Measuring Limits of Arbitrage in Fixed-Income Markets

Jean-Sébastien Fontaine  Guillaume Nolin
Bank of Canada

April 2018

Abstract

We use relative value to measure limits of arbitrage in fixed-income markets in a way that is simple, intuitive and model-free. We construct an index of relative value to measure limits to arbitrage for the US, UK, Japan, Germany, Italy, France, Switzerland and Canada. Relative value indices exhibit strong commonality across countries and high correlations with volatility and funding conditions within countries. The price of risk estimate for relative value is negative and robust in the cross-section of bond and option returns. Relative value shows meaningful economic value as trading signal. Overall, relative value is a cleaner measure of limits of arbitrage in fixed-income markets than the most common alternative. The relative value indices are updated regularly and available publicly.

Keywords: Limits of Arbitrage, Sovereign Bonds

JEL Classification: G12

We thank Bruce Chen, Jens Christensen, Darrell Duffie, Bruno Feunou, Meredith Fraser-Ohman, Étienne Giroux-Léveillé, Toni Gravelle, Sermin Gungor, Madhu Kalimipalli, Natasha Khan, Gitanjali Kumar, Stéphane Lavoie, James Pinnington, Francisco Rivadeneira, Adrien Verdhelan, Harri Vikstedt, Jun Yang, Jing Yang and Bo Young Chang for comments suggestions. We also thank James Pinnington for invaluable research assistance. Email: gnolin@bankofcanada.ca
Introduction

Bonds from the same issuer and with the same cash flows should have the same prices. This is the law of one price. It may be a surprise that deviations of prices from the law of one price are pervasive in bond markets. Panels (A) and (B) of Figure 1 show the yields to maturity of all United States Treasury bonds for a day in 2008 and in 2014, respectively. Deviations can be large—as in 2008—or small—as in 2014—but they are rarely absent.

We expect that arbitrage activities will compress price deviations. Instead, deviations in fixed income markets are common, persistent and significant, even after accounting for direct transaction costs (see Amihud and Mendelson 1991 or, for a review, Fontaine and Garcia 2015). Hence, deviations reveal limits to arbitrage that are due to market frictions, including funding frictions and market segmentations (Duffie 1996; Vayanos and Weill 2008; Vayanos and Vila 2009).

Deviations like those in Figure 1 can be aggregated to reveal variations of limits to arbitrage over time. This is valuable information. Using a dynamic model, Fontaine and Garcia (2012) show that an index of deviations predict excess bond returns across fixed-income markets. D’Amico and Pancost (2017) also use a dynamic model, and argue that repo risk can limit arbitrage. The most common methods use a static parametric yield curve. Hu et al. (2013) (HPW thereafter) show that the “noise” measure—an index of fitting errors—is priced in financial markets. Malkhozov et al. (2016) study the role of funding constraints in international financial markets. Musto et al. (2017) study the feedback between liquidity and investor clienteles. HPW’s measure (or a variant) is often used in policy discussions about the bond market, notably by Bisias et al. (2015), Adrian et al. (2015) and Dudley (2016).

We introduce a model-free measure of limits to arbitrage that is intuitive and easy to compute. For any bond in our sample, we use a small number of comparable bonds
to form a replicating portfolio with the same duration and convexity. Our measure of relative value is the yield difference between a bond and its replicating portfolio. The basic idea is that a bond and its replicating portfolio carry the same risk and should offer the same expected return. This idea builds on the simple butterfly trade, a common strategy used by investors to exploit price deviations. We first compute this measure for every bond and then we aggregate bond-level measures into a relative value index. We repeat this exercise for the US, Canada, UK, Germany, France, Italy, Switzerland and Japan. These indices are available publicly and updated regularly.

Our relative value index is distinct from existing measures because it is model-free. We perform three distinct checks that the relative value index is a meaningful measure of limits to arbitrage. First, we check that a higher value of the index is correlated with other proxies for limits to arbitrage. We find that relative value in each country is highly correlated with volatility in the local market (e.g., the VIX in the US). We also find that a relative value in each country is correlated with the local interbank funding rate (e.g., LIBOR-OIS spread for the US). In the cross-section, we find that relative value indices are highly correlated with each other as well as with US proxies for volatility and funding rates. This points to a common relative value factor related to global sources of risk, consistent with evidence in Malkhozov et al. (2016) based on the noise index.

Second, we test whether relative value is a proxy for risks that are priced in the cross-section of asset returns. In the cross-section of US corporate bonds, we find that relative value carries a negative price of risk. The estimate is highly significant and robust. Relative value explains a large share of dispersion in returns (close to 65%). Relative value also carries a negative price of risk in the cross-section of S&P

---

1 Alternatively, a relative value trade could involve the perfect replication of cash flows. However, this is impractical and costlier than replicating interest rate risk, as it requires extensive trading in the relatively illiquid market for bond strips.

2 See the authors’ page on the Bank of Canada’s website.
500 call and put options, where we use de-levered returns from Constantinides et al. (2013). The estimate is negative, significant and close to results using corporate bonds. Relative value explains a large share of the dispersion in returns, especially in the challenging case of put options. For both asset classes, a change by two standard deviations in the cross-section $\beta$s increases average returns by around 6 percent annually. The results strongly suggest that relative value proxies for risks that are priced in financial markets, since limits to arbitrage play an important role in the asset classes we considered. By comparison, results obtained using noise provide mixed evidence: the estimate is sometimes insignificant or has the wrong sign, the share of dispersion explained is small, or the spread in $\beta$ coefficients is small.

Third, we compare the returns from a pseudo-trading strategy that uses the signal from relative value to exploit deviations in bond prices. The returns from trading on this signal is an objective measure of its economic content, and the returns from holding a portfolio of trades is an objective measure of economic content for the relative value index. We expect that a signal that correctly identifies limits to arbitrage generates positive profits, precisely because the pseudo-trading strategy ignores other risks and constraints that arbitrageurs faced. Higher returns mean that costs and the risks of arbitrage were larger.

In the US Treasury bond market, returns on trades that use the signal from relative value exceed returns using noise. Accounting for bid-ask spreads and assuming a conservative level of capital at risk, relative value produces an average monthly return of 0.09% between 1988 and 2017. By contrast, using noise as a signal to initiate trades but implementing the same trading strategy produces an average return of $-0.19\%$. Repeating this comparison in every 2-year sub-sample produces similar results. Repeating this comparison in the US, UK, Japan, Switzerland, Germany, Italy, France and Canada produces similar results.
Using a model-free measure removes preliminary estimation of parameters in a yield curve model and eliminates an important source of noise that is due to sampling uncertainty and model misspecification. In itself, the large variety of curve-fitting methods used by practitioners suggests a certain level of arbitrariness in the estimation process. To check this idea, we repeat the trading exercise, but increasing the threshold that a signal must meet before we implement a new trades. We find that a small threshold on noise (a few basis points) quickly produces positive returns that improves upon a risk-free strategy. But relative value still over-performs for any level of the threshold.

Our approach is also distinct from a strategy that only matches the duration of a bond. For instance, Longstaff (2004) matches Treasury bonds with bonds issued by the Resolution Funding Corporation (Refcorp), which are guaranteed by the US Treasury. Krishnamurthy (2002) matches recently issued Treasury bonds. Instead we use information from several bonds to form a replicating portfolio and we also match convexity. This approach is more robust and more widely applicable.

This paper is structured as follows. Section 1 details the model-free relative value measure for individual bonds. Section 2 describes the data, the construction of the relative value indices, compares these indices to other proxies and describes the results of asset pricing tests. Section 3 compare relative value and noise using a pseudo-trading strategy.

