Decomposing Geopolitical Risk: Wavelet-Based Time-Series Evidence and Cross-Sectional Implications for Expected Stock Returns

Davide La Cara*

July 2024

Abstract

The increasing concern surrounding geopolitical risk underscores its growing impact on contemporary markets. Leveraging the novel GPR index developed by Caldara and Iacoviello (2022), this paper constructs a new tradable factor based on historical betas, demonstrating its role as a priced state variable with a sizable monthly premium of 0.69% in the US equity market. Notably, this factor is among the few of its kind based on information orthogonal to economic indicators and explicitly linked to geopolitics. This cross-sectional factor analysis supports the main time-series findings, which are examined using an event-driven IV and a novel application of wavelet decomposition. While the former addresses inference concerns, the latter offers a dynamic lens into the frequency-domain behavior of geopolitical risk. It reveals that its effect on equity indices, such as NASDAQ, is negative and transitory. A behavioral interpretation grounded in recency, salience, and availability biases is also provided.

Keywords: wavelet multiresolution analysis, geopolitical risk, asset pricing.
JEL Codes: C22, G12, H56

*Department of Finance, London School of Economics. d.la-cara@lse.ac.uk

1 Introduction

Geopolitical risk has always been a key protagonist in history and politics. Unfortunately, recent events have reminded us once again with the war in Ukraine how also the present and the future are shaped by this particular source of uncertainty. As financial economists, the question that genuinely arises is how financial markets are influenced by this relatively unstudied risk and whether stock returns price in those shocks.

It feels almost natural and instinctive to think that geopolitical risk may be indeed disruptive for firms, entrepreneurs and more generally for the global economy through different channels such as supply chain disturbances, increased uncertainty, and loss of capital stock. However, despite the numerous narratives that could be developed on this topic, scant research has been conducted thus far. In fact, in the literature, the main limitation has always been posed by the difficulty in quantifying such a measure imposing as a consequence strict research boundaries. Attempts such as Dimic et al. (2015) and Diamonte et al. (1996) exploited for instance analyst estimates of political risk but did not have a fully coherent and unique measure yet. Then, Caldara and Iacoviello (2022) made a stark improvement by providing a credible index that would capture geopolitical risk as conveyed by newspapers by exploiting text analysis methods.

Building on their seminal work, literature started flourishing focusing mainly on boosting the forecasting power of previous models by enlarging them with GPR indexes such as GPR, GPRA and GPRT (i.e., respectively geopolitical risk, geopolitical risk actions and geopolitical risk threats). As an illustration, Ma et al. (2022) showed that those indices could help predict stock returns particularly during periods of expansions, whereas Nonejad (2022) suggested that there is no benefit in the out-of-sample forecast of returns volatility when augmenting previous models with the GPR time-series.

However, due to the nature of forecasting research, not much work has been done in terms of credible identification strategies. In particular, the literature has not focused on two key problems related to geopolitical risk which are measurement error and possible endogeneity (Born and Pfeifer, 2014) of this relatively unstudied source of uncertainty with stock returns. An attempt to deal with the latter issue has been made by Ha et al. (2022) who tried, with an IV-VAR approach, to estimate the impact of those shocks in the context of the Korean peninsula on South Korea's macroeconomic outcomes. Nevertheless, the external validity of their conclusions seemed to be limited to a special geo-economic setting that could not be used for general inference.

This paper provides a reasonable alternative to reduce the measurement error and endogeneity related to the geopolitical risk index by exploiting part of the information used in Ha's study. In particular, I assess the impact of geopolitical risk as conveyed by newspapers on the US equity market showing that the negative impact of such shocks is limited in time suggesting a behavioral interpretation of investors' reaction. Indeed, the impact on returns will be limited in time reverting within a few months to its original level despite the new information related to the higher probability of disaster (Barro, 2006). The approach followed in this study thus hinges upon two different pillars: instrumental variable and wavelet multiresolution analysis. The former deals with measurement error, whereas the latter helps assessing the transitory feature of these shocks.

As per first pillar, as a basis for the instrument, I use a dummy representing North Korea's missile tests. The idea is that (almost) every time North Korea's dictator shoots a missile into the sea, he is trying to show his new technological advancement: a new weapon that may do more damage or, even more, something whose range of action is larger than before. Each time, the probability of disaster should be updated by investors factoring in the new technological power of North Korea. Crucially for the story, this military improvement is considered to be permanent with no possibility to revert back. This reasoning is sound, as it is improbable for an army to lose technological knowledge once it has been attained.

Then, to complete the IV construction, I interact this dummy with the number of Google searches made by US citizens relating to topics such as 'Nuclear war'. The intuition is that this makes the IV capture all the ballistic tests the public is aware of. Later in the analysis, we will see that the dummy is a good predictor of those Google searches enforcing this view.

