

Research Statement

Gabriel E. Cabrera Guzmán 

August 2026

Research agenda

My research studies how information is generated, extracted, and incorporated into asset prices. I approach this question using empirical methods from asset pricing and financial econometrics, combining structured and unstructured data with machine learning and computational methods to extract information from both market prices and firms' communications. My research has two complementary directions that have increasingly converged during my work. The first examines the information contained in asset prices and the extent to which it can be extracted and used for investment decisions. The second examines how information is generated and communicated by firms and subsequently incorporated into prices by market participants.

The first direction is statistical. Given a large set of candidate predictors and flexible methods for searching through them, how much genuine predictability is there in asset returns and volatility, and is it large enough to matter to an investor? Seven published papers take up this question across commodities, equities, real estate, and international markets, while also analysing portfolio choice dynamics under return predictability. Their cumulative answer is more sceptical than the prevailing tone of the literature. The statistical gains are real but modest, they concentrate at short horizons, and they often fail to translate into economic value once the benchmark is chosen honestly and transaction costs are taken into account.

The second direction emerged from this finding. If predictability is difficult to extract from prices alone, a natural next question is where the information embedded in prices comes from in the first place. My current work therefore studies the process through which firms generate and communicate information, including when firms choose to speak, what they choose to say, and how quickly and completely markets absorb it. I focus on voluntary disclosure, specifically Analyst/Investor Days, where the timing of a release is a firm's own decision rather than a regulatory requirement.

Job market paper

When Managers Speak, Markets Listen: Evidence from Analyst/Investor Days

with O. Kolokolova (LUMS) and S. Sarah Zhang (AMBS) · 2026

Latest draft Aug 2026.

Setting and data. My job market paper studies how markets incorporate information during Analyst/Investor Days, a distinctive form of voluntary disclosure in which managers present strategy, performance, and outlook over several hours before taking questions. Unlike a conventional announcement, this setting allows us to observe prices adjusting while information is being released. I combine sentence-timestamped transcripts with tick-level trades and quotes for 2,490 U.S. events between 2014 and 2025, allowing me to examine not only when prices move, but how they adjust.

What we find. Price discovery is heavily front-loaded. Approximately 80% of the event-window return is incorporated before the Q&A begins, during the scripted presentation. The adjustment takes two distinct forms. Discrete price jumps respond to the content of what is said, particularly forward-looking and future-oriented statements, with jumps also more likely following negative-tone speech and when the CEO is speaking. Gradual price adjustments, identified as locally explosive drift bursts, respond instead to signed order-flow imbalance and are largely unrelated to language. Decomposing order flow further shows that retail trading has the strongest effect on prices while managers are speaking, whereas non-retail trading contributes more consistently throughout the event.

Why it matters. The distinction between these two forms of price adjustment is the paper’s central point. Both appear as rapid movements in aggregate returns, but they reflect different mechanisms. A jump captures the market incorporating the content of a particular statement, whereas a drift burst reflects information being absorbed through the trading process. Separating the two therefore changes how we understand price discovery at the minute horizon. It also provides within-event evidence for a corporate-governance prediction that has previously been studied mainly at daily frequency, namely that who speaks matters for how markets respond, and not only what is said.

Published research

The following published research comes from my work as a research assistant before my PhD.

Commodities. My earliest work asks whether flexible methods improve commodity price and volatility forecasts. *A Random Walk Through the Trees* evaluates random forests and gradient boosting against the regression trees used in prior work. The ensembles significantly outperform single trees, but a random walk remains competitive against all of them out-of-sample, which is the paper’s real finding and the source of its title. *Economic Drivers of Commodity Volatility* turns to copper volatility, using Bayesian Model Averaging to handle parameter and model uncertainty jointly. It finds that several economic variables forecast significantly relative to an autoregressive benchmark, with predictability varying across the business cycle. *Gold Risk Premium Estimation with Machine Learning Methods* extends the question to 186 predictors and reaches a sharper negative result. Machine learning methods and forecast combinations have limited ability to beat the historical mean. Technical indicators predict during expansions, while business-cycle and geopolitical variables predict during recessions, and once transaction costs are charged, investors should anticipate limited portfolio gains.

Equity and firm volatility. *Machine-Learning Stock Market Volatility* evaluates regularisation methods, tree ensembles, and neural networks against the HAR benchmark for S&P 500 realised volatility. Machine learning wins at the one-month horizon and loses at six and twelve months. Predictor quality matters more than algorithm choice, with financial and macroeconomic uncertainty proxies carrying most of the signal. From an economic perspective, however, the predictive models do not deliver substantial gains over the benchmark. *Forecasting the Volatility of US Oil and Gas Firms with Machine Learning* takes the question to firm-level realised volatility in the energy sector, where certain machine learning techniques do improve on HAR, and where the one-month Treasury bill and the aggregate VIX emerge as significant drivers.