1 Methodology

The model-free measure of relative value is inspired by the butterfly bond-trading strategy, used by market participants to profit from deviations between the prices of similar bonds. The butterfly strategy typically involves the combination of opposite positions in the target bond and in a portfolio of bonds with similar duration, which
are the wings of the butterfly. One bond with a lower duration and one bond with a higher duration, with weights to match the duration of the target bond.

For our measure, the wings include three bonds with weights computed to match the duration, convexity and par value of the target. Matching the duration and convexity neutralizes differences in maturities and coupon rates. Otherwise, these bonds are identical. Therefore, the portfolio replicates the risk of the target bond: assuming the absence of arbitrage, the target bond and the replicating portfolio should have the same expected return.

The relative value of a bond with respect to its replicating portfolio is a proxy for limits that prevents arbitrageurs from driving deviations to zero. Arbitrage only works if investors can use risk-free and self-financing trading strategies to profit. Rebalancing costs and other frictions can limit arbitrage in bond markets (Duffie, 1996; Vayanos and Weill, 2008; Vayanos and Vila, 2009).

1.1 Calculating Portfolio Weights

The replicating portfolio combines three bonds from the same issuer. To select these bonds, we sort outstanding bonds by duration. The first bond has the closest but lower duration than the target bond. The second bond has the closest but higher duration than the target bond. That is, the first two bonds are closest to the target bond but on opposite sides of it on the duration axis. The third bond is on the same side of the duration axis as the bond with the smallest (absolute) difference in duration. The bonds in the replicating portfolio are very close substitutes to the target bonds.

We label the target bond with $b$ and label the bonds in the replicating portfolio with $i = 1, 2, 3$. The duration and convexity of the bonds are denoted by $d_i$ and $c_i$ respectively. We compute modified duration and convexity using mid yields. Every
day and for every bond, the weights $w_i$ in the replicating portfolio as a fraction of par value are chosen to satisfy:

$$d_1 w_1 + d_2 w_2 + d_3 w_3 = w_b d_b$$  \hspace{1cm} (1)

$$c_1 w_1 + c_2 w_2 + c_3 w_3 = w_b c_b$$  \hspace{1cm} (2)

$$w_1 + w_2 + w_3 = w_b = 1.$$  \hspace{1cm} (3)

This system of equations requires that the replicating portfolio has the same duration (Equation 1), the same convexity (Equation 2), and the same par value than the target bond (Equation 3). Without loss of generality, we normalize the weight $w_b$ to 1.

To eliminate extreme portfolio weights, we impose the following conditions on the weights:

$$0 \leq w_i \leq 2/3 \quad \forall i \in \{1, 2, 3\}.$$  \hspace{1cm} (4)

The left hand side of this constraint means that the replicating portfolio does not mix short and long positions. The right hand side of this constraint means that relative value averages information from the prices of several neighbouring bonds. The system defined by Equations (1-3) may have no solution if the inequality in Equation 4 is binding. In this case, we relax the equality constraint in Equation 2 but we minimize the difference in convexity between the target and the replicating portfolio:

$$\arg\min_{w_1, w_2, w_3} (c_1 w_1 + c_2 w_2 + c_3 w_3 - c_b),$$  \hspace{1cm} (5)

subject to Equations 1, 3 and 4.\footnote{We use the \texttt{lsqlin} function with the interior-point algorithm in MATLAB to solve our constrained linear least-squares problem. Our initial point for the solution process is a vector with equal weights 1/3.}
1.2 Computing Relative Value with Bid and Ask Prices

We compute the yield to maturity of the replicating portfolio using the weights $w_1, w_2, w_3$ for this day as well as the bid, mid and ask yields $y_{it}^{bid}$, $y_{it}^{mid}$ and $y_{it}^{ask}$ of bond $i$ at time $t$:

\[
y^{bid}_t = y^{bid}_{1t} w_1 + y^{bid}_{2t} w_2 + y^{bid}_{3t} w_3 \quad (6)
\]

\[
y^{mid}_t = y^{mid}_{1t} w_1 + y^{mid}_{2t} w_2 + y^{mid}_{3t} w_3 \quad (7)
\]

\[
y^{ask}_t = y^{ask}_{1t} w_1 + y^{ask}_{2t} w_2 + y^{ask}_{3t} w_3. \quad (8)
\]

When the observed yield for the target bond is higher than that of the replicating portfolio, the arbitrage strategy to profit from this deviation is to buy the target and short the replicating portfolio. The opposite strategy is profitable when the observed yield for the target is lower than that of the replicating portfolio. Therefore, the relative value is $r_{vt}^+ = y^{ask}_b - y^{bid}_t$ when the target bond is expensive relative to its replicating portfolio and the relative value is $r_{vt}^- = -y^{bid}_t + y^{ask}_t$ when the target bond is cheap relative to its replicating portfolio. Note that $r_{vt}^+$ and $r_{vt}^-$ cannot be strictly positive at the same time: the target cannot be simultaneously expensive and cheap relative to its replicating portfolio. However, $r_{vt}^+$ and $r_{vt}^-$ can be negative at the same time because of bid-ask spreads. The relative value of target bond $b$ is defined as:

\[
r_{vt} \equiv \begin{cases} 
  r_{vt}^+ & \text{if } r_{vt}^+ > 0 \\
  0 & \text{if } r_{vt}^+ < 0 \text{ and } r_{vt}^- < 0 \\
 -r_{vt}^- & \text{if } r_{vt}^- > 0,
\end{cases} \quad (9)
\]
which implies that $rv_{bt}$ takes negative value $-rv_{bt}$ when the target is relatively cheap. Note that $rv_{bt}$ is set to zero if $rv_{bt}^+$ and $rv_{bt}^-$ are negative, since the absence of arbitrage holds in this case once we account for transaction costs. An alternative definition simply uses mid yields:

$$
\tilde{rv}_{bt} = -y_{bt}^{mid} + \hat{y}_t^{mid}.
$$

This definition produces larger relative values than our base case and overstates the presence of limits of arbitrage.

2 International Relative Value Indices

In this section, we compute a daily relative value index for bonds issued by Canada, France, Germany, Italy, Japan, Switzerland, the UK and the US, respectively. All indices are updated regularly and available on the authors’ page on Bank of Canada’s website.

2.1 Data

We use end-of-day clean prices and yields to maturity for US, British, Canadian, French, German, Italian, Japanese and Swiss government bonds. Following HPW, we include bonds that have between 1 and 10 years remaining to maturity, but we exclude inflation-linked bonds, variable-rate bonds and bonds with embedded options. Other restrictions and filters are described in the Appendix.

For the US, we use prices and yields of Treasury bonds data between 1988 and December 2017, from the Center for Research in Security Prices database (CRSP). Before 1987, the different tax treatments of coupon and principal payments affected the relative valuation of bonds\footnote{In the US, the Tax Reform Act of 1986 introduced equal tax treatment of coupon and principal payments. Before that, tax deductions for coupon payments made higher coupon bonds more expensive. See Green and Ødegaard (1997) for a complete discussion.}. For Canada, we use daily prices between 1994 and
December 2017 from StatPro before April 2009 and from FTSE TMX thereafter. Similar to the US, coupon and capital gains had different tax treatments before 1994. For bonds issued by the British, French, German, Italian, Japanese and Swiss governments, we use daily data from Markit Evaluated Bonds between 2005 and December 2017. Table 1 reports the number of observations for each country as well as summary statistics for the number of bonds available on any given day. Japan and the US have the largest number of bonds, Switzerland the lowest.