As per second pillar, wavelet multiresolution analysis is exploited to try to study the different impacts of this source of uncertainty on different frequency components of NASDAQ's returns. I employ this relatively new method in the economic literature, used mainly for forecasting purposes (e.g., Rua, 2011), to prove that the effect on returns of geopolitical shocks captured by the instrument will be transitory. In short, I show how the impact can be divided into the low persistency, high-frequency component being negative (that is, immediate negative effect) and the high persistency, low-frequency part being positive (i.e., a reverting trend to the initial state).

Concerning the theoretical framework, I provide a simple story to convey the intuition. Imagine living in Los Angeles and being aware that North Korea's missile range is 2000 miles. Therefore, you cannot be reached by a possible attack. In this simple example, the probability of disaster in such an infrequent and high-impact event is zero. Then, the dictator shoots a missile into the sea with a range of 10000 miles. Now, you update your information set and the probability of being hit: it moves from zero to a strictly positive number. This probability, no matter what will happen later, may change but cannot fully revert to zero as this technological advancement will not revert as well. Consequently, according to 'rational pricing', it should also be reflected in prices by showing a permanent, though even small, decline.

However, despite the prior given by the information hypothesis, this is not supported by the evidence. The effect is transitory and a behavioral interpretation, conflicting with the rational view, is needed. Behavioural theory suggests possible solutions to this puzzle. In particular, investors, biased by availability and recency heuristics, may just not factor anymore into prices the information they received a few months earlier. Within this framework, I provide a behavioural hypothesis, consistent with the empirical evidence, of an initial overreaction that is later followed by a full reversal in returns.

The time-series analysis so far described thus shows the impact of geopolitical risk on financial indexes such as the NASDAQ. However, one may think that such an aggregate effect might be reflected not just in time but also in the cross-section of the US stock market. To foster and, in fact, prove this hypothesis, a cross-sectional analysis is run in parallel with the time-series one of the above paragraphs. The idea is that the two analyses are complementary and enforce one another. The existence of a geopolitical effect on the index in time should leave cross-sectional evidence and the presence of cross-sectional equity premia due to geopolitical risk should lead to time-series fluctuations in returns as well. Moreover, previous literature did not focus on the impact of geopolitical risk on the cross-section expanding the scope and the contribution of this paper.

Indeed, to my knowledge, none of the previous papers tried to answer the question of whether geopolitical risk was actually priced as a factor in the US equity market. The closest attempt by Cakici et al. (2020) relies on forming 'high minus low' portfolios containing stocks from countries with the highest variation in their national GPR. However, this method has the limitation of not sorting on individual stocks but on country-level risk weakening the possibility of tailoring the factor portfolio. The approach in this study precisely overcomes this issue and I provide a new factor capturing geopolitical risk that is priced in the cross-section of the equity market.

The related literature is massive and finds its roots in the seminal paper by Fama and French (1993). Later, Harvey et al. (2015) cataloged many anomalies that could surge as potential factors in asset pricing models and Feng et al. (2020) even listed the members of the so-called 'factor zoo' reaching an impressive number of 150 candidates. Crucial for the interest of this paper is also to realize that all 150 factors in that list are directly related to some economic concept or variable such as profit margin (Soliman, 2008) or earnings volatility (Francis et al., 2004). None of them is non-strictly related to financial data. However, in this work, a factor fully orthogonal to this type of information is provided; to my knowledge, one of the few. Indeed, I try to capture the sensitivity of the stock market to geopolitical risk with a procedure similar to the one used for liquidity in Pastor and Stambaugh (2003) building a beta-sorted portfolio that should resemble this non-economically driven risk.

In particular, each year, I sort stocks by historical betas coming from regressing the geopolitical risk index of threats, consistently with the forecasting literature that signals this to be the one with the highest predictive power, on the excess returns of CRSP stocks. This procedure generates spreads between high betas and low betas portfolios leading to a 'high minus low' type of factor. The t-stat of the time series mean of the factor will be significant and positive throughout the sample that goes from 1990 to 2023, monthly data.

Moreover, to test whether the factor is priced, consistently with Fama and French (2018), I compare the nested models of Fama-French five-factor (Fama and French, 2015) plus Momentum (Jegadeesh and Titman, 1993) with and without the geopolitical risk factor. In short, regressing my factor on the six factors, the constant plays the key role: a significant constant would suggest a priced factor as it is the case. Various robustness checks are provided such as changing the frequency of sorts (e.g., monthly instead of annually) or adding and deleting controls in the regressions from which the historical betas are coming, and then testing the new factors again. Results are robust.

