rejection of the alternative explanation (there can always be omitted risk factors), they are at least consistent with the intermediary constraints having an unique effect on the aggregate risk premium.

Our intermediary constraint measure is related to a list of funding liquidity measures in the literature. In particular, it is significantly related to the funding liquidity measure of Fontaine and Garcia (2012) and the growth rate in broker-dealer leverage, a funding constraint measure advocated by Adrian and Shin (2010) and Adrian, Moench, and Shin (2010). When regressing market excess returns jointly on lagged $PNBO$ and other funding constraint measures, the coefficient on $PNBO$ remains significant, which suggests that $PNBO$ contains unique information about the aggregate risk premium relative to the other funding liquidity measures.

The analysis of the price-quantity dynamics suggests that when financial intermediaries switch from sellers of DOTM SPX puts to buyers (e.g., in the months following the Lehman Brothers bankruptcy), it is likely that the tightening of constraints are forcing the intermediaries to aggressively hedge their tail risk exposures, rather than the intermediaries accommodating an increase in public investors' demand to sell crash insurance. To understand the risk sharing mechanism, it is important to identify who among the public investors (retail or institutional) are the "liquidity providers" during times of distress: reducing the amount of crash insurance acquired from financial intermediaries or even starting to provide the insurance to the latter. We answer this question by comparing public investors' demand in the markets of SPX vs. SPY options (SPY options are options on the SPDR S&P 500 ETF Trust, which has a significantly higher percentage of retail customers than SPX options). Our results suggest that it is the institutional investors who are selling the DOTM puts to the financial intermediaries during periods of distress.

Our paper builds on and extends the work of Garleanu, Pedersen, and Poteshman (2009) (henceforth GPP) to incorporate supply shocks into the options market. In a partial equilibrium setting, GPP demonstrate how exogenous public demand shocks affect option prices when risk-averse dealers have to bear the inventory risks. In their model, the dealers' intermediation capacity is fixed, and the model implies a positive relation between the public demand for options and the option premium. Like GPP, the limited intermediation capacity of the dealers is a key assumption of ours, but we introduce shocks to the intermediary risk-sharing capacity and then consider the endogenous relations among public demand for options, option pricing, and aggregate market risk premium. In our empirical analysis, we separate the effects of public demand shocks and shocks to intermediary constraints, and show that the latter is linked to the time-varying risk premia for a wide range of financial assets. Our empirical strategy based on the price-quantity dynamics is motivated by Cohen, Diether, and Malloy (2007), who use a similar strategy to identify demand and supply shocks in the equity shorting market.

The recent financial crisis has highlighted the importance of understanding the potential impact of intermediary constraints on the financial markets and the real economy. Following the seminar contributions by Bernanke and Gertler (1989), Kiyotaki and Moore (1997), and Bernanke, Gertler, and Gilchrist (1999), recent theoretical developments include Gromb and Vayanos (2002), Brunnermeier and Pedersen (2009), Geanakoplos (2009), He and Krishnamurhty (2012), Adrian and Boyarchenko (2012), Brunnermeier and Sannikov (2013), among others. We provide a reduced-form approach to model the intermediary constraint, which helps us derive quantitative predictions on the effects of time-varying intermediary constraints in a tractable way.

In contrast to the fast growing body of theoretical work, there is relatively little empirical work on measuring the intermediary constraints and studying their aggregate effects on asset prices. The notable exceptions include Adrian, Moench, and Shin (2010) and Adrian, Etula, and Muir (2012), who show that changes in aggregate broker-dealer leverage is linked to the time series and cross section of asset returns. Our paper demonstrates a particular mechanism (the crash insurance market) through which intermediary constraints affect aggregate risk sharing and asset prices. By utilizing the price-quantity information, we can better isolate shocks to intermediaries' risk sharing capacity. Moreover, compared to intermediary leverage changes, our measure has the advantage of being forward-looking and available at higher (daily vs. quarterly) frequency.

The ability of option volume to predict returns has been examined in other contexts. Pan and Poteshman (2006) show that option volume predicts near future individual stock returns (up to 2 weeks). They find the source of this predictability to be the nonpublic information possessed by option traders. Our evidence of return predictability applies to the market index and to longer horizons (up to 4 months), and we argue that the source of this predictability is time-varying intermediary constraints.

Finally, our paper is related to several studies that have examined the role that derivatives markets play in the aggregate economy. Buraschi and Jiltsov (2006) study option pricing and trading volume when investors have incomplete and heterogeneous information. Bates (2008) shows how options can be used to complete the markets in the presence of crash risk. Longstaff and Wang (2012) show that the credit market plays an important role in facilitating risk sharing among heterogeneous investors. Chen, Joslin, and Tran (2012) show that the aggregate market risk premium is highly sensitive to the amount of sharing of tail risks in equilibrium.

2 Research Design

Our goal is to measure how constrained financial intermediaries are through the ways they manage their exposures to aggregate tail risks. The market of DOTM SPX put options are well-suited for this purpose. First, this market is large in terms of the economic exposures it provides for aggregate tail risks.[2] Second, compared to other over-the-counter derivatives that also provide exposures to aggregate tail risks, the exchange-traded SPX options have the advantages in liquidity and limited counterparty risk (short of exchange failure). Third, the Options Clearing Corporation (OCC) classifies exchange option transactions by investor types, which is essential for measuring of the net exposures for the specific type of investors we focus on, the financial intermediaries.

Specifically, the OCC classifies each option transaction into one of three categories based on who initiates the trade. They include public investors, firm investors, and market-makers. Transactions initiated by public investors include those initiated by retail investors and those by institutional investors such as hedge funds. Trades initiated by firm investors are those that securities broker-dealers (who are not designated market-makers) make for their own accounts or for another broker-dealer. Since we focus on financial intermediaries as a whole, it is natural to merge firm investors and market-makers as one group and observe how they trade against public investors.

We classify DOTM puts as those with strike-to-price ratio $K/S \leq 0.85$. For robustness, we also consider different strike-to-price cutoffs, as well as cutoffs that adjust for option maturity and the volatility of the S&P 500 index (which is similar to cutoffs based on option delta). Another feature of option transaction is that an order can either be an open order (to open new positions) or a close order (to close existing positions). We will focus on open orders in our analysis of option transactions, because they are less likely to be mechanically influenced by existing positions (see Pan and Poteshman, 2006).

Since options are in zero net supply, the amount of net buying by public investors is

[2] For example, based on data in December 2012, Johnson, Liang, and Liu (2014) estimate that the change in total index option value is on the order of trillions of dollars following a severe market crash, the majority of which contributed by out-of-money SPX puts.

equal to the amount of net selling by firm investors and market-makers. Thus, we construct a measure of the public net buying-to-open volume for DOTM SPX puts (abbreviated as $PNBO$). In period t (e.g., a day or a month), $PNBO_t$ is defined as

$$PBNO_t \equiv \text{public total open-buy volume}_t - \text{public total open-sell volume}_t. \tag{1}$$

$PNBO$ represents the amount of new DOTM SPX puts bought (sold if negative) by public investors in a period. Due to the growth in size of the options market and the trading volume, $PNBO$ could potentially be non-stationary. Thus, we also consider an alternative measure, which normalizes $PNBO$ by the average monthly volume of all SPX options traded by public investors over the past three months,[3]

$$PNBON_t \equiv \frac{PNBO_t}{\text{Average monthly public SPX volume over past 3 months}}. \tag{2}$$

While $PNBON$ helps address the potential issue with growth in the size of market, $PNBO$ has the advantage that it better captures the actual magnitude of the tail risk exposures transferred between public investors and intermediaries, which should matter for measuring the degree of intermediary constraints. Considering this tradeoff, we conduct all of our main analyses using both $PNBO$ and $PNBON$.

It is well documented (see e.g., Bollen and Whaley, 2004) that public investors are net buyers of index puts while financial intermediaries are net sellers during normal times. All else equal, when financial intermediaries become more constrained, their willingness to supply crash insurance (DOTM SPX puts) to the market will be reduced. It is thus tempting to infer how constrained financial intermediaries are based on the net amount of crash insurance they sell to public investors each period, as captured by $PNBO$ defined above. However, besides intermediary constraint, weak public demand can also cause the equilibrium amount of crash insurance traded between public investors and financial intermediaries to be low. The challenge is to separate the effects of supply from demand.

[3]For robustness check, we also define $PNBON$ using past 12-month average public trading volume in the denominator, which generates similar results.

We address this problem by proposing two ways to identify periods when variations in $PNBO$ are likely to be mainly driven by shocks to intermediary constraints. First, we exploit the price-quantity relation in daily data to identify periods when supply shocks are likely to be significant, which we explain in detail below.

Second, a main reason that intermediary constraint matters for the supply of DOTM index puts by financial intermediaries is the difficulty to hedge the aggregate tail risk embedded in their inventory positions. Thus, high tail risk will amplify the effect of a shock to intermediary constraints on the equilibrium quantity of options traded. To the extent that public demand does not become more volatile during such times (the level of demand could become higher), variations in $PNBO$ will then be more associated with shocks to intermediary constraints when aggregate tail risk is high, which we capture using the market jump risk measure proposed by Andersen, Bollerslev, and Diebold (2007).

Price-quantity relation We follow the identification strategy of Cohen, Diether, and Malloy (2007), who exploit the price-quantity relation to identify supply shocks and demand shocks in the securities shorting market.

To measure the expensiveness of SPX options, one would ideally like to calculate the difference between the market price of an option and its hypothetical price without any market frictions. The latter is not observable and can only be approximated by adopting a specific structural model. For simplicity and robustness, we use the variance premium (VP) in Bekaert and Hoerova (2014) as a proxy for overall expensiveness of SPX options, which is the difference between VIX^2 and the expected physical variance of the return of the S&P 500 index.[4]

Cohen, Diether, and Malloy (2007) identify shifts in shorting demand vs. shorting supply by examining the relation between the changes in the loan fee (price) and the changes in the percentage of outstanding shares on loan (quantity). In their setting, a

[4] The expected physical variance one month ahead (22 trading days) is computed using Model 8 in Bekaert and Hoerova (2014): $E_d\left[RV_{d+1}^{(22)}\right] = 3.730 + 0.108\frac{VIX_d^2}{12} + 0.199RV_d^{(-22)} + 0.33\frac{22}{5}RV_d^{(-5)} + 0.107 \cdot 22RV_d^{(-1)}$, where $RV_d^{(-j)}$ is the sum of daily realized variances from day $d-j+1$ to day d. The daily realized variance sums squared 5-minute intraday S&P500 returns and the squared close-to-open return.

simultaneous increase (decrease) in the price and quantity indicates *at least* an increase (decrease) in shorting demand, whereas an increase (decrease) in price coupled with a decrease (increase) in quantity indicates *at least* a decrease (increase) in shorting supply.

The same logic applies in our setting. The demand pressure theory of GPP predicts that a positive and exogenous shock to the public demand for DOTM SPX puts forces risk-averse dealers to bear more inventory risks. As a result, the dealers will raise the price of the option (a move along the upward-sloping supply curve). Thus, demand shocks generate a positive relation between changes in prices and quantities. Alternatively, if there are intermediation shocks that raise the degree of constraints facing financial intermediaries (e.g., due to loss of capital or tightened capital requirements), they will become less willing to provide crash insurance to public investors. Then, the premium for the DOTM SPX puts rises while the equilibrium quantity of such options traded falls (a move along the downward-sloping demand curve).

Based on this idea, we run the following regression using daily data in each month t:

$$VP_{i(t)} = a_{VP,t} + b_{VP,t}\ PNBO_{i(t)} + d_{VP,t}\ J_{i(t)} + \epsilon^v_{i(t)}, \tag{3}$$

where $i(t)$ denotes day i in month t. The presence of jumps in the underlying stock index can affect VP even when markets are frictionless. Thus, when examine the relation between VP and $PBNO$, we control for the level of jump risk J in the S&P 500 index based on the measure in Andersen, Bollerslev, and Diebold (2007).

Recall that our goal is to identify periods when variations in $PNBO$ are mainly driven by shocks to intermediary constraints. The negative coefficient $b_{VP,t} < 0$ in a month suggests that supply shocks are the dominant driver of price-quantity relations in that month, and we expect $PNBO$ to be more informative about the variation in intermediary constraints during such times. It is important to note the limitation of this strategy. A negative coefficient $b_{VP,t} < 0$ does not identify any particular supply shocks. In fact, it does not rule out the presence of demand shocks in the same month, but supply shocks must have occurred and are likely to be significant relative to demand shocks. Similarly,

$b_{VP,t} > 0$ does not rule out the presence of supply shocks in the same month, but demand shocks must have occurred.

In summary, in a month when the price-quantity relation is largely negative ($b_{VP,t} < 0$) or when the level of jump risk J_t is high, we interpret small (or negative) value for $PNBO_t$ (cumulative net-buying by public investors for the month) as sign of tight intermediary constraints.

Intermediary constraints and risk premia According to the theory of financial intermediary constraints (see e.g., Gromb and Vayanos (2002) and He and Krishnamurhty (2012)), variations in the aggregate intermediary constraints not only affect option prices, but also drive the risk premia of other financial assets. This theory implies that low $PNBO_t$, when associated with tight intermediary constraints ($b_{VP,t} < 0$ or high J_t), should imply high future expected excess returns on the market portfolio. That is, we expect $b_r < 0$ in the following predictive regressions:

$$r_{t+j\to t+k} = a_r + b_r \; I_{\{b_{VP,t}<0\}} \times PNBO_t + \epsilon_{t+j\to t+k} \tag{4a}$$

$$r_{t+j\to t+k} = a'_r + b'_r \; I_{\{J_t>\bar{J}\}} \times PNBO_t + \epsilon'_{t+j\to t+k} \tag{4b}$$

where r denotes market excess return, and the notation $t+j \to t+k$ indicates the leading period from $t+j$ to $t+k$ ($k > j \geq 0$).[5] We will set the threshold $\bar{J}$ to be the median value in the sample. Besides the market portfolio, the above prediction should apply to other risky assets as well.

The empirical strategy and testable hypotheses above are mainly based on economic intuition. In Appendix A, we present a dynamic general equilibrium model featuring time-varying intermediary constraints. The model not only helps formalize the main intuition, but generates more rigorous predictions about how intermediary constraints affect the equilibrium price-quantity dynamics in the crash insurance market, the aggregate risk premium, and intermediary leverage. Moreover, we can use the calibrated model to

[5]All of our empirical results about predictability remain essentially unchanged if we add the term $I_{\{b_{VP,t}<0\}}$ to the regression (4a) or $I_{\{J_t>\bar{J}\}}$ to (4b).

examine the quantitative effects of intermediary constraints on asset prices.

3 Empirical Results

In this section, we present the empirical evidence connecting the trading activities of S&P 500 index options between public investors and financial intermediaries to the constraints of the financial intermediaries, the pricing of index options, and the risk premium of the aggregate stock market.

3.1 Data

Figure 1 plots the monthly time series of $PNBO$ and its normalized version $PNBON$. Consistent with the finding of Pan and Poteshman (2006) and GPP, the net public purchase of DOTM SPX puts was positive for the majority of the months prior to the financial crisis in 2008, suggesting that broker-dealers and market-makers were mainly supplying crash insurance to public investors. A few notable exceptions include the period around the Asian financial crisis (December 1997), Russian default and the financial crisis in Latin America (November 1998 to January 1999), the Iraq War (April 2003), and two months in 2005 (March and November 2005).[6]

However, starting in 2007, $PNBO$ became significantly more volatile. It turned negative during the quant crisis in August 2007, when a host of quant-driven hedge funds experienced significant losses. It then rose significantly and peaked in October 2008, following the Lehman Brothers bankruptcy. As market conditions continued to deteriorate, $PNBO$ plunged rapidly and turned significantly negative in the following months. Following a series of government interventions, $PNBO$ bottomed in April 2009, rebounded briefly, and then dropped again in December 2009 when the Greek debt crisis escalated. During the period from November 2008 to December 2012, public investors on average sold 44,000 DOTM SPX puts to open new positions each month. In contrast, they bought on average 17,000 DOTM SPX puts each month in the period from 1991 to 2007.

[6]The GM and Ford downgrades in May 2005 might be related to the negative $PNBO$ in 2005.

One reason that the $PNBO$ series appears more volatile in the latter part of the sample is that the options market (e.g., in terms of total trading volume) has grown significantly over time. As the bottom panel of Figure 1 shows, after normalizing $PNBO$ with the total SPX volume (see the definition in (2)), the $PNBON$ series no longer demonstrates visible trend in volatility over time.