2.2 The Relative Value Index

The relative value index \( rv^j_t \) for country \( j \) is the root mean square of the individual relative values \( rv_{bt} \) for this country:

\[
rv^j_t = \sqrt{\frac{1}{N^j_t} \sum_{b=1}^{N^j_t} rv_{bt}^2},
\]

where \( N^j_t \) is the number of bonds in our sample for country \( j \) at time \( t \). Note that bonds with \( rv_{bt}^+ < 0 \) and \( rv_{bt}^- < 0 \) are excluded from this index, since then \( rv_{bt} = 0 \). Using root mean square parallels how HPW computes the noise index. Also following HPW, the index excludes observations where \( rv_{bt} \) is greater than four times the sample standard deviation.

2.3 Commonality

To motivate the construction of relative value indices, we check that relative value is highly correlated across bonds in each country. Table 2 reports results of a principal components analysis conducted on a panel of bonds in each country. Indeed, relative

\footnote{The dataset available from FTSE TMX was formerly available as PC-Bond and DEXIn Canada. Tax deductions for principal payments made higher coupon bonds cheaper before 1994, when the lifetime capital gains deduction was eliminated for property acquired after February 22nd, 1994.}

\footnote{We apply principal component analysis to balanced panels of relative values in each country, where the number of bonds in the panel for each country is calibrated to the period with the smallest
values are highly correlated: we find that the share of variations explained by the
first two principal components ranges between 40 and 80 percent depending on the
country.

Table 3 reports the correlations between the relative value indices of different
countries. The indices for Germany, Switzerland, the UK and Canada are all corre-
lated with the US dispersion index. The correlation is 0.86 with the UK index and
0.48 with Canada’s index. Panel [A] of Figure 2 shows the relative value indices
for Canada, Germany, Switzerland, the US and the UK. As expected, the relative
value indices exhibit very similar patterns: a calm period with low relative value be-
fore 2007, followed by a gradual increase in 2007 with a dramatic peak in 2009 and
smaller variations since then.

This pattern suggests that the limits of arbitrage are largely common for these
large advanced economies’ sovereign issuers. However, the indices for Italy and France
are correlated with each other but not with the indices of other countries. Panel [B]
of Figure 2 compares the relative value indices of France and Italy. The index for
Italy exhibits a striking peak at the height of the euro area sovereign debt crisis, and
the index for France also increases during that period, suggesting that the Euro crisis
had a large impact on limits to arbitrage in these two markets. Finally, the relative
value index for Japan appears to be largely idiosyncratic.

2.4 Relative Value, Funding Costs and Volatility

Figure 3 shows that the relative value indices in the US, UK, Germany and Canada
are strongly related to measures of equity market volatility index in each country
(Panels [A][D]). We use the VIX in the US, the Euro Stoxx 50 volatility index, the
number of outstanding issues. At each date in our sample, we populate the panel starting with the
bonds with remaining time to maturity closest to 10 years and adding bonds with shorter maturity
until the fixed size of the panel is reached. The results are robust to other specifications.
FTSE 100 volatility index and the Nikkei volatility index. These volatility indices are forward-looking. Table 4 reports the correlation. The correlation between each index and the local measure volatility is high in most cases. The correlation between each index and the US VIX is also very high in most cases, consistent with arbitrage capital that is scarcer when financial markets are more volatile. The correlations are lower for France and Italy, presumably because the euro area measure of volatility does not accurately reflect limits to arbitrage in these markets. The correlation is very low in Japan, which stands out as an exception.

Table 4 also reports the correlations of relative value indices with proxies for funding conditions in each country. We use the spread between interbank borrowing rates and overnight indexed swap (OIS) rates, computed locally and for the US. As expected, the relative value indices are positively correlated with funding conditions in essentially all countries. Japan stands out again as the exception.

2.5 Asset Pricing Tests

Relative value indices are strongly correlated with measures of volatility and funding conditions. This is consistent with existing results for similar proxies (e.g., Fontaine and Garcia 2012, Hu et al. 2013). It is important to check that the relative value index proxies for risks that are priced in the cross-section of returns. We want to know whether exposures to the relative value index ($\beta$s) match average returns. This establishes that relative value is a useful proxy for risk: it describes an important risk-return relationship embedded in market prices. The results for the noise index are mixed, which may be due to the preliminary estimation of a yield curve (i.e., sampling uncertainty).

To check this, we first consider the cross-section of corporate bonds. Corporate bonds are good candidates for our test, since they are notoriously illiquid and
et al. (2011) find that a proxy for aggregate liquidity explains individual bond yield spreads with large economic significance. Fontaine and Garcia (2012) find that bonds with lower ratings have more exposures to their funding liquidity measure.

We double-sort corporate bonds on their exposures to changes of market-wide illiquidity and market-wide volatility, respectively. The exposure of each corporate bond is measured by the $\beta$ of monthly bond returns. The $\beta$s are estimated using a rolling window with 3 years of data. At the beginning of each year, we form $5 \times 5 = 25$ portfolios. We then track their returns for one year, before rebalancing the portfolios again at the start of the next year.

Table 5 reports estimates of the prices of risk obtained for relative value $\Delta rv$ and for noise $\Delta noise$. In all cases, we include the market returns as a risk factor and among the test assets, to discipline the estimate for the price of market risk (see Lewellen et al. 2010). Panel (A) reports results using corporate bond portfolios as test assets. A simple message emerges: the price of risk for relative value is large, negative and significant. By contrast, the price of risk for noise is small, not statistically significant and changes sign in a specification that includes both relative value and noise.

Inspection of the portfolio reveals negative $\beta$s: an increase of the relative value index is associated with lower returns (unreported). Moving between portfolios with $\beta$s of 0 and -0.4, approximately two standard deviations in the cross-section $\beta$s, increases average annual returns by around 6 percent.

We also consider the cross-section of option returns. Constantinides et al. (2013) create portfolios of call and put options that are de-levered. Removing the effect of leverage in option contracts makes these portfolios appropriate in linear asset pricing tests. More importantly, Constantinides et al. (2013) find that the returns of

---

7Following Bai et al. (2015) closely, we combine several databases to construct a long sample of monthly individual corporate bond returns.
de-levered portfolios are puzzling from the perspective of the simple CAPM model, especially put options, but a proxy for illiquidity goes a long way to resolve the puzzle.

Panel (B) of Table 5 reports results using option portfolios as test assets. Results are tabulated for Put and Call options, Put only and Call only, respectively. The price of risk for relative value is essentially the same than in Panel (A) for bonds. In fact, the dispersion of relative value $\beta$s is also similar in the cross-section of bonds and options, so that the spread in returns explained by relative value is very similar for bonds and options. Augmenting the CAPM with relative value raises the cross-section $R^2$ from 24 to 85 percent for puts and from 56 to 85 for calls. In this augmented CAPM, we also find that the dispersion of relative value $\beta$s is wide for put options but much less so for call options. This is consistent with Constantinides et al. (2013) findings that the dispersion of average returns is puzzlingly large for puts. Results based on noise remain mixed. The price of risk estimate is significant in one case, but sometimes flips sign. The $R^2$s are also lower, especially for puts.

3 Relative Value as a Trading Signal

We use a pseudo-trading strategy to measure the economic content of relative value. This is important, since the bond-level signals constitute the relative index in Equation (10). The results show that relative value improves on noise as a signal used to implement pseudo-trading strategy. The results also help explain why the index of relative values performed better in asset pricing tests (Section 2.5).

We expect that trading profits from this pseudo-strategy to be positive when using a signal that correctly identifies limits to arbitrage. Duarte et al. (2007) show that these strategies amount to more than “picking up nickels in front of a steamroller.” A patient unconstrained investor should earn a positive return over time when following this strategy. It’s also important to remember that these are pseudo-trading
strategies, since our computation of returns does not account for funding costs, the costs of borrowing bonds and other costs that investors would face in realistic settings. This is a feature and not a bug. The results provides an objective measure of the economic content in relative value. In particular, our goal here is not to find the optimal trading strategy.