Previous paragraphs may suggest that the main contribution of the paper is methodological, that is, developing a new interpretative framework in the context of wavelet analysis matched with an IV to solve previously unaddressed measurement error and endogeneity. Actually though, this framework shows also how the market behaves not in a fully rational way with regards to this source of risk. Moreover, the further evidence and its relative size provided by the cross-sectional analysis show how critical it is not to neglect geopolitical risk and how crucially we need, as I try here, to quantify the impact, the investors' behavior and the market performance in periods of high geopolitical uncertainty.

Overall, the two sub-studies are complementary providing a 360-degree overview. They are both time-series and cross-sectional evidence of the impact of geopolitical risk in the equity market. Each leg has its own robustness checks but together they are also a clue of the presence of the other thus enforcing the view that geopolitical risk has a marked and sizeable impact on stocks.

In short, this paper makes several contributions to the geopolitical risk literature. First, it quantifies the cross-sectional size of the GPR factor showing that is far from being negligible (i.e., 0.69% monthly). Second, it shows how to exploit a wavelet framework (Rua, 2011) to prove in a IV context whether an effect is going to be temporary thus providing a new interpretative tool for multiresolution analysis. Finally, it solves previously unaddressed problems related to the GPR literature such as measurement error and endogeneity (Caldara and Iacoviello, 2022; Ma et al., 2022). [1]

The study unfolds as follows. Section 2 presents the geopolitical index exploited in the analysis and more generally the data used. Section 3 explains in detail the methodology employed to assess the impact of GPR on equity returns. Section 4 presents the results and the robustness checks. Finally, Section 5 concludes. Time series and cross-sectional analysis will run in parallel.

[1] Several other papers contributed to the existing literature. For instance, Drakos (2004) exploits the 9/11 shock on airline stocks. Yang and Yang (2021) use mixed-frequency GPR to forecast stock market returns. Hoque et al. (2020) study nonlinear effects of GPR uncertainty on stock returns. Das et al. (2019) examine whether emerging equity markets react to geopolitical stress. Glick and Taylor (2005) delve into the effect of wars on international trades. Monge et al. (2023) evaluate the impact of GPR on oil prices. Jung et al. (2021) use North Korea GPR index to show how corporate stocks react to national shocks.

2 Data

This study has two data frames. The time-series part is based on a sample period going from 2004 to 2023, monthly data; whereas the cross-sectional data goes from 1990 to 2023, monthly data. For the former, it is a period of twenty years that encompasses various periods of geopolitical tensions (e.g., Ukraine's war). Central in this study is the geopolitical risk index provided by Caldara and Iacoviello (2022) which should represent and quantify those periods of uncertainty. Indeed, the goal of this variable is to try to capture the perceived geopolitical risk conveyed by mass media such as newspapers. In their paper, they show with a VAR how their index may be negatively linked to macroeconomic variables like investment and employment.

The way they calculate this index exploits a news-based approach, that is, they estimate the number of newspaper articles linked to geopolitical risk. The method is close to the one used earlier by Baker et al. (2016) for the economic policy uncertainty index (EPU). In particular, they build dictionaries of words (e.g., 'war', 'terrorism', and 'weapon') for which they look for their occurrence in articles. Those collections of words rely on their definition of GPR; that is: "as the threat, realization, and escalation of adverse events associated with wars, terrorism, and any tensions among states and political actors that affect the peaceful course of international relations".

On a sample of about 25 million news articles published in English, they construct three main indices differentiating among them by using a different dictionary every time. In short, the authors split the original main dictionary into two subcategories: acts and threats. This leads to the three variables: GPR, GPR-A, and GPR-T. The first one encompasses the words used in the other two, the second focuses on geopolitical acts (e.g., terrorist acts, beginning or escalation of wars), and the latter on geopolitical threats (e.g., military buildup, nuclear and terrorist threats). Consistently with the forecasting literature, I use GPR-T since the threat index appears to be the one with the highest impact on the predictability of variables of economic interest (Ma et al., 2022). GPR data can be found on GPR's website. A summary statistics table can be found in the *Appendix*.

Concerning the IV data, I construct a list of the months in which a Korean missile has been shot by looking into CNN and Bloomberg articles. Clearly, this list is affected by measurement error but in the context of this analysis, this is not a major issue. Indeed, if a missile test has not been communicated massively by media (and it is therefore difficult to find traces of it), then it should be irrelevant for the purpose of this study and its absence from the dummy should not impact in any way the analysis. The Google searches instead are downloaded by Google Trends setting the parameter to 'US searches' and monthly data. A summary statistics table can be found in the *Appendix*.

In addition to the above, the analysis is mainly carried on the NASDAQ monthly returns but for robustness checks, I also estimate the same for the SP500 and the value-weighted CRSP index. The former is downloaded by the FRED website, whereas the latter comes from CRSP (WRDS). The analysis predominantly focuses on the NASDAQ index, despite similar findings being observed with the SP500 and CRSP indices. This choice is justified by the rationale that major multinational corporations with extensive global supply chains, such as the prominent technology firms listed on NASDAQ, are likely to be among the first to experience the impact of geopolitical tensions and conflicts.