Table 1 reports the summary statistics of the option volume and pricing variables. From January 1991 to December 2012, the public net buying-to-open volume of DOTM SPX puts ($PNBO$) is close to 10,000 contracts per month on average (each contract is 100 times the index). In comparison, the average total open interest for all DOTM SPX puts is around 0.9 million contracts during the period from January 1996 to December 2012, which highlights the significant difference between $PNBO$ and open interest. The option volume measures have relatively modest autocorrelations at monthly frequency (0.61 for $PNBO$ and 0.48 for $PNBON$) compared to standard return predictors such as dividend yield and term spread. In addition, $PNBO$ is weakly positively correlated with industrial production growth (0.17) and negatively correlated with the unemployment rate (-0.48).

Figure 2 provides information about the trading volume of SPX options at different moneyness. Over our entire sample, put options account for 63% of the total trading volume of SPX options. Among put options, out-of-the-money puts account for over 75% of the total trading volume; in particular, DOTM puts (with $K/S < 0.85$) account for 23% of the total volume. These statistics demonstrate the importance of the market for DOTM SPX puts.

While financial intermediaries can partially hedge the risks of their option inventories through dynamic hedging, the hedge is imperfect and costly. This is especially true for DOTM SPX puts, because they are highly sensitive to jump risk that are difficult to hedge. To demonstrate this point, we regress put option returns on the returns of the corresponding hedging portfolios at both weekly and daily horizons. The R^2s of these regressions demonstrate how effective the hedging methods are. We restrict the options to be between 15 and 90 days to maturity to ensure liquidity. We consider two hedging portfolios, one based on delta hedging (using the S&P 500 index) and one based on delta-gamma hedging.[7]

As Table 2 shows, with daily (weekly) rebalancing, delta hedging can capture around 72% (76%) of the return variation of ATM SPX puts, but only 41% (34%) of the return variation of DOTM puts. With delta-gamma hedging, the R^2 for ATM puts can exceed 90%, but it is still below 60% for DOTM puts.[8] These results imply that when holding non-zero inventories of DOTM SPX puts, financial intermediaries will be exposed to significant inventory risks even after dynamically hedging these positions. It is because of such inventory risks that financial intermediaries become more reluctant to supply crash insurance to the public investors when they are more constrained.

3.2 Option volume and the expensiveness of SPX options

We start by investigating the link between $PNBO$ and the expensiveness of SPX options as proxied by the variance premium (VP) in Bekaert and Hoerova (2014). Before constructing the measure $b_{VP,t}$ in Equation (3) for the price-quantity relation based on daily data, we first examine the relation between $PNBO$ and VP at monthly frequency.

Table 3 reports the results. In both the cases of $PNBO$ and $PNBON$, the coefficient b_{VP} is negative and statistically significant, consistent with the hypothesis that shocks to intermediary constraints generate a negative relation between the equilibrium quantities of DOTM SPX puts that public investors purchase and the expensiveness of SPX options. After adding the interaction between $PNBO_t$ and the jump risk measure J_t into the regression, we see that the coefficient c_{VP} of the interaction term is significantly negative. This result implies that $PNBO$ and VP are more likely to be negatively related during times of high jump risk, and their relation can turn positive when jump risk is sufficiently low. This result is also consistent with the intermediary constraint theory, as the effects of shocks to intermediary constraints on the supply of DOTM SPX puts by financial intermediaries tend to strengthen when the aggregate tail risk is high.

[7] For delta-gamma hedging, we use at-the-money puts expiring in the following month in addition to the S&P500 index.

[8] In this paper, we denote unadjusted R-squared by R^2, and denote adjusted R-squared by $\bar{R}^2$.

In contrast, when we replace $PNBO$ with the public net buying volume for all other SPX options excluding DOTM puts ($PNBO_{ND}$), not only are the R^2 of the regressions smaller, but the regression coefficients b_{VP} and c_{VP} are no longer significantly different from zero. As we have demonstrated in Table 2, DOTM SPX puts are more difficult to hedge and hence expose intermediaries to higher inventory risks. Thus, the trading activities of DOTM SPX puts are likely to be more informative about the fluctuations in intermediary constraints compared to those for other options.

As Garleanu, Pedersen, and Poteshman (2009) show, exogenous public demand shocks can generate a positive relation between net public demand for index options and measures of option expensiveness. They find support for this prediction using data from October 1997 to December 2001. Bollen and Whaley (2004) also find evidence of the effects of public demand pressure on option pricing in daily data.

Our results above are not a rejection of the effect of demand shocks on option prices. In Appendix B, we replicate the results of Table 2 in GPP and show that the different time periods is the main reason for the opposite signs of the price-quantity relation in the two papers. Conceptually, the demand pressure theory and the intermediary constraint theory share the common assumption of constrained intermediaries, and both can be at work in the data. For instance, our results in Table 3 show that the price-quantity relation is more likely to be negative (positive) when the jump risk in the market is high (low), indicating that the supply (demand) effects tend to become dominant under such conditions.

Next, we estimate the monthly price-quantity relation measure $b_{VP,t}$ from regression (3) using daily data. The fact that demand effects and supply effects are both present in the data is again evident. Out of 264 months, the coefficient $b_{VP,t}$ is negative in 159 (significant at 5% level in 44 of them), and positive in 105 (significant at 5% level in 24 of them). These statistics suggest that, according to the price-quantity relation, shocks to intermediary constraints are present in a significant part of our sample period. The months that have significantly negative price-quantity relations include periods in the Asian financial crisis, Russian default, the 2008 financial crisis, and several episodes during the European debt crisis.

3.3 Option volume and risk premia

We now examine the predictions from Section 2 linking $PNBO$ and risk premia in the financial markets.

First, we run the return-forecasting regressions in (4a)-(4b). Panel A of Table 4 shows that $PNBO$ has strong predictive power for future market excess returns up to 4 months ahead in months when $b_{VP,t} < 0$. The coefficient estimate b_r for predicting one-month ahead market excess returns is -21.13 (with a t-stat of -2.63),[9] with $\bar{R}^2$ of 3.2%. For 4-month ahead returns ($r_{t+3\rightarrow t+4}$, or simply r_{t+4}), the coefficient estimate is -16.71 and marginally statistically significant (with a t-stat of -1.72), and $\bar{R}^2$ drops to 1.9%. From 5 months out, the predictive coefficient is no longer statistically significant. When we aggregate the effect for the cumulative market excess returns in the next 3 months, the coefficient b_r is -80.54 (with a t-stat of -3.22) and $\bar{R}^2$ is 15.4%. The economic significance that this coefficient estimate implies is striking. When $b_{VP,t} < 0$, a one-standard deviation decrease in $PNBO$ is associated with a 4.0% (non-annualized) increase in the future 3-month market excess return.

Since Figure 1 indicates that non-stationarity might be a potential concern for $PNBO$, we also use the normalized $PNBO$ to predict market excess returns. Table 4 shows that, like $PNBO$, $PNBON$ also predicts future market returns negatively. The coefficient estimate b_r remains statistically significant up to 3 months ahead, but with lower R^2 at all horizons than $PNBO$. The difference in R^2 between $PNBON$ and $PNBO$ shows that we should interpret the high R^2 for $PNBO$ with caution, which could partially be due to its volatility trend. As for economic significance, when $b_{VP,t} < 0$, a one-standard deviation decrease in $PNBON$ is associated with a 2.7% increase in the future 3-month market excess return.

Next, Panel B of Table 4 shows that, when the average level of daily jump risk in the current month is above the median level in the full sample,[10] $PNBO$ (and $PNBON$)

[9] All the standard errors for the return-forecasting regressions are based on Hodrick (1992). We provide additional results on statistical inference in the Internet Appendix, including Newey and West (1987) standard errors with long lags, bootstrapped confidence intervals, and the test statistic of Muller (2014). See Ang and Bekaert (2007) for further discussion on long-horizon statistical inference.

[10] Setting $\bar{J}$ to be the median value of J_t of the full sample introduces future information into the

again has significant predictive power for future market excess returns up to 4 months ahead (3 months for $PNBON$). In such months, a one-standard deviation decrease in $PNBO$ ($PNBON$) is associated with a 4.6% (2.6%) increase in the future 3-month market excess return.

Table 4 shows that $PNBO$ demonstrates similar predictive power for future market excess returns during months of negative price-quantity relation in the crash insurance market ($b_{VP,t} < 0$) and during months of high jump risk ($J_t > \bar{J}$). However, the two criteria are in fact almost independent of each other in our sample. Recall that both strategies are designed to identify periods when supply shocks are the main drivers of variations in $PNBO$, and both are bound to be imperfect. To understand how they separately contribute to the predictive power of $PNBO$, we further split the full sample into 4 sub-sample periods based on the two criteria ($b_{VP,t} < (\geq) 0$, $J_t < (\geq) \bar{J}$) and re-estimate the return-forecasting regressions.

Among the months with $b_{VP,t} < 0$, 80 have jump risks that are above the median level, 79 below. Among the months with $b_{VP,t} \geq 0$, 52 have above-median jump risk, 53 below. Table 5 shows that for both $PNBO$ and $PNBON$, the predictive power of the option volume measure is concentrated in the periods when $b_{VP,t} < 0$ and is particularly strong when the level of jump risk is high. In the periods when $b_{VP} \geq 0$ but jump risk is high, the coefficient estimate b_r is again significant for $PNBO$, but insignificant for $PNBON$. If $b_{VP} \geq 0$ and jump risk is low, then the return-forecasting coefficient becomes insignificant for both measures and the (adjusted) R^2 drops to near zero.

In summary, the sub-sample results suggest that our strategy based on the price-quantity relation does effectively identify periods when $PNBO$ are more connected to variations in intermediary constraints. The fact that $PNBO$ still has some predictive power (with significant regression coefficient and sizable R^2) in the sub-sample with $b_{VP,t} \geq 0$ but $J_t \geq \bar{J}$ does not contradict with our identification strategy. First, it suggests that the level of aggregate tail risk is also helpful in identifying supply-driven $PNBO$ variations. Second, as we explained in Section 2, $b_{VP,t} < 0$ does not rule out the presence of demand

return-forecasting regression. Our results are robust to changing $\bar{J}$ to only using past information.

shocks in the same period, and $b_{VP,t} > 0$ does not rule out the presence of supply shocks. Thus, while we expect the predictive power of $PNBO$ to be stronger when $b_{VP,t} < 0$ than when $b_{VP,t} > 0$, which is confirmed by Table 5, our identification does not imply that there should be no predictive power for $PNBO$ in the latter case.

To investigate whether the predictability results above are useful in forming real-time forecasts, we follow Welch and Goyal (2008) and compute the out-of-sample R^2 for $PNBO$ and $PNBON$ based on various sample-split dates, starting in January 1996 (implying a minimum estimation period of 5 years) and ending in December 2007 (with a minimum evaluation period of 5 years). We consider the wide range of sample-split dates because recent studies suggest that sample splits themselves can be data-mined (Hansen and Timmermann (2012)). In forming the return forecasts, we first estimate the predictability regression (4a) during the estimation period (from date 1 to t), and then use the estimated coefficients to forecast the 3-month future market excess return for $t+1$. After obtaining all the return forecasts, we then compute the mean squared forecast errors for the predictability model (MSE_A) and the historical mean model (MSE_N) in various evaluation periods that begin at the sample-split dates and end at the end of the full sample. The out-of-sample R^2 is given by

$$\widehat{R}^2 = 1 - \frac{MSE_A}{MSE_N}.$$

Panel A of Figure 3 shows the results. $PNBO$ achieves an out-of-sample R^2 above 10% for all the sample splits and remains above 20% from 2003 onward. Its normalized value $PNBON$ has an an out-of-sample R^2 above 5% for all the sample splits and remains above 10% in the later period. All of the out-of-sample R^2 are significant at the 5% level (1% level since 2000) based on the MSE-F statistic by McCracken (2007).

Panel B of Figure 3 plots the in-sample R^2 from the predictive regressions of $PNBO$ and $PNBON$ using 5-year moving windows. The two R^2s vary significantly over time. They are lower in the early parts of the sample, being less than 5% for the majority of the time prior to 2006. Both R^2 reached 10% in the period around the Asian financial crisis and Russian default in 1997-98, and approached 10% again around 2002. During the financial crisis period, the R^2 rose to close to 40%. These high R^2 for the return-forecasting regressions would translate into striking Sharpe ratios for investment strategies that try to exploit such predictability. For example, Cochrane (1999) shows that the best unconditional Sharpe ratio s^* for a market timing strategy is related to the predictability regression R^2 by

$$s^* = \frac{\sqrt{s_0^2 + R^2}}{\sqrt{1 - R^2}},$$

where s_0 is the unconditional Sharpe ratio of a buy-and-hold strategy. Assuming the Sharpe ratio of the market portfolio is 0.5, then an R^2 of 40% implies a Sharpe ratio for the market timing strategy that exceeds 1. Such high Sharpe ratios could persist during the financial crisis because of the presence of severe financial constraints that prevent arbitrageurs from taking advantage of the investment opportunities.

Theories of intermediary constraints argue that variations in such constraints not only will affect the risk premium of the market portfolio, but also the risk premia on any financial assets for which the financial intermediaries are the marginal investor. Having examined the ability of $PNBO$ to predict future market excess returns, we now apply the predictability regressions to other asset classes.

Among the assets we consider are (1) a high-yield bond portfolio (based on the Barclays U.S. Corporate High Yield total return index), (2) a hedge fund portfolio (based on the HFRI fund-weighted average return index), (3) a carry trade return portfolio (constructed by Lustig, Roussanov, and Verdelhan, 2011, using the exchange rates of 15 developed countries), (4) a commodity portfolio (based on the Goldman Sachs commodity index excess return series), (5) the 10-year US Treasury, and (6) variance swap for the S&P 500 index returns (with the excess return defined as the log ratio of the realized annualized return variance over the swap rate; see e.g., Carr and Wu, 2009).[11]

As Table 6 shows, our option-volume measures predict the future excess returns for a variety of assets. $PNBO$ interacted with the indicator $b_{VP,t} < 0$ predicts negatively and

[11]Data for the returns on high yield bonds, commodity, and hedge funds are from Datastream. Government bond return data are from Global Financial Data.

statistically significantly (at least at the 10% level) the future 3-month returns of high yield bonds, hedge funds, carry trade, and commodity (for $PNBON$, the coefficient b_r for carry trade and commodity returns are still negative but not significant). Thus, like the market portfolio, the risk premia on these assets tend to rise when the intermediary constraints tighten.

Next, both for $PNBO$ and $PNBON$, the predictive coefficient b_r is positive and significant for the 3-month excess returns of the 10-year Treasury and the S&P 500 variance swap. The result on Treasury risk premium is consistent with Fontaine and Garcia (2012), who find the deterioration in funding liquidity predicts lower risk premia for Treasury securities. Intuitively, when financial intermediaries become constrained, Treasury values tend to rise ("flight to quality") as does the volatility in the market. Thus, Treasuries and (pay-fix) variance swaps provide a hedge against negative shocks to intermediary constraints. Our results suggest that their risk premia become lower (or more negative) as the intermediary constraints tighten.

3.4 An alternative hypothesis

The above return predictability results have two alternative interpretations. It is possible that financial intermediaries become more constrained when the market risk premium rises (e.g., due to higher aggregate uncertainty in the real economy), which in turn reduces their capacity to provide market crash insurance to public investors. As a result, a low $PNBO$ today would be associated with high future market returns, even though a tighter intermediary constraint does not *cause* the market risk premium to rise in this case. Alternatively, it is possible that intermediary constraints directly affect the aggregate market risk premium, which is a central prediction in intermediary asset pricing theories.

To distinguish between these two interpretations, we compare $PNBO$ against a number of financial and macro variables that have been shown to predict market returns. If $PNBO$ is merely correlated with the standard risk factors and does not directly affect the risk premium, then the inclusion of the proper risk factors into the predictability regression should drive away the predictive power of $PNBO$. The variables we consider include the variance risk premium measure (IVRV) in Bollerslev, Tauchen, and Zhou (2009),[12] the log dividend yield ($d - p$) of the market portfolio, the log net payout yield (lcrspnpy) by Boudoukh, Michaely, Richardson, and Roberts (2007), the Baa-Aaa credit spread (DEF), the 10-year minus 3-month Treasury term spread (TERM), the tail risk measure (Tail) by Kelly (2012), the slope of the implied volatility curve (IVSlope), and the consumption-wealth ratio measure ($\widehat{cay}$) by Lettau and Ludvigson (2001). All the variables are available monthly except for $\widehat{cay}$, which is at quarterly frequency.