Real estate and the macroeconomy. *Real Estate Returns and the Macroeconomy* applies tree-based methods to excess real estate returns in the US, Canada, and the UK, at one-month and two-year horizons. The drivers are heterogeneous across countries. Demand-side, labour-market,

and interest-rate variables are important in the US and Canada, while a composite leading indicator, unemployment, and economic policy uncertainty are more important in the UK. Their importance also varies substantially over time. Out-of-sample predictability is weak in the US and Canada and reasonably good in the UK, which we interpret as evidence that the North American markets are more informationally efficient.

Preferences. *Time-Varying Risk Aversion and International Stock Returns*, with Massimo Guidolin and Erwin Hansen, is where the economic-value test is passed most clearly. We estimate aggregate time-varying risk aversion from options, stock returns, and macroeconomic data for a panel of eight countries. The measure is counter-cyclical in most countries and forecasts monthly index returns up to twelve months ahead, and the effect survives controls for the variance risk premium, investor sentiment, and economic uncertainty. Critically, it is also tradeable. A mean-variance allocation built on the estimated measure delivers significantly positive returns.

Taken together, these papers reflect a consistent position. Flexible methods can uncover useful patterns in asset prices, but the more important question is whether those patterns capture economically meaningful information. This led me to look beyond the ability to predict prices and towards how information is generated, extracted, and incorporated into markets.

Ongoing Work and Future Agenda

My current research on voluntary disclosure in financial markets extends beyond the job market paper. During my PhD, I have also worked with my supervisors, O. Kolokolova and S. Sarah Zhang, on two related papers.

Product Market Competition and Voluntary Disclosure examines when firms choose to hold an Analyst/Investor Day at all. Firms are more likely to host one when current product market competition is low but the threat from potential rivals is high, a pattern we call competition timing. Events chosen this way are more informative. Firms emphasise innovation, forward outlooks, and firm value, and the market response is stronger, with a more positive speaker tone, higher short-term abnormal returns, improved liquidity, sustained increases in institutional ownership, and a subsequent fall in entry pressure from potential competitors. The paper shows that these events can serve as a tool of competitive positioning, rather than disclosure alone.

Save the Date: Analyst/Investor Days as a Trading Signal, moves from the event itself to its announcement. Prices react positively when an event is announced and continue to rise until it takes place. Firms that hype the event with additional disclosure beforehand see a larger run-up followed by a reversal, while non-hyped events are followed by a genuine positive reaction driven by both cash-flow and discount-rate news. A long-short strategy trading these patterns, implemented with a deep reinforcement learning algorithm, earns 0.13% per day.

Work in progress. *Why Is 1/N So Hard to Beat? Adaptive Learning and the Limits of Dynamic Portfolio Choice* revisits a persistent puzzle in portfolio choice. Even when investors can condition their portfolio decisions on market conditions, simple equal-weighted diversification remains remarkably difficult to outperform. The paper examines whether adaptive learning can overcome this limitation when investors must learn both which portfolio rule works and when to use it. I study this using a Q-learning agent that makes monthly portfolio choices based on market conditions and learns strictly online, without access to future information. This provides a natural framework for studying adaptive portfolio choice and the sequential decision-making relevant to algorithmic trading.

Miscalibrated or Informative? Asymmetric Forecast Uncertainty in Predictive Regressions for Stock Returns, with Juan Díaz and Erwin Hansen at the University of Chile, Faculty of Economics and Business, we examine the standard equity-premium prediction framework from a different angle. Rather than asking only whether predictors forecast returns, the paper asks whether the uncertainty around those forecasts is correctly measured and informative for investors. This is important because conventional predictive intervals rely on assumptions about the distribution of forecast errors that may be poorly suited to stock returns. We use conformal prediction to calibrate forecast intervals using past out-of-sample errors without imposing a distributional assumption, and allow the calibration to differ across the lower and upper tails, capturing both marginal and tail-specific coverage. We then examine whether this improved measure of forecast uncertainty contains information that matters for portfolio choice across predictors, horizons, and frequencies.

Where this goes. My future research will continue to focus on asset pricing, with a particular interest in how information affects prices and portfolio decisions. My earlier work studies the predictability and economic value of signals in asset returns, volatility, and risk. My current work extends this question to the information-generation process itself, using voluntary disclosure to study when firms release information, what they communicate, and how different investors incorporate it into prices. I plan to build on these settings to study price discovery, investor learning, and the economic value of information, while continuing to use modern empirical and computational methods to address these questions.

Research infrastructure

I make the empirical tools developed in my research available to other researchers. Two R packages are on CRAN, MIT-licensed, and developed openly on GitHub with documentation sites.

`sdim` implements supervised dimension reduction for asset pricing and macroeconomic forecasting. It provides methods for extracting factors from large sets of predictors or from panels with time-varying characteristics, together with tools for evaluating the predictive and economic value of the extracted factors.

`forecastdom` implements tests of out-of-sample predictive ability, evaluating whether one forecasting model genuinely outperforms another. It is organised around the taxonomy of Li, Liao and Quaadvlieg (2022, *Review of Economic Studies* 89: 843–875), which distinguishes between whether one model is merely equal to a competitor or genuinely superior, and whether that comparison is unconditional or conditional on the current state. These two distinctions give four cases, and the package implements all four, alongside tests for nested models, forecast encompassing, return predictability, and parameter instability.