CRSP's data (WRDS) is also used to represent the US equity market in the cross-sectional part. In particular, I use the excess returns of CRSP's stocks to build the traded factor resembling geopolitical risk. Those returns are in excess compared to the 1-month T-bill (from Ibbotson and Associates). Cross-sectionally, I have more than 29,000 distinct stocks in the sample. Considering only the stocks listed in the market in any given month, the sample shrinks but it still allows proper estimation and portfolio formation (i.e., minimum of three-hundred stocks per portfolio). Moreover, also the number of outstanding shares and their relative price are used to assess market capitalization for the formation of the value-weighted portfolios.

Finally, for testing, I exploit Fama and French's five factors augmented with Momentum downloaded from Kenneth R. French's website. The five factors (Fama and French, 2015) are: market (i.e., excess returns on the market portfolio), SMB (i.e., Small minus Big: size effect), HML (i.e., High minus Low: BE/ME effect), RMV (i.e., Robust minus Weak: profitability) and CMA (i.e., Conservative minus Aggressive: investment). Moreover, consistently again with Fama and French (2018), I will be adding also Momentum (i.e., the tendency of winners to remain winners in the short run) to their five-factor model.

3 Methodology: time-series and cross-sectional analysis

In the first half of this section, I study whether NASDAQ's returns are linked to geopolitical risk as measured by Caldara's threat index. First, I discuss the wavelet multiresolution method and the IV assumptions. Then, I show the estimation procedure and present the robustness checks. Furthermore, in the second half of this chapter, I show how the geopolitical risk factor is constructed and the procedure used to test whether it is priced in the equity cross-section.

3.1 *Time series: Wavelet multiresolution analysis and the IV*

Wavelet multiresolution analysis

The wavelet multiresolution method is a statistical technique that allows to decompose a time series into orthogonal components differentiated by their frequency level. It can be thought of as a Fourier transform with two key aspects that make it diverge from the well-known transformation: finite energy and compact support (Rua, 2011). That can be beneficial when analyzing financial time series affected by heteroskedasticity, discontinuities and volatile trends because of its more flexible approach.

Therefore, using the DWT, discrete wavelet transform, (Percival and Walden, 2000) instead of the Fourier seems natural in the context of financial time series. In particular, with this method, I split the original NASDAQ's returns series into high-frequency components resembling the short-term dynamics and low-frequency components representing the long-term behavior. Later, those different layers will be the dependent variables on which I study the impact of geopolitical risk at different horizons.

In detail, the DWT relies on two kinds of wavelets: 'mother' wavelets ψ and 'father' ϕ wavelets. The former captures the high-frequency components, whereas the latter the long-run smooth part of the time series. The wavelet approximation has the following form:

$$r(t) = \sum_k S_{J,k}\phi_{J,k}(t) + \sum_k d_{J,k}\Psi_{J,k}(t) + \sum_k d_{J-1,k}\Psi_{J-1,k}(t) + ... + \sum_k d_{1,k}\Psi_{1,k}(t) \tag{1}$$

That is:

$$r(t) = S_J(t) + D_J(t) + D_{J-1}(t) + ... + D_1(t) \tag{2}$$

Where J is the number of multi-resolution scales and k goes from one to the number of coefficients in the respective component. S_J and D_j are the smooth and the details, respectively (each capital letter summarizes a summation part of equation *(1)*). Notation is consistent with Rua (2011).

Overall, this definition of the wavelet transform allows for the mapping from the time domain of the time series (i.e., the original) to the timescale domain. Indeed, *(2)* is the decomposition of the returns *r(t)* into its orthogonal components: details and smooth capturing the higher frequencies and lower frequency waves, respectively. An important side note for what will follow in the time-series part of the analysis is to recall that high frequency components tend to be less persistent and lower frequency ones have the opposite tendency.

IV assumptions and estimation

The Instrumental variable setting is the following. The dependent variables to be studied are the details of the wavelet decomposition (i.e., NASDAQ's returns), the independent variable is the Caldara and Iacoviello GPRT index and finally, the instrument is the interaction of Google searches with the missiles' dummy.

The first issue that the IV has the goal to solve is measurement error. Geopolitical risk is a very broad concept that is difficult to quantify in time. Caldara's attempt has provided a key starting point but it may still suffer from incomplete choices of dictionaries or by relying too heavily on only one source of information, that is, newspapers. For this reason, I use the interaction part given by Google searches; i.e., a different system of information. Having different media sources should at least decrease the measurement error since the noises in both variables are less likely to be correlated. For this reason, I do not use other GPR indexes such as GPRA as an instrument to solve for measurement errors. The error in the latter would be systematically correlated with the noise in GPRT given the same source of information and the same type of method. For this reason, I exclude this possibility and I exploit Google searches consistently with Ha et al. (2022).