Table 7 shows that, with the inclusion of the various competing variables, the predictive coefficients for $PNBO$ and $PNBON$ remain significantly negative and remarkably stable in magnitude. Comparing the R^2 from the bivariate regressions in Table 7 and the univariate regressions in Table 4, we see that the incremental explanatory power for future market excess returns mostly comes from $PNBO$ (interacted with the price-quantity relation measure).

In theory, variation in risk premium due to intermediary constraints should affect variables such as the dividend-price ratio. Why are $d - p$ and lcrspnpy not showing significant predictive power in Table 7? There are two reasons. First, the dividend-price ratio (and payout yield) is affected by both the variations in discount rates and expected dividend growth. Since the statistical evidence for the predictive power of PD is not that strong to begin with (see e.g., Welch and Goyal, 2008; Cochrane, 2008), it is not surprising that it becomes insignificant once another variable that is informative about the variation in risk premium is included into the regression.

Second, transitory fluctuations in the discount rate have limited effects on prices. Instead, the dividend-price ratio should be more sensitive to changes in longer-term discount rates. We find that the predictive power of $PNBO$ lasts only for up to 4 months into the future, which is consistent with the intuition that episodes of tight intermediary constraints are likely to have temporary effects on the risk premium. In comparison, evidence on the predictive power for dividend-price ratio and related measures in the literature are typically at longer horizons.

[12] Using the variance premium (VP) in Bekaert and Hoerova (2014) instead of IVRV generates qualitatively the same results.

In summary, the results from Table 7 show that the option trading activities of public investors and financial intermediaries contain unique information about the market risk premium that is not captured by the standard macro and financial factors. This result is consistent with the theories of intermediary constraints driving asset prices. Of course, the evidence above does not prove that intermediary constraints actually drive aggregate risk premia. It is possible that $PNBO$ is correlated with other risk factors not considered in our specifications.

3.5 Option volume and measures of funding constraints

Recently, several measures of funding constraints for financial intermediaries have been proposed in the literature. They include the balance sheet growth measures advocated by Adrian, Moench, and Shin (2010) (Δlev), the fixed-income market based funding liquidity measures by Fontaine and Garcia (2012) (FG) and Hu, Pan, and Wang (2013) (Noise), the CBOE VIX index (VIX), the TED spread (TED, the difference between the 3-month LIBOR and the 3-month T-bill rate), and the LIBOR-OIS spread (LIBOR-OIS, the difference between the 3-month LIBOR and the 3-month overnight indexed swap rate). While VIX reflects the volatility of the stock market, TED spread and LIBOR-OIS spread measure the credit risk of banks. We now compare our option volume-based measure with these measures of funding constraints.

We first run OLS regressions of our constraint measure on the funding constraint measures in the literature. As Panel A of Table 8 shows, $I_{b_{VP,t}<0} \times PNBO_t$ is significantly *positively* related to the TED spread. This positive relation is mainly due to the fact that $PNBO$ rose significantly along with the TED spread during the early part of the financial crisis. Subsequently, while $PNBO$ turned significantly negative, the TED spread fell to and remained at low levels. This result points out a potential weakness of the TED spread (and LIBOR-OIS) as a measure of funding constraint. The TED spread could become lower because of the cautionary measures banks take to reduce their risk exposures, which includes aggressive deleveraging and buying crash insurance, but they do not necessarily imply that banks are no longer constrained (it could in fact be just the opposite).

Next, Our constraint measure is significantly negatively related to the measure FG, which suggests that financial intermediaries are more constrained in periods when the value of funding liquidity rises according to FG. In the quarterly regression, $I_{b_{VP,t}<0} \times PNBO_t$ is significantly positively related to the growth rate in broker-dealer leverage Δlev, with $\bar{R}^2$ of 7.8%. That is, intermediary constraint is tight while broker-dealers are de-leveraging.

In Panel B of Table 8, we further examine the ability of the various funding constraint measures to predict aggregate market returns. Adrian, Moench, and Shin (2010) show that the year-over-year change in broker-dealer leverage (Δlev) has strong predictive power for excess returns on stocks, corporate bonds, and treasuries. In a univariate regression (unreported), Δlev indeed predicts future market excess returns with a significant negative coefficient in our sample period. However, in a bivariate regression with $PNBO$, the coefficient on Δlev becomes insignificant, while that on $PNBO$ remains significant. The $\bar{R}^2$ of the bivariate regression is only slightly higher than that in the univariate regression for $PNBO$. Similarly, when the other funding constraint measures are used in place of Δlev, the coefficient on $PNBO$ is essentially unaffected. These results suggest that relative to other funding constraint measures, our intermediary constraint measure contains unique information about future market risk premia.

3.6 Public investors: retail vs. institutional

As Figure 1 shows, while financial intermediaries typically sell DOTM SPX puts to public investors during normal times, the roles are often reversed during crisis times, most notably during the 2008-09 financial crisis. To understand the risk sharing mechanism between financial intermediaries and public investors, it is important to find out who among the public investors are reducing the amount of crash insurance acquired from the intermediaries or even starting to provide the insurance to the intermediaries when the latter become constrained. The SPX volume data from CBOE do not provide further information about the types of public investors behind a given transaction (e.g., retail vs. institutional investors). We tackle this question by comparing the trading activities of the public investors in SPX options with those in SPY options.

While SPX and SPY options have essentially identical underlying asset, it is well known among practitioners that institutional investors account for a significantly higher percentage of the trading volume of SPX options than do SPY options. Compared to retail investors, institutional investors prefer SPX options more due to a larger contract size (10 times as large as SPY), cash settlement, more favorable tax treatment, as well as being more capable of trading in between the relatively wide bid-ask spreads of SPX options due to the stronger bargaining power. As in SPX options, we construct $PNBO_{SPY}$ for SPY options. Our SPY options volume data are from the CBOE and the International Securities Exchange (ISE), and cover the period from 2005/05 to 2012/12. Unlike SPX options which trade exclusively on the CBOE, SPY options are cross-listed at several option exchanges. Our $PNBO_{SPY}$ aggregates the volume data from the CBOE and ISE, which account for about half of the total trading volume for SPY options.

Figure 6 compares the $PNBO_{SPY}$ and $PNBO_{SPX}$ (equivalent to $PNBO$) series. During the period of 2005/05 to 2012/01, $PNBO_{SPY}$ is positive in the majority of the months. From 2008/09 to 2010/12, $PNBO_{SPX}$ is negative in 22 out of 28 months, whereas $PNBO_{SPY}$ is negative in just 7 of the 28 months.

A systematic way to examine the difference in how the equilibrium quantities of trading in the two markets are connected to the intermediary constraints and market risk premium is through the return-forecasting regressions. As Table 9 shows, the predictive coefficient for $PNBO_{SPY}$ is insignificant in the period of 2005/05 to 2012/12. In the same period, the coefficient on $PNBO_{SPX}$ is significantly negative, and the R^2 is much larger. The fact that $PNBO_{SPY}$ appears unrelated to variations in market risk premium suggests that retail investors and institutional investors behave differently as the degree of intermediary constraint changes. In particular, when constrained financial intermediaries start buying crash insurance, they appear to be buying the insurance from the public (institutional) investors in the SPX market and not from the public (retail) investors in the SPY market.

4 Robustness Checks

In this section, we report the results of several robustness checks for our main results.

4.1 Financial crisis and outliers

One potential concern regarding the predictive power of $PNBO$ is that it might be driven by a small number of outliers, in particular those observations during the 2008-09 financial crisis. To address this concern, we re-estimate the predictive regressions for 3-month market excess returns after removing a certain number of the most extreme observations of $PNBO$ and $PNBON$ (in terms of the absolute value). Figure 4 shows the 95% confidence intervals of the b_r coefficient after the removal of the extreme observations. The predictive coefficients have relatively stable point estimates and remain significantly negative even after deleting up to 50 extreme observations (19% of the sample).

Table 10 reports the results of the return-forecasting regressions in two sub-samples: pre-crisis (1991/01-2007/11) and post-crisis (2009/06-2012/12). The predictive powers of $PNBO$ and $PNBON$ remain statistically significant in both sub-samples. The economic significance of the predictive power is weaker in the pre-crisis period than in the full sample (in terms of smaller magnitude of b_r and lower R^2) than the full sample, but it is quite strong in the post-crisis period. Thus, while the relation between $PNBO$ and market risk premium is indeed stronger during the financial crisis, it is not just a crisis phenomenon. The weaker predictive power for $PNBO$ in the earlier sample period could be due to the fact that intermediary constraints are not as significant and volatile in the first half of the sample. It could also be that the SPX options market was less developed in the early periods and did not play as important a role in facilitating risk sharing as it does today.

4.2 Moneyness

Next, we examine how the predictive power of $PNBO$ changes based on option moneyness. Our baseline definition of DOTM puts uses a simple cutoff rule $K/S \leq 0.85$. Panel A of Figure 5 plots the coefficient b_r and the confidence intervals as we change this cutoff value for SPX puts. The coefficient b_r in the return forecast regression is significantly negative for a wide range of moneyness cutoffs. The point estimate of b_r does become more negative as the cutoff becomes smaller. Because DOTM puts are more difficult to hedge than ATM puts, they expose financial intermediaries to more inventory risks. Hence, $PNBO$ measure based on DOTM puts should be more informative about intermediary constraints and in turn the aggregate risk premium than $PNBO$ based on ATM puts. At the same time, because far out-of-money options are less liquid, the $PNBO$ series becomes more noisy when we further reduce the cutoff, which widens the confidence interval on b_r. In contrast, Panel B shows that for essentially all moneyness cutoffs, $PNBO$ based on SPX call options does not predict market returns.

A feature of our definition of DOTM puts above is that a constant strike-to-price cutoff implies different actual moneyness (e.g., as measured by option delta) for options with different maturities. A 15% drop in price over one month might seem very extreme in calm periods, but it is more likely when market volatility is high. For this reason, we also examine a maturity-adjusted moneyness definition. Specifically, we classify a put option as DOTM when $K/S \leq 1 + k\sigma_t\sqrt{T}$, where k is a constant, σ_t is the daily S&P return volatility in the previous 30 trading days, and T is the days to maturity for the option. This is similar to using option delta to define moneyness, but does not require a particular pricing model to compute the delta. Panel C of Figure 5 shows that this alternative classification of DOTM puts produces qualitatively similar results as our simple cutoff rule. Panel D shows again that $PNBO$ based on SPX calls and this alternative moneyness cutoff does not predict returns.

4.3 Volume vs. open interest

In our construction of $PNBO$, we focus on the net amount of new DOTM index puts that public investors buy in a period. An alternative way to gauge the economic exposure for public investors and financial intermediaries is to examine the net open interest for the two groups. Between new volume and open interest, which one better represents the degree of intermediary constraints?

We use the volume-based measure in our main analysis for two reasons. First, when financial intermediaries become constrained, it is easier (e.g., due to transaction costs) to adjust the quantity of new DOTM puts traded than to make significant changes in their established positions. That makes the volume-based measure potentially more sensitive to changes in intermediary constraints than open interest-based measure. Second, taking on an option position that is originally near the money but later becomes DOTM due to market movements is different from taking on a new DOTM option. In the former case, the intermediaries can put on hedges against tail risk over time (via OTC options, VIX options etc.), which means these positions are less sensitive to changes in intermediary constraints.

Nonetheless, we provide two robustness checks. First, we examine an alternative measure based on the end-of-month public net open interest for DOTM SPX puts ($PNOI$). The results are discussed below. Second, we construct a $PNBO$ measure using only one-month options, for which the monthly measures of net volume and open interest are equivalent. We find that the main results hold for the measure based on short-dated options (see Table IA3 and IA4 in the Internet Appendix).

For the period of 1991 to 2001, we use daily long and short open interest provided by CBOE to compute $PNOI$. CBOE stopped providing the open interest data after 2001. Thus, from 2001 onward, we compute daily net open interest from the volume information as follows:

$$NOI_d^{K,T} = NOI_{d-1}^{K,T} + \text{openBuy}_\text{d}^\text{K,T} - \text{openSell}_\text{d}^\text{K,T} + \text{closeBuy}_\text{d}^\text{K,T} - \text{closeSell}_\text{d}^\text{K,T}, \tag{5}$$

where $NOI_d^{K,T}$ is the public investor net open interest of options with strike price K and maturity T on day d, openBuy is the public investor buying volume from initiating long positions, openSell is the selling volume from initiating short, closeBuy is the buying volume from closing existing short positions, and closeSell is the selling volume from closing existing long positions. We then aggregate $NOI_d^{K,T}$ to compute daily net open interest of DOTM puts ($PNOI$). We also consider a normalized version of $PNOI$ ($PNOIN$), which is $PNOI$ divided by the sum of public long and short open interest for all SPX options.

Table 1 shows that the end-of-month $PNOI$ is around 20,000 contracts on average. $PNOI$ has higher auto- correlations than $PNBO$. As Table 11 shows, like $PNBO$, $PNOI$ predicts future market excess returns negatively, with significant predictive coefficient up to 3 months in the future. The R^2 of the regressions are also similar to $PNBO$. Furthermore, Table 12 shows that the sub-sample predictive power of $PNOI$ is also similar to $PNBO$, which is the strongest in periods when the price-quantity relation is negative ($b_{VP,t} < 0$) and when market jump risk is high ($J_t \geq \bar{J}$).

5 Conclusion

We provide evidence that the trading activities of financial intermediaries in the market of DOTM SPX put options are informative about the degree of intermediary constraints. In periods when supply shocks are likely to be the main force behind the variations in the price-quantity relation in the DOTM SPX put market, our public investor net-buying volume measure, $PNBO$, has strong predictive power for future market excess returns and the returns for a range of other financial assets. The predictive power of $PNBO$ is stronger during periods when the market jump risk is high, and it is stronger for DOTM puts. $PNBO$ is also associated with several funding liquidity measures in the literature. Moreover, the information that $PNBO$ contains about the market risk premium is not captured by the standard financial and macro variables. These results suggest that time-varying intermediary constraints are driving the supply of crash insurance by financial intermediaries and the risk premia in financial markets.

To explain these findings, we build a general equilibrium model of the crash insurance market, which captures the time-varying intermediary constraints in reduced form. For future work, further investigation of the sources of the variation in intermediary constraints, both theoretically and empirically, will help us better understand how intermediary constraints affect the financial markets.

Appendix

A A Dynamic Model

In Section 3, we present empirical evidence connecting the trading activities in the market for DOTM SPX puts to option pricing, market risk premium, and various measures of intermediary constraints. In particular, there is time-varying equilibrium demand for crash insurance from public investors. The equilibrium demand is inversely related to the relative price of DOTM put options—times in which the equilibrium demand is low are generally times when the protection is very expensive. The demand for crash insurance is also informative about future stock market returns over and above the information in standard macro and financial variables. Finally, the demand for crash insurance is positively related to the changes in broker-dealer leverage. We now examine an equilibrium model consistent with these empirical facts.

A.1 Model setup

We consider an aggregate endowment in the economy which follows a jump diffusion process where the endowment is subject to both a diffusive risk and a jump risk. In particular, sudden severe drops in the aggregate endowment are a source of disaster risk in this economy. There are two types of agents in the economy: small public investors and competitive dealers. We assume there exists a representative public investor, who is denoted by agent P, and a representative dealer, denoted by agent D. To induce the two types of agents to trade, we assume that they have different beliefs about the probability of disasters. Such differences in beliefs capture in reduced form the advantages that dealers have in bearing disaster risk, whether it is due to differences in technology, agency problems, or behavioral biases.[13]

Specifically, we assume that both agents believe that the log aggregate endowment,

[13] Examples include government guarantees to large financial institutions and compensation schemes that encourages managers to take on tail risk. See e.g., Lo (2001) and Malliaris and Yan (2010).