3.1 Implementing the Trading Strategy

We implement the trading strategy separately in each country. Every day, we use either relative value or the noise measure as a signal to determine whether a bond becomes the target of a new trade. This requires that we also estimate the noise measure in each country. In practice, we construct the noise measure to account for the bid-ask spreads (see Appendix A.3). Otherwise, the original noise measure would be at a disadvantage if used as a trading signal.

We initiate a new trade when the absolute value of the signal is greater than some threshold: $|rv_{bt}| > \tau$ or $|noise_{bt}| > \tau$, respectively. If the bond is a target, we implement the same butterfly strategy irrespective of which signal is used. Each trade initially has zero duration, near-zero convexity and thus carries little interest rate risk.\(^8\)

If $rv_{bt} > \tau$ the target bond appears expensive, the trading strategy is to buy one unit of this bond and to sell one unit of the replicating portfolio (with weights given by Equations 1 to 4. The opposite strategy is implemented when the target appears cheap. We track the performance of this trade over time. This position must be liquidated before this bond is targeted again. This is consistent with the construction of the index. We repeat the same exercise for using $noise_{bt} > \tau$ as a signal.

The baseline results use $\tau = 0$ because we want to evaluate the economic content

---

\(^8\)Trades are initiated when there are more than 5 days of pricing data remaining for the target bond and each of the constituents of the replicating portfolio.
of trades included in the relative value index (Equation 10). We also report results for \( \tau = 1, \ldots, 40 \) bps. Note that in the case \( \tau = 0 \), relative value \( |rv_{bt}| > \) or noise \( noise_{bt} > \tau \) suggest potential profits at least as large as the bid-ask spread. Larger values of \( \tau > 0 \) requires larger potential profits.

3.2 Liquidating Trades

Following [Duarte et al. (2007)], all trades are held until convergence: when the mid of the target bond \( y_{bt}^{mid} \) is equal to the mid of the replicating portfolio \( \hat{y}_{t}^{mid} \) or the fitted value \( \hat{y}_{bt} \) derived from the estimated yield curve. This realizes the potential profits of the arbitrage opportunity, since the bid-ask spread was accounted for when entering a new trade. Positions are also liquidated (i) when the target bond matures in less than three months, (ii) when data become unavailable, (iii) after one year or (iv) at the end of the sample.

3.3 Computing Trade Returns

We track the performance of each trade over time to compute an index of returns across all trades for each signal. It is very important to account for the cost of carry in the computation of returns. The carry varies across trades due to different initial investments, due to different coupon cash flows, and due to interest earned or spent.

We follow [Duarte et al. (2007)] to compute returns. For this purpose, it is useful to think of a specialized fund that targets a single bond. The equity value \( E_t \) of this fund at time \( t \) is given by:

\[
E_t = E_0 + B_t + P_t, \tag{11}
\]

where \( E_0 \) is the initial equity or margin set aside to cover for potential losses; \( B_t \) is the fund’s cash position (net borrowing if \( B_t < 0 \)); and \( P_t \) is the marked-to-market value.
of the target bond and its replicating portfolio, accounting for the bid-ask spread.\footnote{We use the following standard settlement convention: one business day for U.S. Treasury bonds and those of other countries except Canada, 2 days for Government of Canada bills and bonds with fewer than 3 years before maturity, 3 days for Government of Canada bonds with more than 3 years remaining before maturity.}

We can follow the equity $E_t$ of this fund given the initial equity $E_0$, the evolution of its cash position $B_t$ and the market value of bonds. For $t > 0$, the marked-to-market value $P_t$ is computed as if the positions were liquidated at end-of-day bid and ask prices, and accounting for accrued interest. At $t = 0$, the market value $P_0$ is recorded at the bid and ask prices of the initial transactions.

The evolution of the cash position $B_t$ is given by:

$$B_{t+1} = B_t(1 + r_{ft}) + C_{t+1}, \tag{12}$$

with initial cash position $B_0 = -P_0$, where $r_{ft}$ is the overnight rate, and $C_{t+1}$ are subsequent coupon received or due.\footnote{We use the Actual/360 basis to calculate the daily interest paid. For simplicity, we use the target overnight inter-bank rate or its equivalent. In the US, we use the effective Fed funds rate if available. Otherwise, we use the mid-point of the Federal Reserve official target range when available. For Canada, we use the Bank of Canada target rate. For the UK, we use the Bank of England’s official Bank rate. For euro area countries, we use the European Central Bank’s main refinancing operation rate. For Switzerland, we use the Swiss National Bank’s official LIBOR target. For Japan, we use the Bank of Japan’s policy rate.}

$Duarte et al. (2007)$ choose the initial equity $E_0$ such that the sample volatility of returns is 10 percent on average across funds. For simplicity and comparability, we choose $E_0 = 10$, which is enough to cover any loss recorded in our sample at any threshold. The volatility implied by this choice varies across countries. We verified that the volatility is around 10 percent or less.

Given the evolution of the equity value $E_{i,t}$, we can compute daily returns $R_{i,t+1}$ for each fund:

$$R_{i,t+1} = \frac{E_{i,t+1}}{E_{i,t}} - 1. \tag{13}$$
Following Duarte et al. (2007), the index returns $\tilde{R}_{t+1}$ is given by:

$$\tilde{R}_{t+1} = \frac{1}{I} \sum_{i} R_{st+1} \quad i f \quad I \geq 1$$  

$$\tilde{R}_{t+1} = r_{ft} \quad i f \quad I = 0,$$  

where $I$ is the number of funds in the index. We compute this index returns separately for trades initiated with relative value and noise, respectively. The index returns corresponds to a portfolio with equal weighted across funds that are rebalanced daily. If no fund is active on a given day, the index return for that day is the risk-free rate. The initial value of the index is set to $V_0 = 100$ and the evolution of the index is simply $V_{t+1} = V_t \times (1 + \tilde{R}_{t+1})$.

### 3.4 Cross-Section of Trade Returns

Figure 4 shows that trades generated with relative value or with noise are broadly similar, except for a small difference in the mean and volatility of returns. Panel (A) reports the duration of each trade in a histogram. For both signals, most trades are short-lived and liquidated within 30 days. Yet, more trades based on noise have to be liquidated after one year, having failed to converged. Panel (B) shows that the number of active trades is broadly similar for both trading signals over time. The distribution of the number of trades over time is generally similar between the two measures, but the noise signal general a larger number of trades in a handful of episodes.

Panel (C) of Figure 4 reports the total gains or losses from each trade in a histograms. For both signals, most trades have total profit between -50 and +50 cents, however relative value generates a larger proportion of trades with gains and a lower proportions of trades with losses. Finally, Panel (D) reports the histograms for the
variance of daily trade returns, showing that the noise measure produces a larger number of trades with very low volatility.

Panel (A) of Table 6 provides summary statistics of returns in weekly percentage terms for the US between 1988 and 2017, using our baseline threshold $\tau = 0$. Recall that this case accounts for the bid-ask spread. Total returns over the life of trades generated with relative value have higher mean and median than trades generated with the noise measure, as well as lower volatility. But the cross-sectional distribution is wide: the standard deviation is 3.84% for relative value and 9.23% for noise. Both signals generate a similar number of trades.

Overall, the results show that the noise measure generates more trades with low variance and small but lower and negative returns (in part because of transaction costs). In contrast, relative value trades have small but modestly positive average returns. One possibility is that these trades are initiated when fitting errors of the parametric yield curve produces false-positive signals and liquidated at a small loss when the fitting errors are reversed. The estimation procedure of the yield curve treats price deviation as measurement error (see Bolder and Gusba (2002) for a discussion).