Below, I discuss instead the assumptions related to inference using the instrumental variable *Z(t)*. First of all, I show the first stage, that is:

$$GPRT(t) = \delta + \alpha Z(t) + \epsilon(t) \quad Where \;\; Z(t) = GT(t) * DummyM(t) \tag{3}$$

Where $Z(t)$ is the IV, $GT(t)$ stands for the Google searches from Google Trends and $DummyM(t)$ is the missiles' dummy. The first stage is econometrically robust with high F-tests and reasonably high R^2 (without indulging into overfitting). However, it may give rise also to possible doubt in terms of the IV being as good as random.

Indeed, a fear that might arise staring at equation *(3)* is that maybe the IV and GPRT are both driven by a general state of uncertainty, a sentiment of fear and escalation that would both drive newspapers and Google searches. Similar concern may be coming from thinking that perhaps both variables are actually driven by attention. To rule this out, I estimate twice the first stage, once as in equation *(3)* and once as below:

$$GPRA(t) = \delta + \alpha_* Z(t) + \epsilon(t) \tag{4}$$

The intuition goes that if attention or a general state of uncertainty were the driving forces of the IV, then we should obtain a strong first stage result also for equation *(4)*; we will see this will be far from evidence in the data. This is because also GPRA, as GPRT, is constructed on newspaper information that could potentially be driven by attention or by a general sentiment of fear fuelled by media.

Moreover, to bring further evidence in favor of the IV being as good as random, I regress the instrument on four economic policy uncertainty indexes as provided by Baker et al. (2016): global, American, South Korean and Japanese. Those are the ones that are related to countries with which North Korea may be politically more involved and may look at when shooting a missile. Japan and South Korea for proximity and obvious reasons, whereas the US is their main ally and, most of all, is the country whose returns are analyzed in this study.

Furthermore, missile tests are sparse across the full sample and are subject to constraint that seems to be unrelated to markets. In particular, those tests are subject to technological and physical restrictions that cannot reasonably be overcome quickly. Additionally, those tests are also linked to internal appraisal as the regime may use them to show its achievement and its power; again, another logic unrelated to the outside world. Finally, the shooting of missiles is decided ultimately by a dictator with a questionable decision-making process.

I present below the structural equations of this analysis.

$$r(t) = \delta + AR(k) + \beta GPRT(t) + \eta(t), \tag{5}$$

$$D_j(t) = \delta_j + AR(k) + \beta_j GPRT(t) + \eta_j(t) \quad with \quad j = 1, ..., J, \tag{6}$$

$$S_J(t) = \delta_J + AR(k) + \beta_J GPRT(t) + \eta_J(t), \tag{7}$$

where *AR(k)* stands for the autoregressive component, *r(t)* are the NASDAQ's returns, $D_j(t)$ are the details, and $S_J(t)$ the smooth. Those structural equations should capture the time-series impact of geopolitical risk as conveyed by media on a broad stock market index such as the NASDAQ. In the results section, a behavioural interpretation based on this evidence is provided.

Given those structural equations, the reason for the wavelet method can finally be made clear (apart from econometrical benefits). With equation *(6)*, I obtain the impact of GPR on the details of different frequencies. Thanks to their orthogonality and linearity, I can back out the overall impact on the return series decomposed by frequency level. In a way, it can be thought of as a partial derivative. Intuitively, recalling equation *(2)*:

$$\frac{\partial r(t)}{\partial GPRT(t)} \approx \frac{\partial S_J(t)}{\partial GPRT(t)} + \frac{\partial D_J(t)}{\partial GPRT(t)} + ... + \frac{\partial D_1(t)}{\partial GPRT(t)} \tag{8}$$

Then, we can see the differentiated impact that GPRT can have on returns at different frequencies. Recalling that high frequencies are less persistent and that low frequencies are more persistent, we see that on the former there is a negative impact and on the latter a positive one. This signals already at time *(t)* that the effect will later revert as the only force left after the disappearance of the high-frequency impact will be the upwardly persistent one. This interpretative framework of wavelet is novel. This overall transitory effect is later confirmed also by projections on returns and cumulative returns at *(t+h)*.

The validity of the exclusion restriction needed for the inference based on the above equations is less cumbersome to show compared to the previous assumption. This stems from the broad and encompassing nature of geopolitical risk. It is reasonable to assume that missile tests—serving primarily as signals of military escalation—impact stock returns only through their effect on geopolitical risk. Absent this uncertainty channel, a missile test is merely a metal object falling into the sea. As for the autoregressive component in each regression, attributing it to the timing of missile tests would imply that the dictator conditions test launches on past stock returns—an implausible assumption. Moreover, if markets were somehow anticipating these events (again, unlikely), any resulting bias would attenuate the estimated effect, rendering the results conservative.