$c_t = \log C_t$, follows the process

$$dc_t = \bar{g}dt + \sigma_c dW_t^c - \bar{d}\, dN_t, \tag{A1}$$

where $\bar{g}$ and σ_c are the expected growth rate and volatility of consumption without jumps, W_t^c is a standard Brownian motion under both agents' beliefs, and $\bar{d}$ is the constant size of consumption drop in a diaster[14]. N_t is a counting process whose jumps arrive with stochastic intensity λ_t under the public investors' beliefs, and λ_t follows

$$d\lambda_t = \kappa(\bar{\lambda} - \lambda_t)dt + \sigma_\lambda \sqrt{\lambda_t} dW_t^\lambda, \tag{A2}$$

where $\bar{\lambda}$ is the long-run average jump intensity under P's beliefs, and W_t^λ is a standard Brownian motion independent of W_t^c. In general, the dealers are more willing to bear the disaster risk because (they act as if) they are more optimistic about disaster risk. We assume that they believe that the disaster intensity is given by $\rho\lambda_t$ with $\rho < 1$. We summarize the public investors' beliefs with the probability measure $\mathbb{P}_P$, and the dealers' beliefs with the probability measure $\mathbb{P}_D$.

Public investors have standard constant relative risk aversion (CRRA) utility:

$$U^P = E_0^P \left[\int_0^\infty e^{-\delta t} \frac{C_{P,t}^{1-\gamma}}{1-\gamma} dt \right], \tag{A3}$$

where we focus on the cases where $\gamma > 1$. The superscript P reflects that the expectations are taken under the public investors' beliefs.

The utility function of the dealers are different. We assume that the dealers face an intermediation constraint that we model in a reduced form directly in terms of their utility. Specifically, we suppose that

$$U^D = E_0^D \left[\int_0^\infty e^{-\delta t} \frac{C_{D,t}^{1-\gamma}}{1-\gamma} e^{-\sum_{n=1}^{N_t} (\alpha_{\tau(n)} - \bar{\alpha})} dt \right], \tag{A4}$$

[14] As in Chen, Joslin, and Tran (2012), one could generalize the model by allowing disaster size to have a time-invariant distribution.

where α_t is a stochastic variable representing the ability of the dealer to intermediate disaster risk. Limited ability to intermediate risk is modeled as increased risk aversion against market crashes. This specification generalizes the state-dependent preferences proposed by Bates (2008) in that it allows the dealers' risk aversion against crashes to rise with the probability of disasters.

Specifically, $\tau(n)$ is the time of the n$^{\text{th}}$ disaster since $t = 0$, $\tau(n) \equiv \inf\{s : N_s = n\}$ Thus, this crash-aversion term remains constant in between disasters. Suppose the dealer's log consumption drops by $d_{D,\tau(n)}$ at the time of the n$^{\text{th}}$ disaster. Then, at the same time, the marginal utility of the dealer jumps up by

$$e^{\gamma d_{D,\tau(n)} - (\alpha_{\tau(n)} - \bar{\alpha})} = e^{\left(\gamma - \frac{\alpha_{\tau(n)} - \bar{\alpha}}{d_{D,\tau(n)}}\right) d_{D,\tau(n)}}, \tag{A5}$$

which implies that the dealer's effective relative risk aversion against the disaster is

$$\gamma_{D,\tau(n)} = \gamma - \frac{\alpha_{\tau(n)} - \bar{\alpha}}{d_{D,\tau(n)}}. \tag{A6}$$

Thus, when $\alpha_t > \bar{\alpha}$, the dealer will have lower aversion to disaster risk than the public investor. As α_t falls, the dealer's effective risk aversion rises.

The intermediation capacity of the dealer may be related to the disaster intensity. We model the intermediation as being driven jointly by the disaster intensity, λ_t, and an independent factor, x_t, so that $\alpha_t = -a\lambda_t + bx_t$. Thus when $a > 0$, the intermediation capacity goes down as the intensity rises and the dealer becomes more averse to disaster risk.

Any jointly affine process for (c_t, λ_t, x_t) would be suitable for a tractable specification. For example, we could suppose that x_t follows an independent CIR process:

$$dx_t = \kappa_x(\bar{x} - x_t)dt + \sigma_x\sqrt{x_t}dW_t^x. \tag{A7}$$

In our calibrations, we will choose the simple specification with $b = 0$ so that the intermediation capacity is perfectly correlated with the disaster intensity.

The main motivation for the dealer's time-varying aversion to crash risk is the time-varying constraint faced by financial intermediaries. Rising crash risk in the economy raises the intermediaries' capital/collateral requirements and tightens their constraints on tail risk exposures (e.g., Value-at-Risk constraints), which make them more reluctant to provide insurance against disaster risk. For example, see Adrian and Shin (2010) and He and Krishnamurhty (2012). In this sense, the shocks to the disaster intensity in the model also serve the purpose of generating time variation in the intermediation capacity of the dealer. We can further generalize the specification by making the dealer's aversion to crash risk driven by adding independent variations in the intermediation shocks.

We also assume that markets are complete and agents are endowed with some fixed share of aggregate consumption ($\theta_P, \theta_D = 1 - \theta_P$). The equilibrium allocations can be characterized as the solution of the following planner's problem, specified under the probability measure $\mathbb{P}_P$,

$$\max_{C_t^P, C_t^P} E_0^P \left[\int_0^\infty e^{-\delta t} \frac{(C_t^P)^{1-\gamma}}{1-\gamma} + \zeta \eta_t e^{-\delta t} \frac{(C_t^D)^{1-\gamma} e^{a \sum_{n=1}^{N_t} (\lambda_{\tau(n)} - \bar{\lambda})}}{1-\gamma} dt \right], \tag{A8}$$

subject to the resource constraint $C_t^P + C_t^D = C_t$. Here, ζ is the the Pareto weight for the dealers and

$$\eta_t \equiv \frac{d\mathbb{P}_D}{d\mathbb{P}_P} = \rho^{N_t} e^{(1-\rho) \int_0^t \lambda_s ds}. \tag{A9}$$

where $\rho = \bar{\lambda}_D / \lambda$, the relative likelihood of a jump under the two beliefs. From the first order condition and the resource constraint, we obtain the equilibrium consumption allocations $C_t^P = f^P(\tilde{\zeta}_t) C_t$ and $C_t^D = (1 - f^P(\tilde{\zeta}_t)) C_t$, where

$$\tilde{\zeta}_t = \rho_t^N e^{(1-\rho) \int_0^t \lambda_s ds + \alpha \sum_{n=1}^{N_t} (\lambda_{\tau(n)} - \bar{\lambda})} \zeta \tag{A10}$$

and

$$f^P(\tilde{\zeta}) = \frac{1}{1 + \tilde{\zeta}^{\frac{1}{\gamma}}}. \tag{A11}$$

The stochastic discount factor under P's beliefs, M_t^P, is given by

$$M_t^P = e^{-\rho t}(C_t^P)^{-\gamma} = e^{-\delta t} f^P(\tilde{\zeta}_t)^{-\gamma} C_t^{-\gamma}. \tag{A12}$$

We can solve for the Pareto weight ζ through the lifetime budget constraint for one of the agents (Cox and Huang (1989)), which is linked to the initial allocation of endowments.

Our key focus will be on risk premiums and on the net public purchase of crash insurance which we relate to the market for deep out of the money puts in our empirical analysis. The risk premium for any security under each agent's beliefs is the difference between the expected return under $\mathbb{P}_i$ and under the risk-neutral measure $\mathbb{Q}$.

$$E_t^i[R^e] = \gamma \sigma_c \partial_B P + (\lambda_t^i - \lambda_t^{\mathbb{Q}}) E_t^d[R], \quad i = D, P, \tag{A13}$$

where we use the shorthand that $\partial_B P$ denotes the sensitivity of the security to Brownian shocks and $E_t^d[R]$ is the expected return of the security *conditional on a disaster*. Since consumption will be relatively smooth in our calibration, the return of securities which are not highly levered on the Brownian risk will be dominated by the jump risk term. Moreover, agents agree about the Brownian risk and have the same risk aversion with respect to these shocks so there will be no variation in the Sharpe ratio for Brownian risk. In light of these facts, we focus on the variation in the jump risk premium, as measured by $\lambda^{\mathbb{Q}}/\lambda_P^{\mathbb{P}}$.

The stochastic discount factor characterizes the unique risk neutral probability measure $\mathbb{Q}$ (see, e.g., Duffie 2001). The risk-neutral disaster intensity, $\lambda_t^{\mathbb{Q}} \equiv E_t^d[M_t^i]/M_t^i \lambda_t^i$, is determined by the expected jump size of the stochastic discount factor at the time of a disaster. When the risk-free rate and disaster intensity are close to zero, the risk-neutral disaster intensity has the nice interpretation of (approximately) the value of a one-year crash insurance contract that pays one at $t+1$ when a disaster occurs between t and $t+1$.

In our setting, the risk-neutral jump intensity is given by

$$\lambda_t^{\mathbb{Q}} = e^{\gamma \bar{d}} \frac{\big(1 + (\rho \tilde{\zeta}_t)^{\frac{1}{\gamma}}\big)^{\gamma}}{(1 + \tilde{\zeta}_t^{\frac{1}{\gamma}})^{\gamma}} \lambda_t. \tag{A14}$$

In order to define the market size, we must consider how the Pareto efficient allocation is obtained. The equilibrium allocations can be implemented through competitive trading in a sequential-trade economy. Extending the analysis of Bates (2008), we can consider four types of traded securities: (i) a risk-free money market account, (ii) a claim to aggregate consumption, and (iii) a crash insurance contracts which pays one dollar in the event of a disaster in exchange for a continuous premium, and (iv) a separate instrument sensitive only to shocks in the disaster intensity. As in Chen, Joslin, and Tran (2012), since agents agree about the Brownian risk and have identical aversion to the risk, they will proportionally hold the Brownian risk according to their consumption share. With the instruments we have specified, this means they will proportionally hold the consumption claim. Thus, the agents will hold proportional exposure to the disaster risk from their exposure to the consumption claim. Motivated by these facts, we define the net public purchase for crash insurance as the (scaled) difference between the consumption loss the public bears in equilibrium minus the consumption loss that the public would bear without insurance. That is, the public purchase for insurance is the difference between $e^{-\bar{d}}(f^P(\tilde{\zeta}_t^d) - f^P(\tilde{\zeta}_{t-}))$ (where $\tilde{\zeta}_t^d$ is the value of $\tilde{\zeta}_t$ conditional on a disaster occurring at time t: $\tilde{\zeta}_t^d = \rho e^{\alpha(\lambda_{t-} - \bar{\lambda})} \tilde{\zeta}_{t-}$) and $e^{-\bar{d}}(f^P(\tilde{\zeta}_t^d) - f^P(\tilde{\zeta}_{t-})$. Thus we define the net public purchase for insurance as

$$\text{net public purchase for crash insurance} = e^{-\bar{d}}\big(f^P(\tilde{\zeta}_t \rho e^{\alpha(\lambda_{t-} - \bar{\lambda})}) - f^P(\tilde{\zeta}_t)\big). \tag{A15}$$

A.2 Net public purchase and risk premia in the dynamic model

We now study the relationship between public purchase and risk premia in the context of our dynamic model. We calibrate our model as in Chen, Joslin, and Tran (2012) and Wachter (2012). The key new parameter that we introduce is the time-variation in aversion to jump risk. We parameterize this by setting $a = \alpha\bar{d}/\sigma_{SS}(\lambda)$, where $\sigma_{SS}(\lambda)$ is the volatility of the stationary distribution of λ. We choose $\alpha = 1$, which together with the other parameters implies that when $\lambda = 2.35\%$ (one standard deviation from the steady state volatility (65 bp) above the long run mean (1.7%)), an economy populated by only the dealer will behave as an economy with a representative agent with relative risk aversion of 5 with respect to jumps (one higher than if he had standard CRRA utility with $\gamma = 4$). As a baseline comparison, we also consider the parameterization with $\alpha = 0$, which corresponds to the stochastic intensity model of Chen, Joslin, and Tran (2012).

In our model, there are two state variable: λ_t, the likelihood of a disaster, and f_t, the public investors consumption share. Figure A1 plots the net public purchase of crash insurance as a function of the public consumption share and the jump intensity. When $\alpha = 0$ (panel B), the amount of risk sharing does not depend on λ as the motivation to share risk depends only on the size of the jump in consumption. The equilibrium public purchase is close to zero when either the public has a low consumption share (the public has limited resources to buy insurance) or when the public has high consumption share (the dealers have limited ability to provide insurance) with a peak in the middle where the public and dealers share a lot of risk. In contrasts, Panel B shows that when the dealers have time-varying aversion to jump risk, the relationship is much more complex. For low levels of the intensity the patterns is the same as before since the dealers are as willing (or even more willing) to sell crash insurance. However, as λ rises, the dealer becomes more averse to jump risk and is less willing to provide insurance. When the intensity becomes high enough, the dealers become so averse to jump risk that they even begin to become buyers of insurance rather than sellers.

Figure A2 plots the jump risk premium, as measured by $\lambda^{\mathbb{Q}}/\lambda$, as a function of the public consumption share and the jump intensity. In the case of $\alpha = 0$, the jump risk premium rises as there are fewer dealers to hold the jump risk. When $\alpha = 1$ and λ is low, the jump risk premium falls as the dealer's consumption share increases. When λ is high enough, this relationship reverses and the premium rises as the dealer gains consumption share. The reason for this is that as λ rises, the dealer becomes more risk averse and eventually demands a higher premium than the public; this relation can be seen be following the curve with $f_t = 0$, corresponding to the case where there is only the dealer.

Next, we examine the relation between the net public purchase of crash insurance and the disaster risk premium in equilibrium. We do so by first holding constant the consumption share of the public investors (f_t) while letting the disaster intensity (λ_t) vary over time. The results are in Figure A3. When $\alpha = 1$, for each of the consumption shares considered ($f_t = 0.9, 0.8, 0.5$), the model predicts a negative relation between the net public purchase for crash insurance and risk premium. This negative relation is consistent with our empirical finding of $b_{VP} < 0$ in Equation (3) and $b_r < 0$ in Equation (4a).

In contrast, when $\alpha = 0$, i.e., when the dealer has constant risk aversion for disaster risk, both the net public purchase and the disaster risk premium remain constant as the disaster intensity changes. This comparison again highlights the key role played by the dealer's time-varying aversion to disaster risk in our model.

When we fix the disaster intensity and let the consumption share vary over time, the relation between the net public purchase of crash insurance and the disaster risk premium is no longer monotonic. Consider first the case with $\alpha = 0$, i.e., the case where the dealer has constant relative risk aversion (see Panel B of Figure A4). In this case, regardless of the disaster intensity, there is a unique relation between the two quantities: as consumption share of the public investors changes from 0 to 1, the net public purchase, as defined in (A15), starts at 0, reaches its peak at 11%, and then falls back to 0 eventually. The limiting case where the public investor own all the aggregate endowment is marked by the red circle on the y-axis. In this process, because the relative amount of risk sharing by the public investor falls as he gains a larger share of consumption, the disaster risk premium in equilibrium rises monotonically until it reaches the limit of $e^{\gamma \bar{d}} = 7.7$.

Consider now the case where the dealer has time-varying risk aversion ($\alpha = 1$). When $\alpha = 1$, (Panel A of Figure A4) the relation between the net public purchase of crash insurance and the disaster risk premium is qualitatively similar to that in Panel B (where $\alpha = 0$) when the disaster intensity is not too high. The equilibrium where public investors own all the endowment is still identical regardless of the level of disaster intensity (again marked by the red circle), but the other extreme where the dealer owns all the endowment has different risk premiums for different disaster intensities. This result is simply due to the dealer's time-varying risk aversion. When λ_t is sufficiently high, the dealer can become so averse to disaster risk that, despite his optimist beliefs about the chances of disasters, the dealer still demands a higher premium than the public investors would. For this reason, when λ_t is sufficiently high, the net public purchase of crash insurance turns negative, i.e., the public investor is now insuring the dealer against disasters, and the disaster risk premium in the economy exceeds the highest level in the case with constant risk aversion.

Panel A of Figure A4 also illustrate that, as the level of disaster intensity λ_t rises, the relation between the disaster risk premium and the net public purchase of crash insurance becomes more negative in the region near the red circle. That is, when public investors own the majority of the wealth in the economy, changes in the public demand for crash insurance are accompanied by larger swings in aggregate risk premium when the level of crash risk is higher. This prediction is confirmed by our empirical finding of $c_{VP} < 0$ in Equation (3).

Figure A5 examines the relationship between the dealer leverage and the consumption share of the public investor and the disaster intensity. Here, we measure the leverage of the dealers as the ratio of the equilibrium consumption loss of the dealers (which includes losses from selling disaster insurance and their equity position) to their consumption loss they would bear without selling any insurance. When the disaster probability is low, so that the dealers are relatively unconstrained and aggressively sell insurance, the dealer leverage increases to a maximum of around two when public investors own the majority of the wealth. For a given level of disaster probability, the dealer leverage decreases as public investors' wealth (consumption) share becomes smaller. As the disaster probability rises, the dealers become more constrained and begin to deleverage. In some cases, the dealer leverage as we define may even fall below one, indicating that the dealers are acquiring insurance to reduce their exposures to disaster risks.