3.5 Time Series of Portfolio Returns

Panel (A) of Figure 5 reports the cumulative index returns for portfolios of trades generated by relative value the and noise measure, respectively. Using noise as a signal yields cumulative returns that are consistently negative over the period. Between 1988 and 2017, this difference in returns accumulates and creates a large difference in outcomes. But note that the scale of the portfolio returns are also determined by the choice of the initial equity $E_0$. Lowering the initial equity increases leverage, produces higher average returns, but also increases volatility. We leave the optimal choice of equity for future research. In any case, it will not alter the ranking in Figure 5.
Table 6 reports the summary statistics of the time series of index returns, in monthly percentage terms. The average monthly returns of the portfolio of trades based on relative value and noise measures is on average 0.09% and -0.19%, respectively.

We checked that using noise yield negative returns for every two-year subsample between 1988 and 2017 and, further, that relative value generates better returns in each of these subsamples. We check that bid-ask spreads do not explain the better performance of relative value. Panel [B] of Figure 5 shows the cumulative index returns when excluding the effect of bid-ask spreads. Relative value continues to outperform noise; if anything its relative performance has improved. Inspection of individual trade returns (unreported) reveals that removing bid-ask spreads only shifts the distribution of returns but essentially nothing else and cannot affect the ranking between relative value and noise.

We also checked that any difference in duration or convexity does not explain the better performance of relative value. Recall that we do not rebalance trades after initiation. At the trade level, the effect is small. One reason for this is that most trades are liquidated within 30 days. Another reason is we use bonds that are very close substitutes in the replicating portfolios. This makes matching duration and convexity a robust approximation. For the aggregate portfolio of trades, we find that duration and convexity are slightly positive, but the differences between portfolios are small. Over the sample period, the average difference in duration between the relative value and noise portfolios is .004 years, or slightly more than a day. The average difference in convexity between the two portfolios is close to the convexity of a 6-month t-Bill. These differences are very small and cannot explain the difference in average returns: the relationship between convexity and expected returns is very small for bills (see e.g., Ilmanen 2000).
3.6 International Trading Performance

In this section, we compare the economic value of relative value and noise for sovereign bond markets in the US, UK, Canada, Germany, France, Italy, Switzerland and Japan. The data cover the period between January 2005 and December 2017. Table 7 reports summary statistics for the cross-section of trades in weekly percentage terms. Relative value trades have higher average returns than noise trades in every country.

Table 8 reports summary statistics for the time series of index returns in each country, reported in monthly percentage terms. Again, relative value outperforms noise in every country. The difference in average monthly returns ranges from 0.1% to 0.4%. Overall, the results confirm that using relative value identifies economically meaningful limits to arbitrage, but that using the noise measure generates a significant share of loss-making signals.

3.7 Portfolio Returns Using Higher Thresholds

In this section, we compare the returns of trades initiated using a higher threshold than in our baseline $\tau = 0$ bps. This is a natural check of the idea that noise produces negative returns because of small fitting errors. Indeed, we find that noise outperforms a risk-free investment in our sample of low values of the threshold $\tau$. However, we find in all cases that relative value outperforms noise.

Using the baseline threshold $\tau = 0$ to compare performances is consistent with the goal of evaluating the economic content of trades included in the relative value index. Raising the threshold excludes trades that targeting small deviations in the price of bonds and offer low potential profits. Excluding these trades is a more realistic trading strategy. Excluding these trades also provides a check that noise under-performs because of small errors which can be corrected with a relatively low threshold.
Figure 6 presents the results when excluding trade from each portfolio using successively higher thresholds $\tau = 1, \ldots, 40$ bps. Panel (A) shows that, as the threshold increases, the number of trades generated by the noise measure drops more rapidly than for relative value. The difference in the number of trades levels at a threshold of $\tau$ between 10 and 15 bps.

Panel (B) shows the value of the index of portfolio returns at the end of the sample period (December 31st 2017), by threshold and for both trading signals. It shows that relative value trades consistently overperforms noise by a wide margin at any threshold. We find that raising the threshold improves the performance of noise until a threshold between 10 and 15 bps again.

Together the results suggest that a threshold filter some of the model-induced noise and eliminate unprofitable trades from the portfolio. Yet, the higher thresholds reduces the number of trades generated by the noise measure and ultimately reduces performance compared to relative value.

**Conclusion**

We introduce the relative value measure of limits of arbitrage in fixed-income markets. Relative value is model-free and offers several benefits. It does not require the estimation of a parametric model. It is computed quickly; parametric methods require non-linear optimization. It requires only a few bonds to compute; parametric models require a large sample of bonds. It emphasizes price deviations that are economically significant; parametric curve produces results that are polluted by fitting errors. These relative value indices are available publicly and will be regularly updated. We hope that these indices will help to answer a number of research questions. In addition, future research could apply our methodology to create relative value indices for supranational, sub-national or corporate bond markets.
References


A Appendix

A.1 Other Data

Local interbank borrowing rates are USD LIBOR for the US, GBP LIBOR for the UK, Canadian Dollar Offered Rate (CDOR) for Canada, CHF LIBOR for Switzerland, JPY LIBOR for Japan and Euribor for euro area countries. OIS rates are USD OIS for the US, Sterling Overnight Index Average (SONIA) for the UK, CAD OIS for Canada, CHF OIS for Switzerland, Tokyo Overnight Average Rate (TONAR) for Japan and Euro OverNight Index Average (EONIA) for euro area countries. Local volatility indices are VIX for the US, VFTSE for the UK, VNKY for Japan and VSTOXX for the euro area countries.

A.2 Data Filter

For each country, we include all bullet government bonds with a remaining term to maturity between 1 and 10 years. To calculate the relative value measure for the shortest or longest maturity bonds, we rely on bills or bonds with slightly shorter or longer maturities. We exclude days when the yield to maturity is available for less than four bonds. We also exclude inflation-linked bonds, callable bonds, fungible bonds, strips, perpetual bonds (UK), retail bonds (Japan) and flower bonds (US).

From an original sample of 2210 bonds in the CRSP US Treasury dataset (1988-2017), the number of bonds considered is 1203, for a total of 573,860 daily observations. Note that the CRSP dataset does not require further filtering for errors, unlike the FTSE TMX, Bloomberg and Markit Evaluated Bonds datasets used for other countries. The filtering procedures described below apply to the latter.

We filter errors in the reported yield to maturity. Every day for every bond, we calculate the median level of mid yields for the 6 bonds with the nearest remaining term to maturity, as well as the absolute median deviation from this median. If the reported mid yield of the bond is larger than 5 bps from the median level and 6 times the median absolute deviation, we recalculate the bid, ask and mid yields to maturity using the quoted bid, ask and mid prices and replace their original values if the new values fall within the thresholds.

We filter errors in the reported bid-ask spread. Every day for every bond, we filter for bonds with a bid-ask spread of more than 30 bps, a bid-ask spread greater than $1 per $100 of par value or a negative bid-ask spread in terms of yields or price. For bonds breaching one of these thresholds, we calculate the median bid, ask and mid yields for the 6 bonds with the nearest remaining term to maturity, as well as the absolute median deviation from the median mid yield. We identify if the bid or ask yields are breaching the thresholds of 5 bps from the median level and 6 times the median absolute deviation away from the mid yield. If both bid and ask prices and yields breach the thresholds, we drop the observation. If only the bid (ask) yield is breaching the threshold, we replace the bid (ask) yield and price by the ask (bid) yield plus (minus) the median bid-ask spread of the neighbouring 6 bonds; we also replace the bid (ask) price by the ask (bid) price minus (plus) the median bid-ask spread of the neighbouring 6 bonds. In all cases, we exclude observations with bid-ask spreads wider than 6 times the 90th percentile of bid-ask spreads of active bonds with more than 200 days remaining until maturity, if this bid-ask spread exceeds 10 bps.