Finally, I present here the reduced form.

$$D_j(t) = \delta + AR(k) + \theta_j Z(t) + \iota(t). \tag{9}$$

That is, this equation captures the impact of the missile tests directly on the returns. The structural equations exploiting GPRT rescale this equation adding also the attempt to solve measurement error. Moreover, the choice of this particular IV is crucial for external validity as well. The geopolitical threat perceived and diffused by newspapers coming from those ballistic experiments is not bound to a single region (motivating the choice of 'Nuclear war' as the main interaction), to a single industry, or a particular supply chain. It is about a general risk of a huge and unlikely event with macroscopic consequences. That is why this instrument in terms of external validity offers more compared to a regional crisis that may only impact some supply chains, some industries, and not others, etc. The risk at stake would encompass and touch everything and everyone making it a reasonable choice of 'broadly speaking' geopolitical risk proxy already in the reduced form of equation *(9)*. Of course, the choice of using the US stock returns (NASDAQ, SP500 and CRSP) is linked to this external validity concern. It is not just a local impact for a local context driven mainly by proximity.

Finally, the estimation is carried out via GMM, Newey-West standard errors (1987) robust to heteroskedasticity and autocorrelation.

Robustness checks

Various robustness checks are performed to show that results are indeed robust to different specifications, interactions, and indexes.

First of all, I do the very same analysis as discussed in the previous section on both the SP500 and the value-weighted index of CRSP showing that coefficients remain overall significant and of similar magnitude. This confirms the idea that geopolitical risk is impacting the US equity market and that results do not depend specifically on a chosen market index.

Then, to check that results are not relying too heavily on the choice of words for the Google search interaction, I do the analysis also using 'North Korea', 'Republic of Korea', and 'Korea' in addition to the main 'Nuclear war'. Also, I regress the Google searches of the former on the dummy showing how missile tests are predictive as expected of how much US citizens look for that on Google. Moreover, as a last check for 'Nuclear war', I regress the time series of those searches on both EPU global and EPU USA obtaining non-significant coefficients and low R^2.

To have stationary low-frequency variables, differencing is needed. To avoid issues related to the endogeneity of the autoregressive parts, I use an Anderson-Hsiao (1982) type of estimator truncating at the second lag for the instrument. For robustness, I change the k in the *AR(k)* obtaining almost unchanged results proving that estimates are robust to the time-series specification.

Furthermore, I do the entire analysis multiple times by changing every time the wavelet family. In particular, this is to rule out data mining. Wavelet decomposition requires a judgment call in terms of the choice of wavelet family. Every family has a wavelet with a slightly different shape. In the main analysis, I use the classic Daubechies least-asymmetric wavelet with four vanishing moments and periodic boundary handling. I do everything also starting with a multiresolution based on the same family but adding or subtracting one vanishing moment. I move then to do the same analysis with the Haar symmetric wavelet, with the classic Daubechies, the Fejér-Korovkin, and the Coiflet wavelets. Results are unaltered.

A final concern is the potential for look-ahead bias stemming from estimating the wavelet decomposition over the full sample. To address this, I re-estimate the decomposition using a rolling procedure updated quarterly. The resulting slopes (unreported for brevity) remain consistent in both sign and magnitude, suggesting, as expected, that the wavelet transformation does not introduce forward-looking bias.

Finally, to confirm the transitory effect already signaled by the low-frequency positive impact, I use the cumulative returns in equation *(5)* and show that the effect vanishes within a few months. I do the same also on the log of cumulative returns, cumulative abnormal returns and log of cumulative abnormal returns and obtain similar results again.

3.2 *Cross-section: Building and testing the Geopolitical factor*

Constructing the factor

Following the steps of the liquidity factor portfolio formation (Pastor and Stambaugh, 2003), I use a similar approach here. The sensitivity of every stock to geopolitical risk is represented by its GPR-T beta, that is, the slope coefficient of a regression whose dependent variable is stock i excess return and whose independent variables are GPR-T index and other factors considered to be relevant by previous literature (e.g., Fama French three-factor model).

In particular, at the end of every year, with a five-years estimation sample, I regress returns on GPR-T and store the betas associated with geopolitical risk in a vector. Then, sorting on those betas, I select the top and bottom deciles related to high and low betas reproducing the different sensitivities of stocks to GPR. This gives rise to sufficient dispersion in geopolitical betas. Given this sorting, I select returns starting at the end of the beta estimation period (i.e., post-formation returns) of those deciles and construct my two main portfolios of high and low betas. Both value-weighted (i.e., weighted based on the stock's market capitalization) and equal-weighted portfolios are estimated based on those returns.