Taken together, the results from Figure A1 and Figure A5 show that our model is able to capture the negative relation between intermediary leverage growth and the net quantity of trading in disaster insurance in the data, as documented in Table 8.

B Comparison with GPP

Our results on the price-quantity relation in the DOTM SPX puts market is related to Garleanu, Pedersen, and Poteshman (2009). However, our results differ from GPP in several aspects. First, we have a longer sample period, from 1991 to 2012, while theirs is from 1996 to 2001. Second, our $PNBO$ measure uses contemporaneous public net-buying volume, while GPP use public net open interest,[15] which is the accumulation of past net-buying volumes. Third, our $PNBO$ measure focuses on DOTM SPX puts, whereas GPP use options of all moneyness (which is similar to the sum of $PNBO$ and $PNBO_{ND}$ in this regard). In this section, we first replicate the main results of GPP (in Table 2, p4287), and then examine the differences between the two studies.

The dependent variable in Table A1, option expensiveness, is the same as the one used in GPP Table 2, i.e., the average implied volatility of ATM options minus a reference model-implied volatility used in Bates (2006).[16] GPP regress the option expensiveness on several measures of SPX non-market-maker demand pressure, including the equal-weighted public open interest (NetDemand), and public open interest weighted by jump risk (JumpRisk). They find the regression coefficients to be positive in the period from 1997/10 to 2001/12.

In Table A1, we obtain very similar results to GPP for the two sub-samples 1996/01-1996/10 and 1997/10-2001/12 for the open interest-based measures. We also construct two net volume-based measures and again find that they are positively related to option expensiveness in the period from 1997/10 to 2001/12.

Next, for the period 2002-2012, we use daily volume data to extend the net open interest measures (constructed using the procedure described in Section 4.3). In this

[15]GPP aggregate the net demand of both public investors and firm investors, and refer to it as the non-market-maker demand.

[16]We thank David Bates for sharing the data on this measure.

subsample, We find that the coefficients on both the open interest and volume-based measures become negative and statistically significant. This finding is consistent with our finding of a negative price-quantity relation in the full sample (see Table 3). The changing signs of the price-quantity relation in different sub-samples suggest that the effects of demand shocks and supply shocks are both present in the SPX option market.

C Additional Robustness Checks

Additional empirical results and robustness checks are provided in the Internet Appendix. They include (1) a table of correlations between $PNBO$ and various macroeconomic and financial variables (Table IA1); (2) a systematic analysis of the statistical significance for the return-forecasting regressions using $PBNO$ and $PNBON$ (Table IA2); (3) return-forecasting regression with PNB (and $PNBN$), which is the public net buy volume for DOTM SPX puts including both open and close transactions (Table IA3); (4) return-forecasting regression with $PNBO_{1month}$ (and $PNBON_{1month}$), which is $PNBO$ constructed using only options with one month or less to maturity (Table IA4); (5) sub-sample return-forecasting regressions with the log dividend-price ratio (Table IA5).

References

Adrian, Tobias, and Nina Boyarchenko, 2012, Intermediary leverage cycles and financial stability, Staff reports, Federal Reserve Bank of New York.

Adrian, Tobias, Erkko Etula, and Tyler Muir, 2012, Financial intermediaries and the cross-section of asset returns, Working Paper.

Adrian, Tobias, Emanuel Moench, and Hyun Song Shin, 2010, Financial intermediation, asset prices, and macroeconomic dynamics, Staff Reports 422 Federal Reserve Bank of New York.

Adrian, Tobias, and Hyun Song Shin, 2010, Liquidity and leverage, *Journal of Financial Intermediation* 19, 418–437.

Andersen, T. G., T. Bollerslev, and F. X. Diebold, 2007, Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility, *Review of Economics and Statistics* 89, 701–720.

Ang, Andrew, and Geert Bekaert, 2007, Stock return predictability: Is it there?, *Review of Financial Studies* 20, 651–707.

Bates, David S., 2006, Maximum Likelihood Estimation of Latent Affine Processes, *Review of Financial Studies* 19, 909–965.

——— , 2008, The market for crash risk, *Journal of Economic Dynamics and Control* 32.

Bekaert, Geert, and Marie Hoerova, 2014, The VIX, the variance premium and stock market volatility, *Journal of Econometrics* 183, 181–192.

Bernanke, Ben, and Mark Gertler, 1989, Agency Costs, Net Worth, and Business Fluctuations, *American Economic Review* 79, 14–31.

Bernanke, Ben S., Mark Gertler, and Simon Gilchrist, 1999, The financial accelerator in a quantitative business cycle framework, in J. B. Taylor, and M. Woodford, ed.: *Handbook* *of Macroeconomics*, vol. 1 of *Handbook of Macroeconomics* . chap. 21, pp. 1341–1393 (Elsevier).

Bollen, Nicolas, and Robert Whaley, 2004, Does net buying pressure affect the shape of implied volatility functions?, *Journal of Finance* 59, 711–753.

Bollerslev, Tim, George Tauchen, and Hao Zhou, 2009, Expected stock returns and variance risk premia, *Review of Financial Studies* 22, 4463–4492.

Boudoukh, Jacob, Roni Michaely, Matthew Richardson, and Michael R. Roberts, 2007, On the Importance of Measuring Payout Yield: Implications for Empirical Asset Pricing, *Journal of Finance* 62, 877–915.

Brunnermeier, Markus K., and Lasse Heje Pedersen, 2009, Market liquidity and funding liquidity, *Review of Financial Studies* 22, 2201–2238.

Brunnermeier, Markus K., and Yuliy Sannikov, 2013, A macroeconomic model with a financial sector, Working Paper, Princeton University.

Buraschi, Andrea, and Alexei Jiltsov, 2006, Model uncertainty and option markets with heterogeneous beliefs, *Journal of Finance* 61, 2841–2897.

Carr, Peter, and Liuren Wu, 2009, Variance risk premiums, *Review of Financial Studies* 22, 1311–1341.

Chen, Hui, Scott Joslin, and Ngoc-Khanh Tran, 2012, Rare disasters and risk sharing with heterogeneous beliefs, *Review of Financial Studies* 25, 2189–2224.

Cochrane, John H., 1999, Portfolio advice of a multifactor world, *Economic Perspectives* pp. 59–78.

——— , 2008, The Dog That Did Not Bark: A Defense of Return Predictability, *Review of Financial Studies* 21, 1533–1575.

Cohen, Lauren, Karl B. Diether, and Christopher J. Malloy, 2007, Supply and Demand Shifts in the Shorting Market, *Journal of Finance* 62, 2061–2096.

Cox, John C., and Chifu Huang, 1989, Optimum consumption and portfolio policies when asset prices follow a diffusion process, *Journal of Economic Theory* 49, 33–83.

Fontaine, Jean-Sebastien, and Rene Garcia, 2012, Bond Liquidity Premia, *Review of Financial Studies* 25, 1207–1254.

Garleanu, Nicolae, Lasse Heje Pedersen, and Allen M. Poteshman, 2009, Demand-based option pricing, *Review of Financial Studies* 22, 4259–4299.

Geanakoplos, John, 2009, The leverage cycle, Cowles Foundation Discussion Papers 1715 Cowles Foundation for Research in Economics, Yale University.

Gromb, Denis, and Dimitri Vayanos, 2002, Equilibrium and welfare in markets with financially constrained arbitrageurs, *Journal of Financial Economics* 66, 361–407.

Hansen, P.R., and A. Timmermann, 2012, Choice of sample split in out-of-sample forecast evaluation, Working Paper, Stanford University.

He, Zhiguo, and Arvind Krishnamurhty, 2012, Intermediary asset pricing, Forthcoming, American Economic Review.

Hodrick, RJ, 1992, Dividend yields and expected stock returns: alternative procedures for inference and measurement, *Review of Financial Studies* 5, 357–386.

Hu, Grace Xing, Jun Pan, and Jiang Wang, 2013, Noise as Information for Illiquidity, *Journal of Finance* 68, 2341–2382.

Johnson, Timothy C., Mo Liang, and Yun Liu, 2014, What drives index options exposure?, Working Paper.

Kelly, Bryan, 2012, Tail risk and asset prices, Working Paper, University of Chicago Booth School of Business.

Kiyotaki, Nobuhiro, and John Moore, 1997, Credit Cycles, *Journal of Political Economy* 105, 211–48.

Lettau, Martin, and Sydney Ludvigson, 2001, Resurrecting the (C)CAPM: a cross-sectional test when risk premia are time-varying, *Journal of Political Economy* 109, 1238–1287.

Lo, Andrew, 2001, Risk management for hedge funds: Introduction and overview, *Financial Analysts Journal* 57, 16–33.

Longstaff, Francis A., and Jiang Wang, 2012, Asset pricing and the credit market, *Review of Financial Studies* 25, 3169–3215.

Lustig, Hanno, Nikolai Roussanov, and Adrien Verdelhan, 2011, Common Risk Factors in Currency Markets, *Review of Financial Studies* 24, 3731–3777.

Malliaris, Steven, and Hongjun Yan, 2010, Reputation Concerns and Slow-Moving Capital, Working Paper, Yale University.

McCracken, Michael W., 2007, Asymptotics for out of sample tests of Granger causality, *Journal of Econometrics* 140, 719–752.

Muller, Ulrich K., 2014, Hac corrections for strongly autocorrelated time series, *Journal of Business & Economic Statistics* 32, 311–322.

Newey, W. K., and K. D. West, 1987, A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix estimator, *Econometrica* 55, 703–708.

Pan, Jun, and Allen Poteshman, 2006, The information in option volume for future stock prices, *Review of Financial Studies* 19, 871–908.

Welch, Ivo, and Amit Goyal, 2008, A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies* 21, 1455–1508.

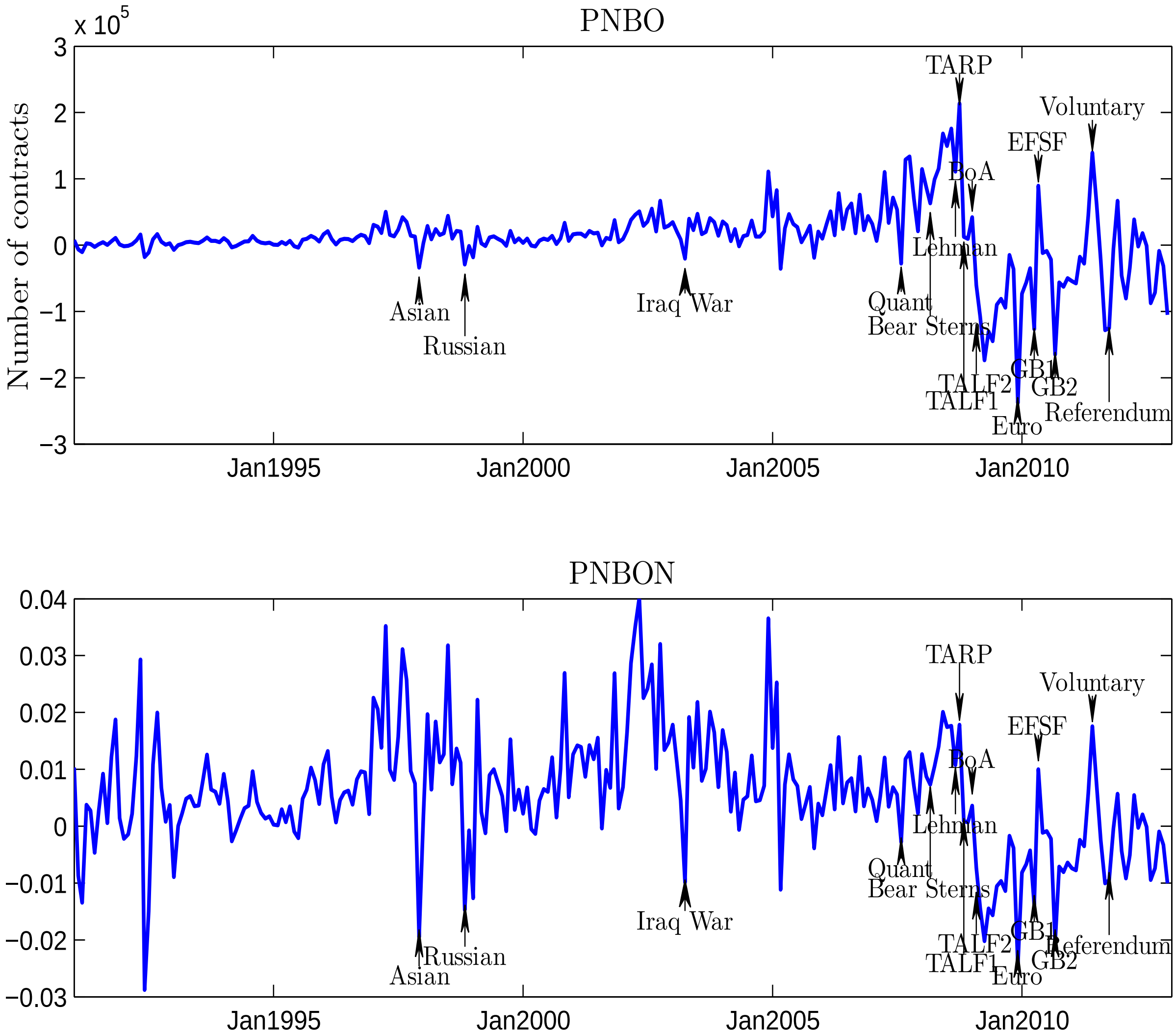

Figure 1: **Time Series of net public purchase for DOTM SPX puts.** $PNBO$ is the net amount of deep-out-of-the-money (DOTM)(with $K/S \leq 0.85$) SPX puts public investors buying-to-open each month. $PNBON$ is $PNBO$ normalized by average of previous 3-month total volume from public investors."Asian" (1997/10): period around the Asian financial crisis. "Russian" (1998/11): period around Russian default. "Iraq" (2003/04): start of the Iraq War. "Quant" (2007/08): the crisis of quant-strategy hedge funds. "Bear Sterns" (2008/03): acquisition of Bear Sterns by JPMorgan. "Lehman" (2008/09): Lehman bankruptcy. "TARP" (2008/10): establishment of TARP. "TALF1" (2008/11): creation of TALF. "BoA" (2009/01): Treasury, Fed, and FDIC assistance to Bank of America. "TALF2" (2009/02): increase of TALF to $1 trillion. "Euro" (2009/12): escalation of Greek debt crisis. "GB1" (2010/04): Greece seeks financial support from euro and IMF. "EFSF" (2010/05): establishment of EFSM and EFSF; 110 billion bailout package to Greece agreed. "GB2" (2010/09): a second Greek bailout installment. "Voluntary" (2011/06): Merkel agrees to voluntary Greece bondholder role. "Referendum" (2011/10): further escalation of Euro debt crisis with the call for a Greek referendum.