We filter obvious outliers. Every day for every bond issuer, we filter for obvious outliers in terms of yields or bid-ask spread. First, we estimate a yield curve using all bonds with more than
200 days remaining until maturity. This curve is estimated by OLS with a constant, the remaining
time to maturity and the square of the remaining time to maturity of every bond as the regressors.
We exclude bonds the absolute difference between the estimated curve and the observed yield to
maturity is more than twice the 90th percentile of regression residuals and if the yield to maturity
is more than 50 bps from the median of yields for the 4 bonds with the closest term to maturity.

A.3 Adjusting the Noise Measure to Incorporate Bid-Ask Spreads

We adapt HPW’s noise measure to make it directly comparable to relative value. HPW noise
measure does not account for the effect of bid-ask spreads, and would be at a disadvantage if
used as a trading signal. The noise measure is computed relative to the yield \( \hat{y}_{bt} \) derived from the
estimated curve using [Svensson (1994)] model\(^{[11]}\). We estimate the parametric curve daily for each
country.

The adjustment cannot be the same as in Equation 9, since there is no bid-ask for the estimated
yield curve. Instead, the distance between the mid quote and the estimated curve, accounting for
the bid-ask spread is given by:

\[
\begin{align*}
\epsilon^+_{bt} &\equiv \hat{y}_{bt} - \hat{y}_{bt} - (y_{ask} - y_{bid}) \\
\epsilon^-_{bt} &\equiv y_{mid} - \hat{y}_{bt} - (y_{ask} - y_{bid})
\end{align*}
\]

where \((y_{ask} - y_{bid})\) is the bid-ask spread for the target bond. Similar to the relative value, the noise
measure is then defined as:

\[
noise_{bt} = \begin{cases} 
\epsilon^+_{bt} & \text{if } \epsilon^+_{bt} > 0 \\
0 & \text{if } \epsilon^+_{bt} < 0 \text{ and } \epsilon^-_{bt} < 0 \\
-\epsilon^-_{bt} & \text{if } \epsilon^-_{bt} > 0 
\end{cases}
\]

\(^{[11]}\)The results are similar if we use [Grkaynak et al. (2007)] parameter estimates from the Federal Reserve
Board’s [website](#).
Figure 1: Dispersion of US Treasury Yields to Maturity
Yields to maturity for US Treasury securities from the CRSP database plotted against each security’s duration. The Grkaynak et al. (2007) parametric par curve (GSW) is plotted for comparison.

(A) 24 November 2008
(B) 24 November 2014

Figure 2: International Relative Value Indices
Aggregate relative value indices for sovereign issuers in different advanced economies. The figures are computed daily using end-of-day yields for bonds with between 1 and 10 years remaining until maturity.

(A) US, UK, Canada, Germany and Switzerland
(B) France and Italy
Figure 3: Relative Value Index and Volatility

Comparison of the sovereign issuer’s relative value index and the local stock market volatility index. Daily values.
Figure 4: Distribution of Profits and Volatility in the US (1988-2017)

Panel [A] histogram of the number of days between the implementation and liquidation of every trade based on the relative value and noise measures. Panel [B] plot of the number of active trades generated by the relative and noise measures, over time. Panel [C] histogram of total profits or losses of every trade based on the relative value or noise measures, including transaction costs. Panel [D] histograms of the volatility of daily trade returns for the relative value or the noise measure, respectively. The standard deviation of daily returns is calculated between the implementation and the liquidation of every trade.
Figure 5: Cumulative Portfolio Returns in the US (1988-2017)
Cumulative returns for the indices of equally weighted portfolios of trades generated using relative value and the noise measure, including (A) or excluding (B) the bid-ask spread. The indices are set to 100 on January 1st, 1988.

Figure 6: Trading Signals Using Thresholds (US 1988-2017)
Results from trading strategies in the US Treasury market using the relative value and noise measures where implementation threshold ranges between $\tau = 1, \ldots, 40$ bps. Panel A total number of trades initiated during the sample period, by threshold and measure. Panel B values of the indices of cumulative returns on December 31st, 2017, by threshold and measure, including the bid-ask spread. The indices are set to 100 in January 1st, 1988.
Table 1: **Summary Statistics for Sovereign Issuers**

<table>
<thead>
<tr>
<th></th>
<th>US</th>
<th>GB</th>
<th>CA</th>
<th>CH</th>
<th>DE</th>
<th>FR</th>
<th>IT</th>
<th>JP</th>
</tr>
</thead>
<tbody>
<tr>
<td># Obs.</td>
<td>502371.00</td>
<td>59881.00</td>
<td>94605.00</td>
<td>35456.00</td>
<td>116231.00</td>
<td>85606.00</td>
<td>130994.00</td>
<td>473412.00</td>
</tr>
<tr>
<td># Bonds, Median.</td>
<td>166.00</td>
<td>18.00</td>
<td>29.00</td>
<td>11.00</td>
<td>36.00</td>
<td>30.00</td>
<td>41.00</td>
<td>139.00</td>
</tr>
<tr>
<td># Bonds, 25th Perc.</td>
<td>103.00</td>
<td>16.00</td>
<td>26.00</td>
<td>10.00</td>
<td>28.00</td>
<td>18.00</td>
<td>26.00</td>
<td>131.00</td>
</tr>
<tr>
<td># Bonds, 75th Perc.</td>
<td>206.00</td>
<td>19.00</td>
<td>30.00</td>
<td>11.00</td>
<td>39.00</td>
<td>31.00</td>
<td>51.00</td>
<td>146.00</td>
</tr>
<tr>
<td>Avg. Duration</td>
<td>3.62</td>
<td>4.28</td>
<td>3.56</td>
<td>4.95</td>
<td>4.17</td>
<td>4.51</td>
<td>3.93</td>
<td>4.39</td>
</tr>
<tr>
<td>Avg. Std. Dev. of Duration</td>
<td>1.91</td>
<td>2.04</td>
<td>1.96</td>
<td>2.23</td>
<td>4.17</td>
<td>4.51</td>
<td>3.93</td>
<td>4.39</td>
</tr>
<tr>
<td>Avg. Coupon</td>
<td>3.32</td>
<td>4.80</td>
<td>4.69</td>
<td>2.93</td>
<td>3.12</td>
<td>3.92</td>
<td>3.80</td>
<td>1.32</td>
</tr>
<tr>
<td>Avg. Std. Dev. of Coupon</td>
<td>1.99</td>
<td>1.94</td>
<td>2.92</td>
<td>0.87</td>
<td>1.20</td>
<td>1.79</td>
<td>1.34</td>
<td>0.86</td>
</tr>
</tbody>
</table>

Table 2: **Principal Components of Relative Value Measures**
Share of the variations in a balanced panel of individual bonds' relative value measures explained by the first two principal components, in each country. The sample period is from 2005 to 2017.

<table>
<thead>
<tr>
<th></th>
<th>US</th>
<th>GB</th>
<th>CA</th>
<th>CH</th>
<th>DE</th>
<th>FR</th>
<th>IT</th>
<th>JP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>0.39</td>
<td>0.40</td>
<td>0.24</td>
<td>0.72</td>
<td>0.24</td>
<td>0.26</td>
<td>0.32</td>
<td>0.13</td>
</tr>
<tr>
<td>2nd</td>
<td>0.06</td>
<td>0.15</td>
<td>0.16</td>
<td>0.11</td>
<td>0.14</td>
<td>0.16</td>
<td>0.12</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Table 3: **International Relative Value Indices–Correlations**
Correlation between the daily relative value indices in each country. The sample period is from 2005 to 2017.