Concerning the regression from which betas are extracted, there are three different specifications. Each specification hinges upon a different set of control variables that play a key role in the factor literature.

$$r_{i,t} = \delta_i + \beta_i^{GPRT} GPRT_t + \epsilon_{i,t}, \tag{10}$$

$$r_{i,t} = \delta_i + \beta_i^{GPRT,3} GPRT_t + \theta_i^{m} MKT_t + \theta_i^{smb} SMB_t + \theta_i^{hml} HML_t + \eta_{i,t}, \tag{11}$$

$$\begin{aligned} r_{i,t} = \delta_i &+ \beta_i^{GPRT,5} GPRT_t + \theta_i^{m} MKT_t + \theta_i^{smb} SMB_t + \theta_i^{hml} HML_t \\ &+ \theta_i^{rmw} RMW_t + \theta_i^{cma} CMA_t + \iota_{i,t} \end{aligned} \tag{12}$$

Where $r_{i,t}$ are monthly excess returns compared to the 1-month T-Bill of asset i, MKT is the excess return of a market portfolio, *SMB* is Small minus Big, *HML* denotes High minus Low, *RMW* and *CMA* are Robust minus Weak and Conservative minus Aggressive respectively.

The first regression *(10)* has no control and uses only GPR-T as independent variable. The second specification adds Fama and French three factors, that is, Market, HML and SMB. Finally, the latter adds all Fama and French five factors as controls. Pastor and Stambaugh (2003) use specification *(11)* for their historical betas to form their liquidity portfolio. Accordingly, I use $\beta_i^{GPRT,3}$ as the main candidate to represent the geopolitical risk factor. By augmenting the specification with controls, $\beta_i^{GPRT,3}$ captures the sensitivity of GPR-T unrelated to the co-movement of other financial variables.

In addition to the above six factors representing geopolitical risk (i.e., one value-weighted and one equally weighted for each of the three specifications), I estimate also factors changing the sorting frequency. In the main specification, consistent with previous literature, I sort on a twelve-month basis with a five-year rolling window for the estimation. Here, as a robustness check, I also estimate value-weighted factors with a monthly frequency of sorts. In short, in the former case, I sort every month and store for the portfolio formation one-month postformation return for each i. Crucially for the implementability of all those factors, the portfolio formation procedure uses only data that was available as of the time of formation.

All this estimation procedure, consistent with previous literature, assumes that the betas are relatively constant in time, at least when considering a moderately short horizon as it is the case here. Nevertheless, in the context of geopolitical risk, this seems to be reasonable as a company i, which is likely to be exposed more than a company j today, is also likely to be more exposed tomorrow. Geopolitical risk exposure can rarely change on an overnight basis, it is more of a slow 'buildup' process that takes some time to compound. Also, we will see later that changing the frequency of estimation and sort does not change the results enforcing this view.

However, to take a further step in the direction of less noisy and more stable betas, I also build one factor based on stocks sorted on significant betas. In short, I sort yearly on significant (90%) betas whose standard errors have been estimated with the Newey-West approach to be robust to heteroskedasticity and autocorrelation using specification *(11)*. The drawback of such an estimation procedure is that it considerably shrinks the number of assets in a portfolio as, clearly, the number of significant betas is strictly smaller than the number of betas.

Testing the factor

Once the factor is estimated, the question is to understand whether it is priced in the cross-section. The outcome of the above procedure is a time series vector of portfolio returns that alone cannot say much about cross-sectional variation in stocks' excess returns. However, before even moving to the cross-sectional tests, I estimate the mean and the t-stat for all my factors resembling geopolitical risk (plural because of the multiple specifications with alternative frequencies of sorts and controls). This is just enough to suggest whether the series of returns is made of noise or whether the equity premium of the geopolitical factor is significantly different from zero. In the results section, I show that it is significant and positive. Moreover, to verify that the variation among deciles is significant, I estimate all the pairwise decile difference t-stats highlitghing that there is enough variation among their returns.

As in Pastor and Stambaugh (2003) and consistently with the more recent Fama and French (2018), I test my value-weighted factor sorted yearly (with betas estimated with the three-factor model as control; i.e., sorting on $\beta_i^{GPRT,3}$) by regressing this geopolitical factor on Fama-French five-factor model augmented with Momentum. The idea is that if the constant is significantly different from zero, then the proposed factor will contribute (in-sample) to explain average returns. In addition to this, I apply this very same procedure also to the Fama-French three-factor model and the classic CAPM. As a robustness check, this analysis is also carried on all other factors estimated in the previous section (i.e., geopolitical risk factors whose formation hinges upon different sorting frequencies or different controls; e.g., β_i^{GPRT} and $\beta_i^{GPRT,5}$). Newey-West's (1987) standard errors, robust to heteroskedasticity and autocorrelation, are used for regression confidence intervals.