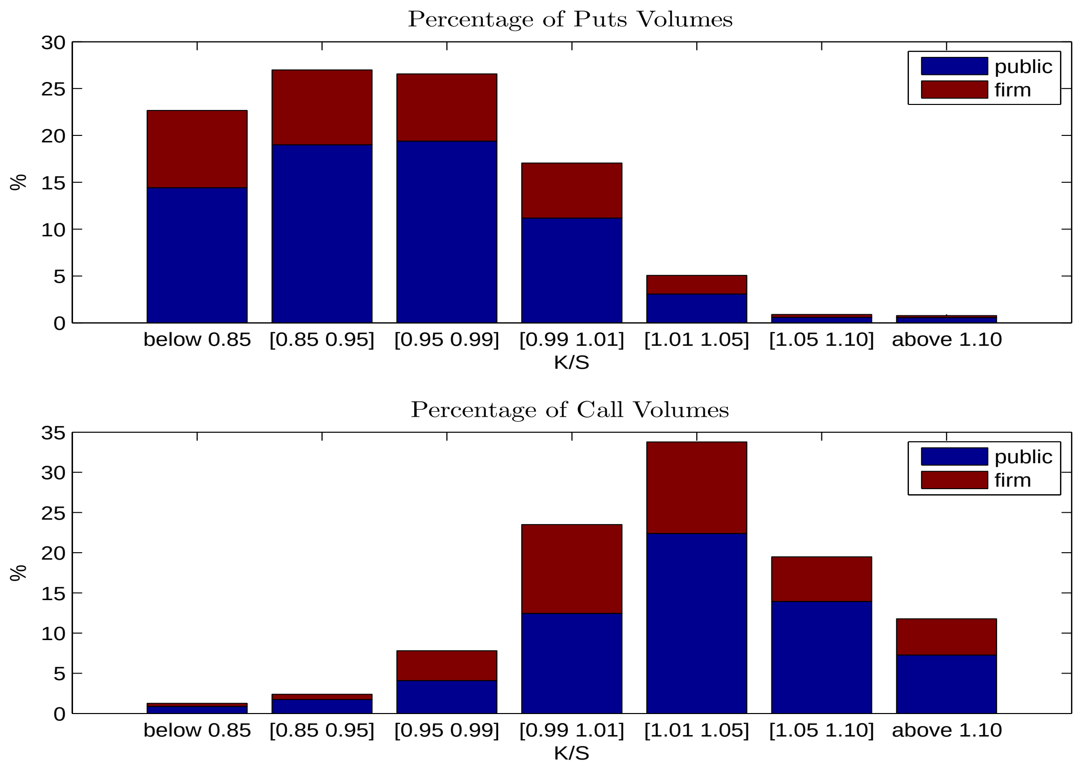

Figure 2: **Percentage of total put and call volumes at different moneyness.** This figure plots the total fraction of volume for calls and puts at different levels of moneyness (measure by strike price, K, divided by spot price S). The height of each bar indicates the fraction of the total market volume at that moneyness level, while the colors indicate the breakdown within the strike between public (blue) and private (red) orders.

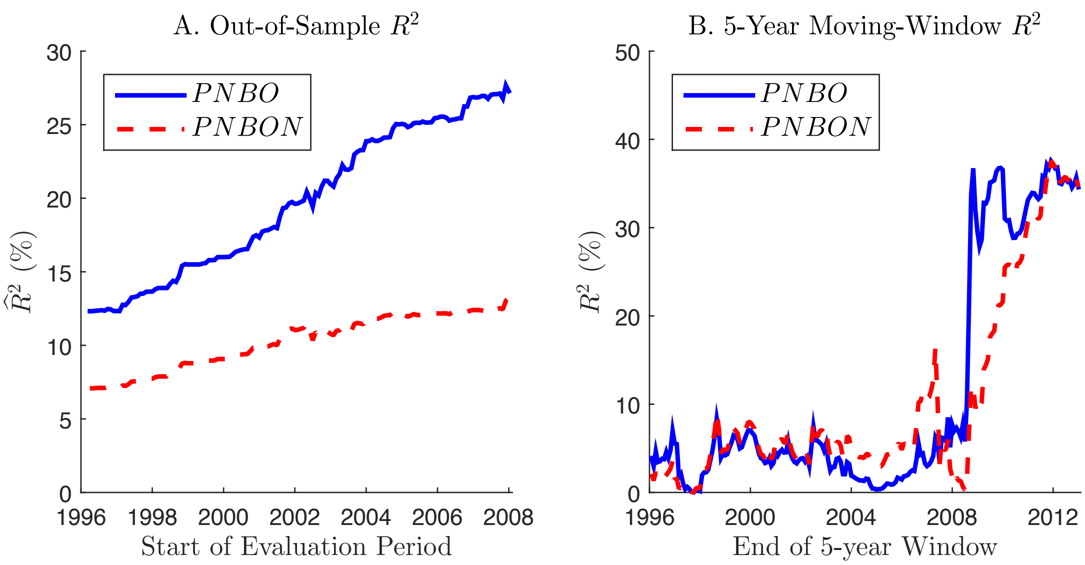

Figure 3: **Out-of-sample R^2 and R^2 from 5-year moving-window regressions.** This figures plots the out of sample R^2 from projecting future 3-month market excess returns on a constant and $PNBO$ (or $PNBON$) interacted with a dummy for the variance premium being negative: $r_{t+1\rightarrow t+3} = a_r + b_r\ I_{\{b_{VP,t}<0\}} \times PNBO_t + \epsilon_{t+1\rightarrow t+3}$. $PNBO$ is the net amount of deep-out-of-the-money (with $K/S \leq 0.85$) SPX puts public investors buying-to-open each month. $PNBON$ is $PNBO$ normalized by average of previous 3-month total volume from public investors.Panel A plots the out-of-sample R^2 as a function of the sample split date. Panel B plots the in-sample R^2 of 5-year moving-window regressions.

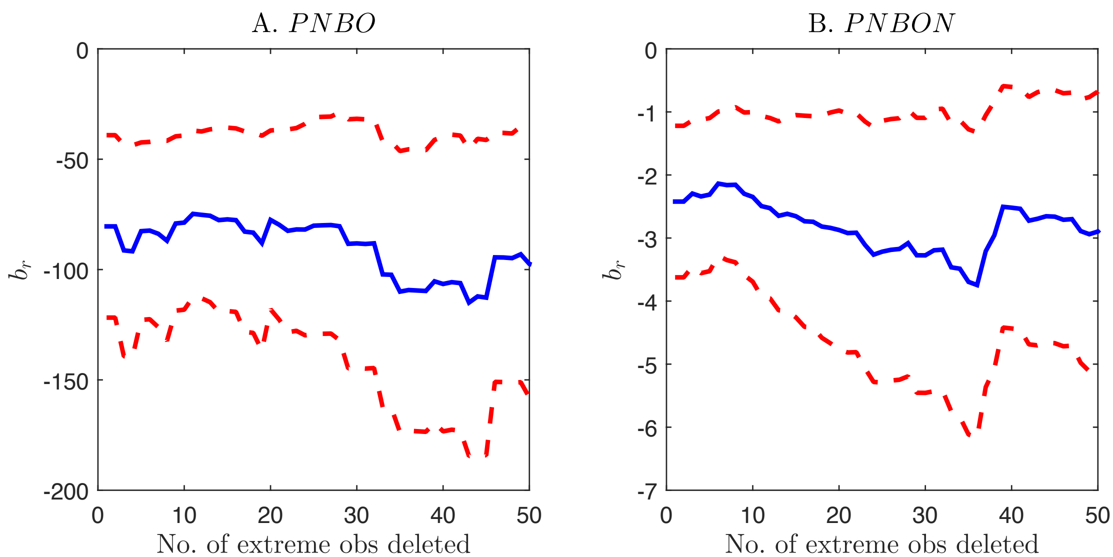

Figure 4: **Predictability of $PNBO$ and $PNBON$ after deleting extreme observations.** This figures plots point estimates and 95% confidence intervals, after removing various numbers of extreme absolute $PNBO$ observations, for the slope, b_r, from projecting future 3-month market excess returns on a constant and $PNBO$ (or $PNBON$) interacted with a dummy for the variance premium being negative: $r_{t+1\to t+3} = a_r + b_r\ I_{\{b_{VP,t}<0\}} \times PNBO_t + \epsilon_{t+1\to t+3}$. $PNBO$ is the net amount of deep-out-of-the-money (with $K/S \leq 0.85$) SPX puts public investors buying-to-open each month. $PNBON$ is $PNBO$ normalized by average of previous 3-month total volume from public investors.

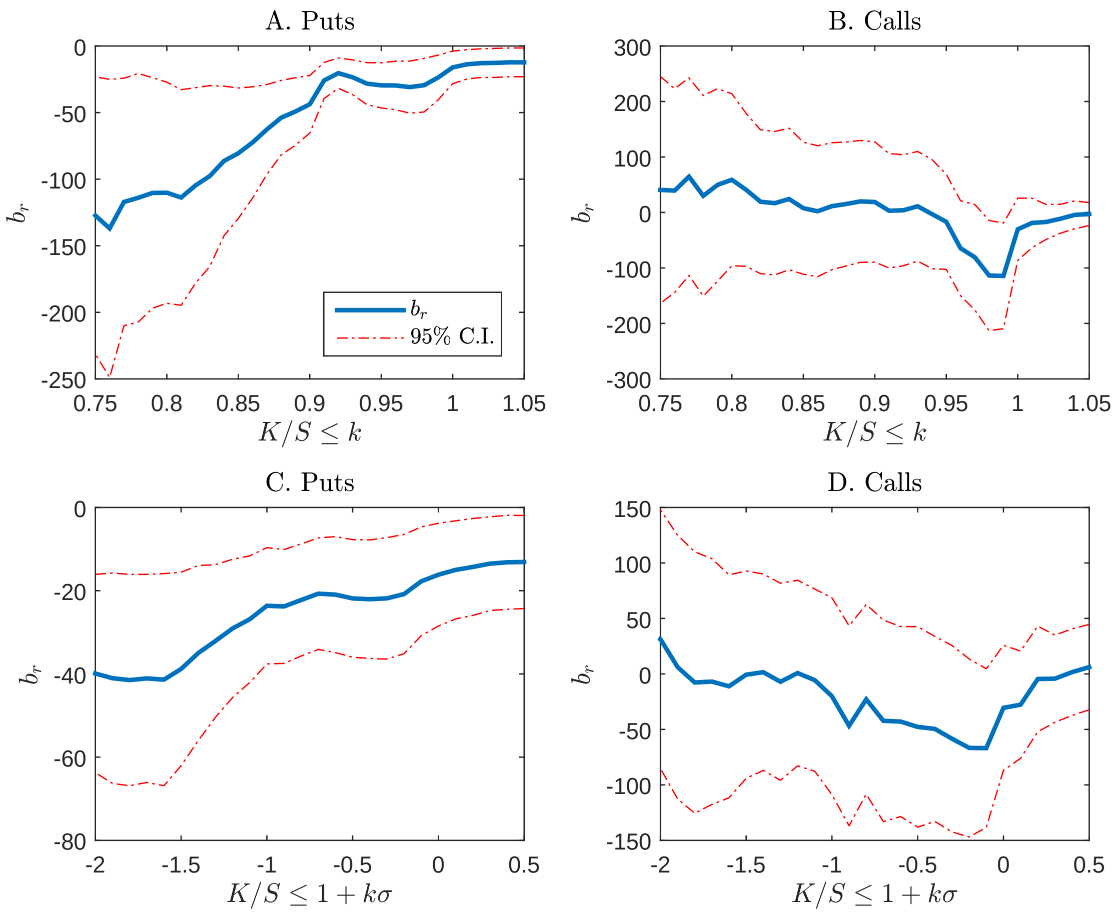

Figure 5: **Predictive power of $PNBO$ at different moneyness.** This figure plots the point estimate and 95% confidence interval of the slope coefficient in a regression of 3-month market excess return on a constant and a measure of public demand for puts (calls) interacted with an dummy for the variance premium being negative: $r_{t+1\to t+3} = a_r + b_r\ I_{\{b_{VP,t}<0\}} \times (\text{demand measure})_t + \epsilon_{t+1\to t+3}$. In Panels A, demand for puts is measured based on public orders to open for puts with K/S less than a constant cutoff k. In Panels C, demand for puts is measured based on public orders to open for puts with K/S less than $1 + k\sigma$, where k is a constant, and σ represents a maturity-adjusted return volatility, which is the daily S&P return volatility in the previous 30 trading days multiplied by the square root of the days to maturity for the option. Panels B and D are analogously defined for calls above the corresponding moneyness level. The 95% confidence intervals are based on Hodrick (1992) standard errors.

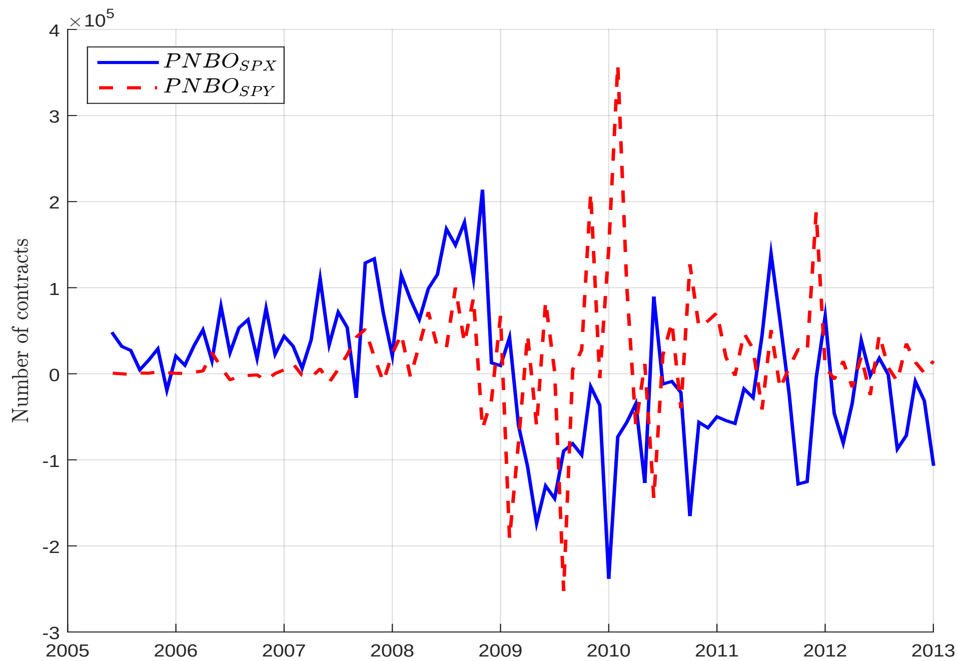

Figure 6: **Comparing $PNBO$ for SPX and SPY options.** This figure plots the public net buying-to-open volume for deep-out-of-the-money (with $K/S \leq 0.85$) puts in the market for SPX options and SPY options for the period of 2005/05 to 2012/12.

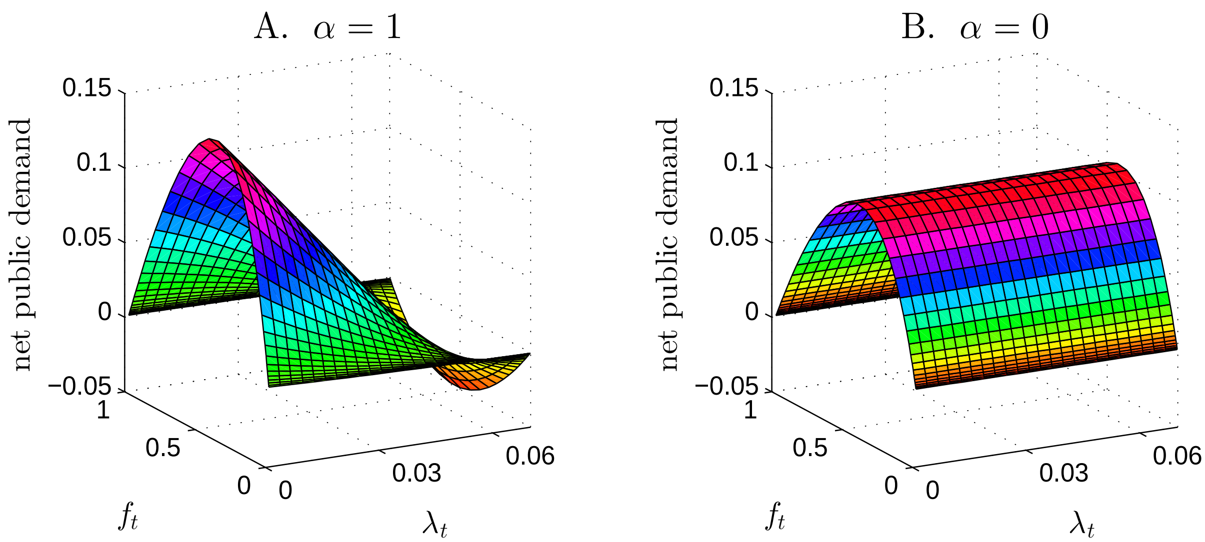

Figure A1: **Net Public Purchase for Crash Insurance.** The two panels plot the net public purchase for crash insurance as a function of the public investor P's consumption share (f_t) and the disaster intensity under P's beliefs (λ_t). Panel A considers the case when $\alpha = 1$, which implies that a 1-standard deviation increase in the disaster intensity from its long-run mean effectively increases the dealers' relative risk aversion against disasters by 1 in the homogeneous-agent economy. Panel B considers the case when $\alpha = 0$.