<table>
<thead>
<tr>
<th></th>
<th>US</th>
<th>GB</th>
<th>CA</th>
<th>CH</th>
<th>DE</th>
<th>FR</th>
<th>IT</th>
<th>JP</th>
</tr>
</thead>
<tbody>
<tr>
<td>US</td>
<td>1.00</td>
<td>0.80</td>
<td>0.53</td>
<td>0.64</td>
<td>0.60</td>
<td>0.15</td>
<td>0.03</td>
<td>-0.15</td>
</tr>
<tr>
<td>GB</td>
<td>0.80</td>
<td>1.00</td>
<td>0.59</td>
<td>0.52</td>
<td>0.44</td>
<td>0.16</td>
<td>-0.00</td>
<td>-0.10</td>
</tr>
<tr>
<td>CA</td>
<td>0.53</td>
<td>0.59</td>
<td>1.00</td>
<td>0.56</td>
<td>0.66</td>
<td>0.29</td>
<td>0.02</td>
<td>-0.31</td>
</tr>
<tr>
<td>CH</td>
<td>0.64</td>
<td>0.52</td>
<td>0.56</td>
<td>1.00</td>
<td>0.44</td>
<td>-0.11</td>
<td>-0.32</td>
<td>-0.35</td>
</tr>
<tr>
<td>DE</td>
<td>0.60</td>
<td>0.44</td>
<td>0.66</td>
<td>0.44</td>
<td>1.00</td>
<td>0.33</td>
<td>0.29</td>
<td>-0.20</td>
</tr>
<tr>
<td>FR</td>
<td>0.15</td>
<td>0.16</td>
<td>0.29</td>
<td>-0.11</td>
<td>0.33</td>
<td>1.00</td>
<td>0.62</td>
<td>0.21</td>
</tr>
<tr>
<td>IT</td>
<td>0.03</td>
<td>-0.00</td>
<td>0.02</td>
<td>-0.32</td>
<td>0.29</td>
<td>0.62</td>
<td>1.00</td>
<td>0.43</td>
</tr>
<tr>
<td>JP</td>
<td>-0.15</td>
<td>-0.10</td>
<td>-0.31</td>
<td>-0.35</td>
<td>-0.20</td>
<td>0.21</td>
<td>0.43</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 4: **International Relative Value Indices, Money-Market Rates and Volatility**
Correlation between the daily value of relative value indices with local or international measures of volatility or funding stress. Appendix A.1 details the additional data. The sample period is from 2005 to 2017.

<table>
<thead>
<tr>
<th></th>
<th>US</th>
<th>GB</th>
<th>CA</th>
<th>CH</th>
<th>DE</th>
<th>FR</th>
<th>IT</th>
<th>JP</th>
</tr>
</thead>
<tbody>
<tr>
<td>3M Local Interbank-OIS spread</td>
<td>0.71</td>
<td>0.66</td>
<td>0.34</td>
<td>0.55</td>
<td>0.36</td>
<td>0.17</td>
<td>0.25</td>
<td>-0.15</td>
</tr>
<tr>
<td>3M US LIBOR-OIS spread</td>
<td>0.71</td>
<td>0.62</td>
<td>0.27</td>
<td>0.58</td>
<td>0.28</td>
<td>0.04</td>
<td>0.05</td>
<td>-0.03</td>
</tr>
<tr>
<td>Volatility Index (local)</td>
<td>0.82</td>
<td>0.73</td>
<td>N/A</td>
<td>0.44</td>
<td>0.05</td>
<td>0.09</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>VIX</td>
<td>0.82</td>
<td>0.67</td>
<td>0.46</td>
<td>0.58</td>
<td>0.51</td>
<td>0.09</td>
<td>0.10</td>
<td>-0.07</td>
</tr>
</tbody>
</table>
Table 5: Relative Values as Proxy for Risks—Asset Pricing Tests
Cross-sectional asset pricing tests based on two-stage Fama-MacBeth regressions. The prices of risk are annualized (multiplied by 12). Standard errors and Shanken-corrected standard errors are reported in parentheses. The market returns is included as test assets at estimation. Monthly data, January 1986–December 2015.

Panel (A) Corporate bonds

<table>
<thead>
<tr>
<th></th>
<th>MKT</th>
<th>t-FM</th>
<th>t-Sh</th>
<th>RV</th>
<th>t-FM</th>
<th>t-Sh</th>
</tr>
</thead>
<tbody>
<tr>
<td>MKT</td>
<td>0.078</td>
<td>0.077</td>
<td>0.074</td>
<td>0.077</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-FM</td>
<td>(2.68)</td>
<td>(2.63)</td>
<td>(2.56)</td>
<td>(2.65)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-Sh</td>
<td>(2.68)</td>
<td>(2.60)</td>
<td>(2.54)</td>
<td>(2.64)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RV</td>
<td>-0.162</td>
<td></td>
<td></td>
<td>-0.162</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-FM</td>
<td>(3.02)</td>
<td></td>
<td></td>
<td>(2.94)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-Sh</td>
<td>(1.94)</td>
<td></td>
<td></td>
<td>(1.89)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noise</td>
<td></td>
<td>-0.063</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-FM</td>
<td>(1.38)</td>
<td></td>
<td></td>
<td>(0.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-Sh</td>
<td>(1.25)</td>
<td></td>
<td></td>
<td>(0.01)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ R^2 \]

Panel (B) S&P500 Options

<table>
<thead>
<tr>
<th></th>
<th>Call &amp; Put</th>
<th>Put</th>
<th>Call</th>
</tr>
</thead>
<tbody>
<tr>
<td>MKT</td>
<td>0.215</td>
<td>0.109</td>
<td>0.085</td>
</tr>
<tr>
<td>t-FM</td>
<td>(4.28)</td>
<td>(3.30)</td>
<td>(2.59)</td>
</tr>
<tr>
<td>t-Sh</td>
<td>(4.11)</td>
<td>(2.66)</td>
<td>(2.56)</td>
</tr>
<tr>
<td>RV</td>
<td>-0.169</td>
<td>-0.266</td>
<td>-0.310</td>
</tr>
<tr>
<td>t-FM</td>
<td>(4.26)</td>
<td>(5.44)</td>
<td>(3.27)</td>
</tr>
<tr>
<td>t-Sh</td>
<td>(2.74)</td>
<td>(2.53)</td>
<td>(1.34)</td>
</tr>
<tr>
<td>Noise</td>
<td>-0.171</td>
<td>-0.543</td>
<td>-0.185</td>
</tr>
<tr>
<td>t-FM</td>
<td>(3.74)</td>
<td>(4.71)</td>
<td>(2.62)</td>
</tr>
<tr>
<td>t-Sh</td>
<td>(2.43)</td>
<td>(1.21)</td>
<td>(1.62)</td>
</tr>
</tbody>
</table>

\[ R^2 \]
Table 6: **Summary Statistics for US Treasury Trade Returns (1988-2017)**

Panel (A) summary statistics for the cross-section of total trade returns, in weekly percentage terms. Panel (B) summary statistics for the time series of portfolio returns, in monthly percentage terms. Both panels present the results accounting for the bid-ask spread and use the baseline threshold for entering trades, $\tau = 0$.