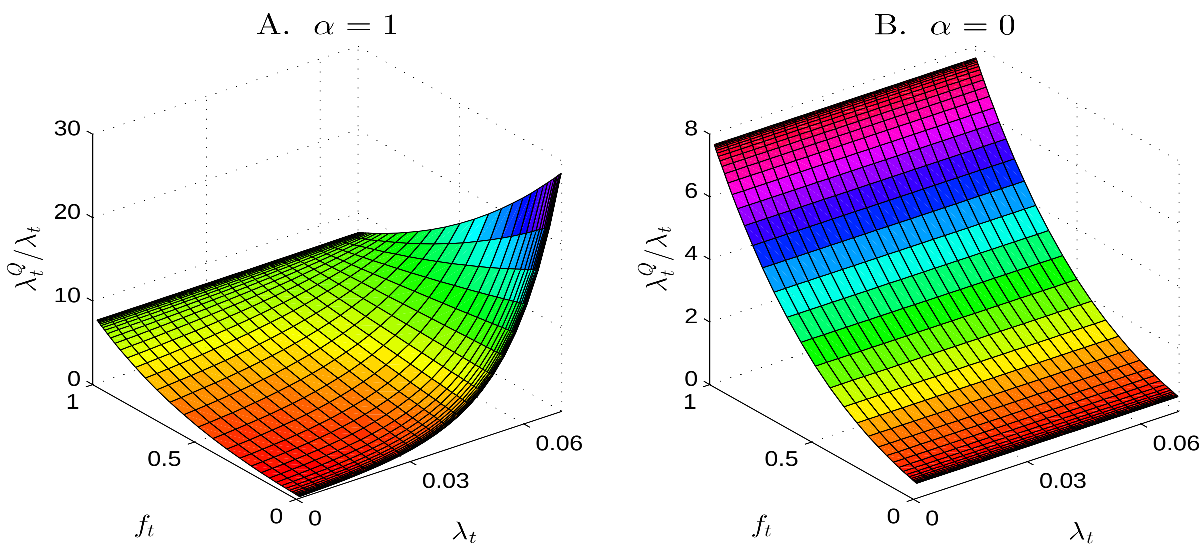

Figure A2: **Disaster Risk Premium.** The two panels plot the conditional disaster risk premium as a function of the public investor P's consumption share (f_t) and the disaster intensity under P's beliefs (λ_t). Panel A considers the case when $\alpha = 1$. Panel B considers the case when $\alpha = 0$. Again, $\alpha = 1$ means that a 1-standard deviation increase in the disaster intensity from its long-run mean effectively increases the dealers' relative risk aversion against disasters by 1 in the homogeneous-agent economy.

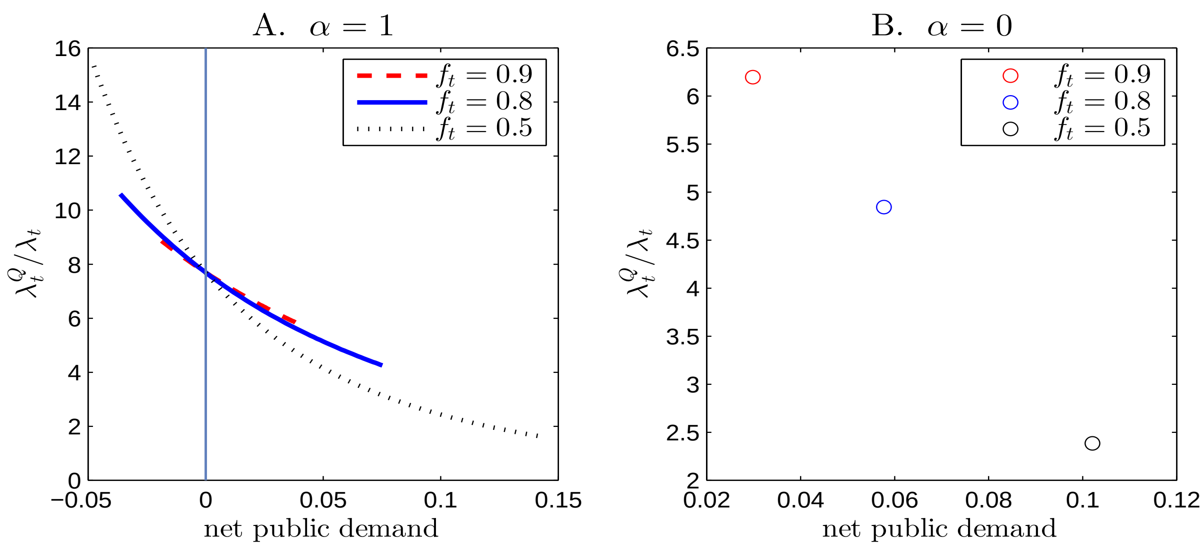

Figure A3: **Net Public Purchase for Crash Insurance and Risk Premium: Fixed Consumption Share.** The two panels plot the conditional disaster risk premium against the net public purchase for crash insurance while holding the public investors' consumption share f_t constant. Panel A considers the case when $\alpha = 1$. Panel B considers the case when $\alpha = 0$.

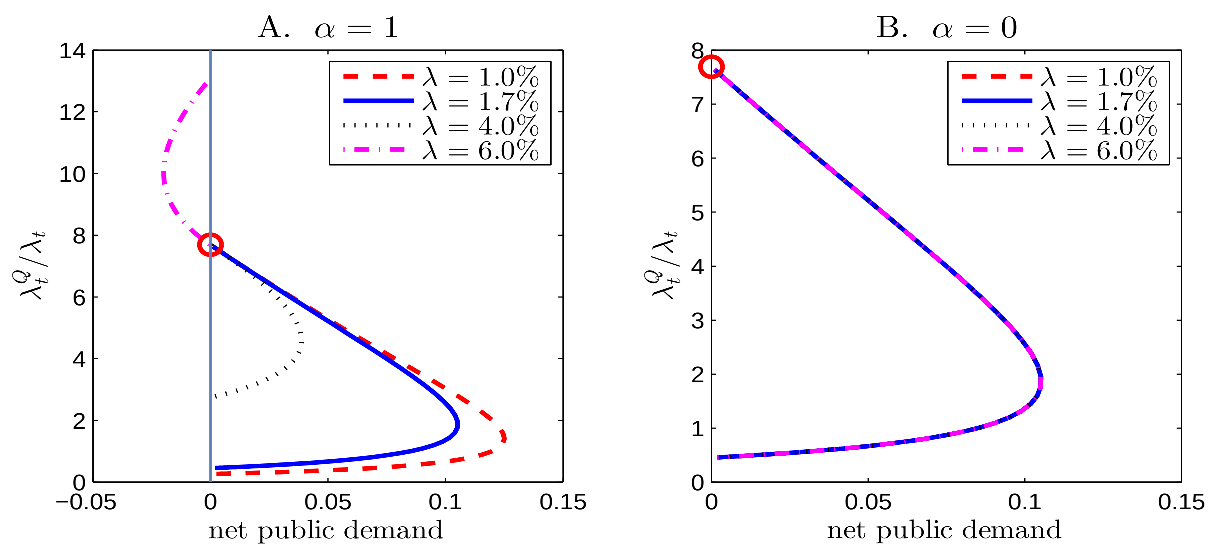

Figure A4: **Net Public Purchase for Crash Insurance and Risk Premium: Fixed Disaster Intensity.** The two panels plot the conditional disaster risk premium against the net public purchase for crash insurance while holding the disaster intensity λ_t constant. Panel A considers the case when $\alpha = 1$. Panel B considers the case when $\alpha = 0$. The red circles mark the limiting cases in the two economies where public investors own all the aggregate endowment.

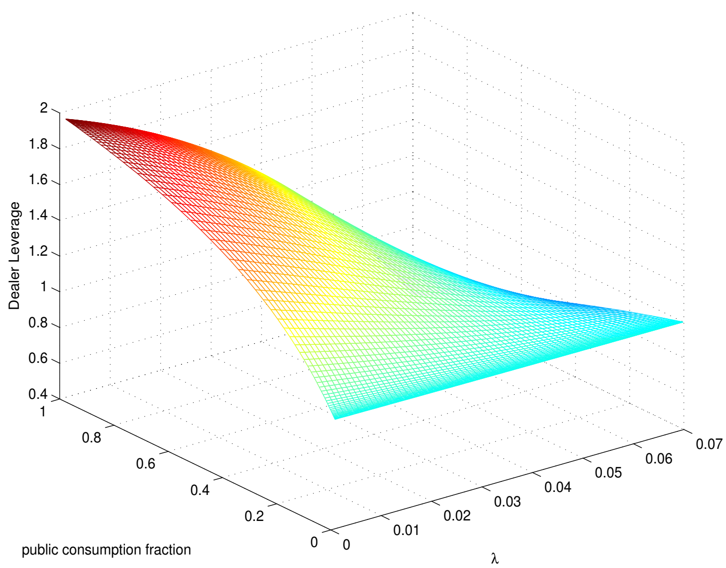

Figure A5: **Dealer leverage.** This figure plots the dealer leverage as a function of the public consumption fraction and the disaster probability.

Table 1: **Summary Statistics**

This table reports the summary statistics for the SPX options volume from public investors and pricing variables in the empirical analysis. $PNBO$: public net open-buying volume of DOTM puts ($K/S \leq 0.85$). $PNBON$: $PNBO$ normalized by average monthly public SPX volume over past 3 months (in million contracts). $PNBO_{ND}$: public net open-buying volume of all SPX options excluding DOTM puts. $PNOI$: public net open interest for DOTM SPX puts (in million contracts). $PNOIN$: $PNOI$ normalized by the total public open interest of all options (long and short). J: monthly average of the daily physical jump risk measure by Andersen, Bollerslev, and Diebold (2007). VP: variance premium based on Bekaert and Hoerova (2014). AC(1) is the first order autocorrelation for monthly time series; pp-test is the p-value for the Phillips-Perron test for unit root. Sample period: 1991/01-2012/12.

	mean	median	std	AC(1)	pp-test
$PNBO \times 10^6$ (contracts)	9996	9665	51117	0.61	0.00
$PNBON$ (%)	0.56	0.53	1.07	0.48	0.00
$PNBO_{ND} \times 10^6$ (contracts)	138075	117103	113637	0.76	0.01
$PNOI \times 10^6$ (contracts)	28788	19847	63212	0.74	0.00
$PNOIN$ (%)	0.18	0.15	0.36	0.70	0.00
J (%)	12.14	10.81	6.17	0.62	0.00
VP	19.37	14.56	12.66	0.54	0.00

Table 2: **Explaining Options Returns with Hedging Portfolios**

This table shows the R^2 from regressing option returns on hedging portfolios returns. The dependent variables are the returns of put options with different moneyness. delta denotes the returns on the delta hedging portfolio for the corresponding put option. delt+gam denotes the returns on the delta-gamma hedging portfolio. The sample period: $1996 - 2012$.

	$\frac{K}{S} < 0.85$		$0.85 < \frac{K}{S} < 0.95$		$0.95 < \frac{K}{S} < 0.99$		$0.99 < \frac{K}{S} < 1.01$	
	delta	delt+gam	delta	delt+gam	delta	delt+gam	delta	delt+gam
Weekly R^2	0.34	0.54	0.45	0.75	0.59	0.86	0.76	0.91
Daily R^2	0.41	0.59	0.46	0.74	0.56	0.82	0.72	0.87

Table 3: $PNBO$ **and SPX Option Expensiveness**

The dependent variable is VP: the variance premium in Bekaert and Hoerova (2014). We use three different measures of the public net-buying volumes: $PNBO$ (public net buying-to-open volume for deep-out-of-the-money (with $K/S \leq 0.85$) puts), $PNBON$ ($PNBO$ normalized by past 3-month average public volume), and $PNBO_{ND}$ (public net buying volume of all SPX options excluding deep-out-of-the-money puts). J is the average of daily physical jump of S&P 500 in Andersen, Bollerslev, and Diebold (2007). Standard errors in parentheses are computed based on Newey and West (1987) with 6 lags. (***,**,*) denote significance at 1%, 5%, and 10%, respectively. Sample period: 1991/01 to 2012/12.

$$VP_t = a_{VP} + b_{VP}\ PNBO_t + c_{VP}\ J_t \times PNBO_t + d_{VP}\ J_t + \epsilon_t^v$$

	$PNBO$		$PNBON$		$PNBO_{ND}$	
a_{VP}	-2.83	-4.80	-1.88	-5.92*	-1.13	-7.08
	(3.05)	(3.17)	(3.35)	(3.45)	(4.23)	(7.40)
b_{VP}	-101.57***	13.16	-3.24***	4.51	-20.22	11.43
	(24.13)	(27.09)	(1.39)	(2.80)	(13.73)	(29.18)
c_{VP}		-6.83***		-0.57**		-2.33
		(1.81)		(0.24)		(2.85)
d_{VP}	1.91***	2.05***	1.90***	2.20***	1.92***	2.38***
	(0.31)	(0.32)	(0.35)	(0.34)	(0.38)	(0.75)
$\bar{R}^2$ (%)	34.1	36.8	30.9	34.5	29.5	30.1

Table 4: **Return Forecasts with $PNBO$**

This table reports the results of the return forecasting regressions using $PNBO$ and $PNBON$. $r_{t+j\to t+k}$ represents market excess return from $t+j$ to $t+k$ ($k > j \geq 0$). $PNBO$ is the net amount of deep-out-of-the-money (with $K/S \leq 0.85$) SPX puts public investors buying-to-open each month. $PNBON$ is $PNBO$ normalized by average of previous 3-month total volume from public investors. VP is the variance premium based on Bekaert and Hoerova (2014). J is monthly average of the daily physical jump risk measure by Andersen, Bollerslev, and Diebold (2007). $\bar{J}$ is the median of monthly J_t for the full sample. Standard errors (σ) in parentheses are computed based on Hodrick (1992). Sample period: 1991/01 – 2012/12. (***,**,*) denote significance at 1%, 5%, and 10%, respectively.

Return	b_r	$\sigma(b_r)$	$\bar{R}^2$ (%)	b_r	$\sigma(b_r)$	$\bar{R}^2$ (%)
A: $r_{t+j\to t+k} = a_r + b_r\ I_{\{b_{VP,t}<0\}} \times PNBO_t + \epsilon_{t+j\to t+k}$						
	$PNBO$			$PNBON$		
r_{t+1}	-21.13***	(8.03)	3.2	-0.85**	(0.37)	2.8
r_{t+2}	-31.45***	(10.26)	7.6	-0.95***	(0.35)	3.6
r_{t+3}	-29.09***	(10.42)	6.4	-0.61*	(0.34)	1.3
r_{t+4}	-16.71*	(9.74)	1.9	-0.51	(0.34)	0.8
$r_{t\to t+3}$	-80.54***	(24.99)	15.4	-2.43***	(0.82)	7.5
B: $r_{t+j\to t+k} = a_r + b_r\ I_{\{J_t>\bar{J}\}} \times PNBO_t + \epsilon_{t+j\to t+k}$						
	$PNBO$			$PNBON$		
r_{t+1}	-30.20***	(9.70)	7.5	-0.61*	(0.36)	1.1
r_{t+2}	-24.44***	(9.42)	4.8	-0.76**	(0.35)	1.9
r_{t+3}	-27.94***	(9.94)	6.4	-1.00**	(0.41)	3.6
r_{t+4}	-14.59*	(8.16)	1.5	-0.42	(0.35)	0.3
$r_{t\to t+3}$	-81.57***	(25.32)	17.0	-2.34***	(0.89)	6.1

Table 5: **Return Forecasts with $PNBO$: Sub-sample Results**

This table reports the sub-sample results of the return forecasting regressions using $PNBO$ and $PNBON$. $r_{t+j\to t+k}$ represents market excess return from $t+j$ to $t+k$ ($k > j \geq 0$). $PNBO$ is the net amount of deep-out-of-the-money (with $K/S \leq 0.85$) SPX puts public investors buying-to-open each month. $PNBON$ is $PNBO$ normalized by average of previous 3-month total volume from public investors. VP is the variance premium based on Bekaert and Hoerova (2014). J is monthly average of the daily physical jump risk measure by Andersen, Bollerslev, and Diebold (2007). $\bar{J}$ is the median of monthly J_t for the full sample. Standard errors (σ) in parentheses are computed based on Hodrick (1992). Sample period: 1991/01 – 2012/12. (***,**,*) denote significance at 1%, 5%, and 10%, respectively.

Sub-sample	b_r	$\sigma(b_r)$	R^2 (%)	b_r	$\sigma(b_r)$	R^2 (%)	obs
$r_{t\to t+3} = a_r + b_r\ PNBO_t + \epsilon_{t\to t+3}$							
	$PNBO$			$PNBON$			
$b_{VP,t} < 0, J_t \geq \bar{J}$	-109.26***	(40.09)	31.2	-3.54***	(1.17)	14.8	80
$b_{VP,t} < 0, J_t < \bar{J}$	-52.42***	(18.43)	16.2	-1.91**	(0.88)	11.9	79
$b_{VP,t} \geq 0, J_t \geq \bar{J}$	-61.01***	(22.46)	20.3	-2.27	(1.62)	9.4	52
$b_{VP,t} \geq 0, J_t < \bar{J}$	17.41	(36.66)	1.3	0.51	(1.72)	0.4	53

Table 6: **Return Forecasts for Other Financial Assets**

This table reports the results of forecasting future excess returns on a variety of assets using $PNBO$ and $PNBON$. $r_{t\to t+3}$ represents excess return from t to $t+3$. $PNBO$ is the net amount of deep-out-of-the-money (with $K/S \leq 0.85$) SPX puts public investors buying-to-open each month. $PNBON$ is $PNBO$ normalized by average of previous 3-month total volume from public investors. VP is the variance premium based on Bekaert and Hoerova (2014). Asset return definitions are given in Section 3.3. Standard errors in parentheses are computed based on Hodrick (1992). Sample period: 1991/01 - 2012/12. (***,** ,*) denote significance at 1%, 5%, and 10%, respectively.

Return	b_r	$\sigma(b_r)$	$\bar{R}^2$ (%)	b_r	$\sigma(b_r)$	$\bar{R}^2$ (%)
	$r_{t\to t+3} = a_r + b_r\ I_{\{b_{VP,t}<0\}} \times PNBO_t + \epsilon_{t\to t+3}$					
	$PNBO$			$PNBON$		
High Yield	-55.15**	(21.56)	16.9	-1.34***	(0.51)	5.3
Hedge Fund	-33.26***	(10.72)	9.0	-1.04***	(0.33)	4.7
Carry Trade	-38.53**	(17.60)	6.7	-0.59	(0.43)	0.5
Commodity	-72.54*	(42.48)	6.0	-1.41	(0.95)	1.0
10-Year Treasury	19.28**	(8.74)	3.6	0.64**	(0.31)	2.1
Variance Swap	1345.16***	(402.87)	19.0	35.83***	(10.65)	7.2

Table 7: **Return Forecasts with $PNBO$ and Other Predictors**

The table reports the results of forecasting 3-month market excess returns with $PNBO$ and other predictors, including the variance risk premium (IVRV), log dividend yield ($d-p$), log net payout yield (lcrspnpy), Baa-Aaa credit spread (DEF), term spread (TERM), tail risk measure (Tail), implied volatility slope measures (Slope) and the quarterly consumption-wealth ratio ($\widehat{cay}$). Standard errors in parentheses are computed based on Hodrick (1992). Sample period is 1991-2012, except for lcrspnpy and Tail (1991-2010), and IVSlope (1996-2012). (***,**,*) denote significance at 1%, 5%, and 10%, respectively.

$$r_{t\to t+3} = a_r + b_r\ I_{b_{VP,t}<0} \times PNBO_t + c_r\ \mathbf{X}_t + \epsilon_{t\to t+3}$$

$PNBO \times I_{b_{VP}}$	-72.63***	-77.46***	-88.03***	-80.97***	-82.04***	-85.04***	-76.36***	-76.50***	-59.91**
	(24.80)	(25.40)	(29.33)	(25.40)	(25.65)	(29.62)	(23.87)	(29.05)	(23.53)
IVRV	11.94***							14.79***	
	(3.67)							(4.18)	
$d-p$		4.17						13.04*	
		(3.12)						(7.45)	
lcrspnpy			5.89					6.45	
			(4.82)					(7.76)	
DEF				-0.70				-5.56	
				(2.92)				(3.84)	
TERM					-0.38			-0.52	
					(0.68)			(0.85)	
Tail						26.19		-52.07	
						(34.71)		(46.89)	
IVSlope							0.22	0.11	
							(0.16)	(0.16)	
$\widehat{cay}$									29.48
									(31.11)
$\bar{R}^2$ (%)	24.1	17.3	17.6	15.2	15.4	15.4	18.1	36.6	10.6
$PNBON \times I_{b_{VP}}$	-2.36***	-2.19***	-2.20***	-2.46***	-2.45***	-2.31***	-2.79***	-2.24**	-1.73**
	(0.82)	(0.81)	(0.84)	(0.83)	(0.82)	(0.84)	(0.94)	(0.97)	(0.88)
IVRV	13.50***							15.46***	
	(3.73)							(4.22)	
$d-p$		3.86						14.21*	
		(3.06)						(7.32)	
lcrspnpy			3.73					1.20	
			(4.88)					(7.72)	
DEF				-0.80				-5.17	
				(2.91)				(3.84)	
TERM					-0.19			-0.04	
					(0.66)			(0.84)	
Tail						26.27		-35.95	
						(34.73)		(45.79)	
IVSlope							0.25	0.18	
							(0.16)	(0.17)	
$\widehat{cay}$									34.96
									(31.45)
$\bar{R}^2$ (%)	19.0	9.1	7.6	7.4	7.3	7.2	12.6	30.5	3.2

Table 8: $PNBO$ **and Measures of Funding Constraints**

Panel A reports the results of the OLS regressions of $I_{b_{VP,t}<0} \times PNBO_t$ on measures of funding constraints. Panel B reports the results of return predictability regressions with $PNBO$ and funding constraint measures. $PNBO$ is the net amount of deep-out-of-the-money (with $K/S \leq 0.85$) SPX puts public investors buying-to-open each month. VP is the variance premium based on Bekaert and Hoerova (2014). TED is the TED spread; LIBOR-OIS is the spread between 3-month LIBOR and overnight indexed swap rates; VIX is the CBOE VIX index; FG is the funding liquidity measure by Fontaine and Garcia (2012); Noise is the illiquidity measure by Hu, Pan, and Wang (2013); Δlev is the broker-dealer balance sheet growth measure by Adrian, Moench, and Shin (2010). Standard errors in parentheses are computed based on Newey and West (1987) with 6 lags. Sample period is 1991-2012, except for the regressions with LIBOR-OIS, which is 2002-2012. (***,**,*) denote significance at 1%, 5%, and 10%, respectively.

	A: $I_{b_{VP,t}<0} \times PNBO_t = a + b\,\mathbf{X}_t + e_t$						
TED	22.04**					52.61***	
	(10.57)					(12.39)	
LIBOR-OIS		0.25					
		(0.16)					
VIX			-0.13			-0.78	
			(0.30)			(0.56)	
FG				-4.89**		-11.61***	
				(2.32)		(2.86)	
Noise					0.54	-2.49	
					(1.68)	(1.67)	
Δlev							42.59***
							(16.68)
$\bar{R}^2$ (%)	3.9	2.2	-0.3	1.1	-0.3	12.3	7.8
	B: $r_{t\to t+3} = a_r + b_r\ I_{b_{VP,t}<0} \times PNBO_t + c_r\ \mathbf{X}_t + \epsilon_{t\to t+3}$						
$PNBO \times I_{b_{VP}}$	-74.96***	-73.56***	-80.34***	-78.75***	-79.98***	-62.51***	-72.19**
	(23.67)	(23.36)	(25.09)	(25.41)	(24.83)	(23.49)	(32.93)
TED	-2.87					-6.09*	
	(3.04)					(3.67)	
LIBOR-OIS		-0.07					
		(0.05)					
VIX			0.04			0.23*	
			(0.13)			(0.14)	
FG				0.60		1.46	
				(0.88)		(0.94)	
Noise					-0.32	-0.34	
					(0.61)	(0.79)	
Δlev							-3.58
							(3.00)
$\bar{R}^2$ (%)	16.8	28.6	28.6	15.2	15.6	20.8	18.4

Table 9: **Comparing SPX vs. SPY Trades**

This table reports the results of the return forecasting regressions based on two measures of public net buying-to-open volume: public net buying to open for SPX options ($PNBO_{SPX}$, denoted elsewhere as simply $PNBO$) and public net buying to open for SPY options ($PNBO_{SPY}$). All options under consideration are DOTM, as defined by $K/S \leq 0.85$. Standard errors (σ) in parentheses are computed based on Hodrick (1992). Sample period: 2005/05 - 2012/12.(***,**,*) denote significance at 1%, 5%, and 10%, respectively.

Return	b_r	$\sigma(b_r)$	$\bar{R}^2$ (%)	b_r	$\sigma(b_r)$	$\bar{R}^2$ (%)
	$r_{t+j\to t+k} = a_r + b_r\ I_{\{b_{VP,t}<0\}} \times PNBO_t + \epsilon_{t+j\to t+k}$					
	$PNBO_{SPX}$			$PNBO_{SPY}$		
r_{t+1}	-18.68***	(8.48)	6.1	-8.03	(14.54)	0.5
r_{t+2}	-31.39***	(11.29)	17.3	-22.13	(13.67)	3.9
r_{t+3}	-33.22***	(11.54)	19.3	3.47	(15.01)	0.1
$r_{t\to t+3}$	-81.54***	(23.44)	31.2	-26.12	(28.17)	1.4

Table 10: **Return Forecasting Outside the Financial Crisis**

This table reports the results of the return forecasting regressions in pre- and post-crisis subsamples. $r_{t\to t+3}$ indicate market excess return one- and three- month ahead, respectively. $PNBO$ is the net amount of deep-out-of-the-money (with $K/S \leq 0.85$) SPX puts public investors buying-to-open each month. VP is the variance premium based on Bekaert and Hoerova (2014). Pre-crisis: 1991/01 - 2007/11. Post-crisis: 2009/06 - 2012/12. Standard errors in parentheses are computed based on Hodrick (1992). (***,**,*) denote significance at 1%, 5%, and 10%, respectively.

Return	b_r	$\sigma(b_r)$	$\bar{R}^2$ (%)	b_r	$\sigma(b_r)$	$\bar{R}^2$ (%)
$r_{t\to t+3} = a_r + b_r\ I_{\{b_{VP,t}<0\}} \times PNBO_t + \epsilon_{t\to t+3}$						
Pre-crisis		$PNBO$			$PNBON$	
$r_{t\to t+3}$	-57.88***	(21.28)	3.1	-1.40**	(0.61)	3.5
Post-crisis		$PNBO$			$PNBON$	
$r_{t\to t+3}$	-47.38**	(23.75)	15.8	-4.57**	(2.32)	18.3

Table 11: **Return Forecasts with** $PNOI$

This table reports the results of the return forecasting regressions using $PNOI$ and $PNOIN$. $PNOI$ is the end-of-month public net open interest for deep-out-of-the-money ($K/S \leq 0.85$) puts. $PNOIN$ is $PNOI$ normalized by the sum of public long and short open interest. VP is the variance premium based on Bekaert and Hoerova (2014). $r_{t+j\to t+k}$ represents market excess return from $t+j$ to $t+k$ ($k > j \geq 0$). Standard errors (σ) in parentheses are computed based on Hodrick (1992). Sample period: 1991/01 – 2012/12. (***,** ,*) denote significance at 1%, 5%, and 10%, respectively.

Return	b_r	$\sigma(b_r)$	$\bar{R}^2$ (%)	b_r	$\sigma(b_r)$	$\bar{R}^2$ (%)
A: $r_{t+j\to t+k} = a_r + b_r\ I_{\{b_{VP,t}<0\}} \times PNOI_t + \epsilon_{t+j\to t+k}$						
	$PNOI$			$PNOIN$		
r_{t+1}	-22.71***	(7.95)	6.3	-3.33***	(1.05)	4.8
r_{t+2}	-19.71**	(9.00)	4.7	-2.27**	(1.14)	2.2
r_{t+3}	-16.19**	(7.20)	3.2	-1.76*	(0.93)	1.3
r_{t+4}	-6.72	(6.63)	0.5	-0.82	(0.96)	0.3
$r_{t\to t+3}$	-57.49***	(21.22)	11.9	-7.31***	(2.13)	7.0
B: $r_{t+j\to t+k} = a_r + b_r\ I_{\{J_t>\bar{J}\}} \times PNOI_t + \epsilon_{t+j\to t+k}$						
	$PNOI$			$PNOIN$		
r_{t+1}	-18.75**	(8.83)	4.0	-2.10*	(1.13)	1.4
r_{t+2}	-11.17	(8.51)	1.2	-0.78	(1.10)	0.0
r_{t+3}	-16.07**	(8.04)	2.9	-1.67*	(1.01)	0.8
r_{t+4}	-9.00	(6.87)	0.6	-1.01	(0.89)	0.0
$r_{t\to t+3}$	-44.26**	(22.33)	7.1	-4.35*	(2.38)	2.0

Table 12: **Return Forecasts with $PNOI$: Sub-sample Results**

This table reports the sub-sample results of the return forecasting regressions using $PNOI$ and $PNOIN$. $PNOI$ is the end-of-month public net open interest for deep-out-of-the-money ($K/S \leq 0.85$) puts. $PNOIN$ is $PNOI$ normalized by the sum of public long and short open interest. VP is the variance premium based on Bekaert and Hoerova (2014). J is monthly average of the daily physical jump risk measure by Andersen, Bollerslev, and Diebold (2007). $\bar{J}$ is the median of monthly J_t for the full sample. Standard errors (σ) in parentheses are computed based on Hodrick (1992). Sample period: 1991/01 – 2012/12. (***,**,*) denote significance at 1%, 5%, and 10%, respectively.

Sub-sample	b_r	$\sigma(b_r)$	R^2 (%)	b_r	$\sigma(b_r)$	R^2 (%)	obs
$r_{t\to t+3} = a_r + b_r\ PNOI_t + \epsilon_{t\to t+3}$							
	$PNOI$			$PNOIN$			
$b_{VP,t} < 0, J_t \geq \bar{J}$	-78.60**	(35.56)	23.9	-9.25**	(3.78)	11.0	80
$b_{VP,t} < 0, J_t < \bar{J}$	-42.12***	(14.59)	14.2	-6.77***	(2.57)	13.1	79
$b_{VP,t} \geq 0, J_t \geq \bar{J}$	-22.79	(19.35)	3.4	-1.25	(3.14)	0.0	52
$b_{VP,t} \geq 0, J_t < \bar{J}$	-11.03	(14.70)	1.7	-2.23	(2.87)	1.8	53

Table A1: **Comparison with GPP Table 2**

This table replicates results in Garleanu, Pedersen, and Poteshman (2009). The table reports regression coefficients from regressing option expensiveness, measured by the average implied volatility of ATM options minus a reference model-implied volatility used in Bates (2006), on measures of option demand. NetDemand and JumpRisk are the equal- and vega-weighted public net open interest (or net volume) for all SPX options. t-stats are computed based on Newey and West (1987) standard errors with 10 lags. (***,**,*) denote significance at 1%, 5%, and 10%, respectively.

	Net Open Interest				Net Volume	
	Replicated Results		GPP Results			
	NetDemand	JumpRisk	NetDemand	JumpRisk	NetDemand	JumpRisk
1996/01-1996/10	2.1×10^{-7}	4.5×10^{-6}	2.1×10^{-7}	6.4×10^{-6}	1.6×10^{-8}	-5.1×10^{-6}
t-stat	(0.72)	(0.49)	(0.87)	(0.79)	(0.10)	(-0.23)
1997/10-2001/12	$3.8\times10^{-7**}$	$2.6\times10^{-5***}$	3.8×10^{-7}	$3.2\times10^{-5***}$	$4.7\times10^{-6**}$	$4.9\times10^{-5**}$
t-stat	(2.17)	(2.80)	(1.55)	(3.68)	(2.25)	(2.57)
2002/01-2012/12	$-9.9\times10^{-8***}$	$-7.8\times10^{-6**}$			$-6.7\times10^{-7**}$	$-5.2\times10^{-5***}$
t-stat	(-3.13)	(-2.40)			(-2.35)	(-3.32)

Table A2: **Model Parameters**

risk aversion: γ	4
time preference: δ	0.03
mean growth of endowment: $\bar{g}$	0.025
volatility of endowment growth: σ_c	0.02
mean intensity of disaster: $\bar{\lambda}$	1.7%
speed of mean reversion for disaster intensity: κ	0.142
disaster intensity volatility parameter: σ	0.05
dealer risk aversion parameter: α	1.0