### Panel (A) Cross-Section of Trade Returns

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>Std. Dev.</th>
<th># Trades</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative Value</td>
<td>0.10</td>
<td>-0.04</td>
<td>3.84</td>
<td>30030</td>
</tr>
<tr>
<td>Noise</td>
<td>-0.15</td>
<td>-0.11</td>
<td>9.23</td>
<td>30590</td>
</tr>
</tbody>
</table>

### Panel (B) Time-Series of Portfolio Returns

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
<th>Skew</th>
<th>Kurt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative Value</td>
<td>0.09</td>
<td>0.87</td>
<td>-5.72</td>
<td>5.10</td>
<td>-0.32</td>
<td>11.75</td>
</tr>
<tr>
<td>Noise</td>
<td>-0.19</td>
<td>0.47</td>
<td>-3.63</td>
<td>2.47</td>
<td>-0.34</td>
<td>16.00</td>
</tr>
</tbody>
</table>

Table 7: **Summary Statistics – Cross-Section of Trade Returns for All Countries (2005-2017)**

Mean, median and standard deviation in the cross-section of trade returns, from implementation to liquidation. Weekly returns, in percentage terms.

<table>
<thead>
<tr>
<th>Country</th>
<th>Relative Value</th>
<th>Mean</th>
<th>Median</th>
<th>Std. Dev.</th>
<th># Trades</th>
</tr>
</thead>
<tbody>
<tr>
<td>US</td>
<td></td>
<td>-0.02</td>
<td>-0.04</td>
<td>1.29</td>
<td>17429</td>
</tr>
<tr>
<td></td>
<td>Noise</td>
<td>-0.21</td>
<td>-0.09</td>
<td>1.45</td>
<td>17691</td>
</tr>
<tr>
<td>GB</td>
<td>Relative Value</td>
<td>0.77</td>
<td>0.04</td>
<td>3.50</td>
<td>909</td>
</tr>
<tr>
<td></td>
<td>Noise</td>
<td>0.76</td>
<td>0.05</td>
<td>2.86</td>
<td>931</td>
</tr>
<tr>
<td>CA</td>
<td>Relative Value</td>
<td>-0.05</td>
<td>-0.03</td>
<td>1.12</td>
<td>1408</td>
</tr>
<tr>
<td></td>
<td>Noise</td>
<td>-0.60</td>
<td>-0.22</td>
<td>1.41</td>
<td>2504</td>
</tr>
<tr>
<td>DE</td>
<td>Relative Value</td>
<td>1.00</td>
<td>0.04</td>
<td>4.13</td>
<td>2310</td>
</tr>
<tr>
<td></td>
<td>Noise</td>
<td>0.70</td>
<td>-0.01</td>
<td>5.19</td>
<td>2365</td>
</tr>
<tr>
<td>CH</td>
<td>Relative Value</td>
<td>1.34</td>
<td>0.06</td>
<td>5.24</td>
<td>373</td>
</tr>
<tr>
<td></td>
<td>Noise</td>
<td>0.66</td>
<td>0.03</td>
<td>3.97</td>
<td>518</td>
</tr>
<tr>
<td>FR</td>
<td>Relative Value</td>
<td>1.80</td>
<td>0.07</td>
<td>9.34</td>
<td>1191</td>
</tr>
<tr>
<td></td>
<td>Noise</td>
<td>1.74</td>
<td>0.09</td>
<td>8.85</td>
<td>1670</td>
</tr>
<tr>
<td>IT</td>
<td>Relative Value</td>
<td>3.20</td>
<td>0.08</td>
<td>12.47</td>
<td>3911</td>
</tr>
<tr>
<td></td>
<td>Noise</td>
<td>1.84</td>
<td>0.01</td>
<td>10.89</td>
<td>3275</td>
</tr>
<tr>
<td>JP</td>
<td>Relative Value</td>
<td>0.13</td>
<td>0.01</td>
<td>1.78</td>
<td>7191</td>
</tr>
<tr>
<td></td>
<td>Noise</td>
<td>0.09</td>
<td>-0.00</td>
<td>1.88</td>
<td>9402</td>
</tr>
</tbody>
</table>
Table 8: **Summary Statistics – Monthly Index Returns (2005-2017)**

Summary statistics for the time series of monthly portfolio returns, in percentage terms.

<table>
<thead>
<tr>
<th>Country</th>
<th>Relative Value</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
<th>Skew</th>
<th>Kurt</th>
</tr>
</thead>
<tbody>
<tr>
<td>US</td>
<td>Relative Value</td>
<td>0.08</td>
<td>1.00</td>
<td>-5.72</td>
<td>5.10</td>
<td>-0.44</td>
<td>13.61</td>
</tr>
<tr>
<td></td>
<td>Noise</td>
<td>-0.17</td>
<td>0.48</td>
<td>-3.63</td>
<td>2.47</td>
<td>-1.14</td>
<td>26.57</td>
</tr>
<tr>
<td>GB</td>
<td>Relative Value</td>
<td>0.37</td>
<td>1.00</td>
<td>-2.70</td>
<td>3.30</td>
<td>0.23</td>
<td>4.00</td>
</tr>
<tr>
<td></td>
<td>Noise</td>
<td>0.25</td>
<td>1.09</td>
<td>-2.79</td>
<td>4.57</td>
<td>1.15</td>
<td>5.81</td>
</tr>
<tr>
<td>CA</td>
<td>Relative Value</td>
<td>0.08</td>
<td>1.10</td>
<td>-3.47</td>
<td>3.41</td>
<td>-0.41</td>
<td>4.29</td>
</tr>
<tr>
<td></td>
<td>Noise</td>
<td>-0.31</td>
<td>0.64</td>
<td>-2.57</td>
<td>1.10</td>
<td>-0.70</td>
<td>4.39</td>
</tr>
<tr>
<td>DE</td>
<td>Relative Value</td>
<td>0.38</td>
<td>0.80</td>
<td>-2.31</td>
<td>2.59</td>
<td>-0.44</td>
<td>4.88</td>
</tr>
<tr>
<td></td>
<td>Noise</td>
<td>0.26</td>
<td>0.60</td>
<td>-1.14</td>
<td>2.39</td>
<td>0.97</td>
<td>4.91</td>
</tr>
<tr>
<td>CH</td>
<td>Relative Value</td>
<td>0.55</td>
<td>1.93</td>
<td>-6.58</td>
<td>11.37</td>
<td>1.69</td>
<td>15.16</td>
</tr>
<tr>
<td></td>
<td>Noise</td>
<td>0.40</td>
<td>2.41</td>
<td>-6.47</td>
<td>11.46</td>
<td>1.88</td>
<td>10.28</td>
</tr>
<tr>
<td>FR</td>
<td>Relative Value</td>
<td>0.86</td>
<td>1.92</td>
<td>-6.78</td>
<td>6.69</td>
<td>0.26</td>
<td>6.24</td>
</tr>
<tr>
<td></td>
<td>Noise</td>
<td>0.73</td>
<td>1.49</td>
<td>-4.86</td>
<td>6.84</td>
<td>0.18</td>
<td>6.15</td>
</tr>
<tr>
<td>IT</td>
<td>Relative Value</td>
<td>1.39</td>
<td>3.41</td>
<td>-2.47</td>
<td>31.81</td>
<td>5.83</td>
<td>47.14</td>
</tr>
<tr>
<td></td>
<td>Noise</td>
<td>0.99</td>
<td>1.87</td>
<td>-2.37</td>
<td>10.93</td>
<td>2.91</td>
<td>14.41</td>
</tr>
<tr>
<td>JP</td>
<td>Relative Value</td>
<td>0.49</td>
<td>0.98</td>
<td>-1.15</td>
<td>4.60</td>
<td>1.79</td>
<td>6.93</td>
</tr>
<tr>
<td></td>
<td>Noise</td>
<td>0.19</td>
<td>0.70</td>
<td>-0.93</td>
<td>3.60</td>
<td>1.98</td>
<td>9.11</td>
</tr>
</tbody>
</table>