Finally, as a last step, I also regress each single decile portfolio based on the yearly sort, three controls factor (i.e., $\beta_i^{GPRT,3}$) on the six-factor model. By storing the alphas of each of those regressions, the GRS test (Gibbons, Ross and Shanken, 1989) is performed to assess whether those ten alphas are significantly different from zero. Crucially, the key advantage of this test lies in the fact that it allows to test whether those alphas are jointly non-zero. This test strongly rejects the zero null.

In short, in this testing section, I heavily rely on what Fama and French (2018) defined the RHS (right-hand-side) approach which indeed consists of spanning proposed factors on known ones. They argue that this is the best solution to test nested models as it is the case here. Moreover, the alternative is the so-called LHS (left-hand-side) method (i.e., comparing models based on their intercepts from regressions of a set of portfolio returns), which has the relevant constraint that considerably depends on the set of test assets chosen. It seemed thus natural to follow the RHS as described above.

4 Results and robustness checks

In this section, I present the results and some of the robustness checks. In particular, in the first half the focus is on time-series (i.e., wavelet analysis) results, whereas in the second it is on the cross-sectional outcome. The remaining robustness checks are left in the *Appendix*.

4.1 *Time-series Results*

First of all, I show that the instruments, regardless of the choice of words, are strong and robust. Below, in *Table (1)*, I report the first stage. We can easily see looking at the F-tests that in all cases we are strictly above ten. As per the instruments coming from the interaction with Google searches, the lowest F-test is a comforting 16.64 which reassures that we are not in a weak-IV setting. Coefficients are all 99% significant.

Table 1: First Stage Regressions

	(1) GPRT	(2) GPRT	(3) GPRT	(4) GPRT
DummyXgtNW	2.394***			
	(0.340)			
DummyXgtK		0.851***		
		(0.204)		
DummyXgtNK			0.587***	
			(0.122)	
Missiles dummy				28.26***
				(8.901)
Constant	94.22***	99.03***	101.5***	100.5***
	(1.889)	(2.024)	(2.376)	(2.596)
F-Test	49.55	16.64	23.10	10.08
R-squared	0.486	0.079	0.131	0.096

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: DummyXgtNW stands for the IV made of the interaction between the dummy and the Google Trend searches of 'Nuclear war', DummyXgtK and DummyXgtNK similarly for 'Korea' and 'North Korea', respectively.

Moreover, as anticipated with equation *(4)*, I estimate the very same first stage substituting the dependent variable with GPRA instead of GPRT (*Table (2)*). In particular, comparing the R^2 and the F-test statistics, we can see a drop, column by column, between the two tables. In the case of 'Nuclear war', the F-test goes from above 40 to below 7, similarly the R^2 drops from above 40% to 6%. This shows clearly that the instrument is not affected by a general build-up that might have been driven by the interaction of Google searches. If that had been the case, we should have not seen a stark difference between the above two regressions. Furthermore, this checks also for potential differences in the attention of market participants. Both indices are built on newspapers news so they should both be affected by this volatile attention; however, the impact of the IV on the first stage is strong only on one of the two indices showing how this potential change in attention/sensitivity is not the driver of the instrument. Thus, this confirms that the main channel of our instrumental variable is missile tests that the public is aware of.

Table 2: Comparison – First Stage

	(1)	(2)	(3)	(4)
	GPRA	GPRA	GPRA	GPRA
DummyXgtNW	0.820**			
	(0.316)			
DummyXgtK		0.106		
		(0.126)		
DummyXgtNK			0.0589	
			(0.101)	
Missiles Dummy				-13.426*
				(7.889)
Constant	91.10***	93.67***	94.03***	110.4***
	(2.515)	(2.504)	(2.447)	(5.962)
F-Test	6.75	0.70	0.34	2.90
R-squared	0.063	0.001	0.000	0.004

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Where DummyXgtNW stands for the IV made of the interaction between the dummy and the Google Trend searches of 'Nuclear war', DummyXgtK and DummyXgtNK similarly for 'Korea' and 'North Korea', respectively.

Then, I move to the main result of the study coming from structural equations *(5), (6)* and *(7)*. In *Table (3)*, I report the impact of geopolitical risk as measured by the threat index provided by Caldara and Iacoviello (2022). Looking at columns (3) and (4), the effect is negative in the short/medium trend frequency and positive in the long-term one (5). There is no effect either on noise (as expected, Column (2)) or on the long-run smooth (Column 6) thus fading away. The regression on returns (Column (1)) indicates a negative impact at time *(t)*. This is consistent with above findings as the magnitude of the short/medium-term frequency is higher in absolute value compared to lower frequency paths. So at time *(t)*, taking and summing partial derivatives with respect to GPRT (equation *(8)*), we get a negative effect overall as confirmed by Column (1). However, with the long-term trend kicking with time, the effect will revert, vanishing.

Table 3: Regression Results: main decomposition

	(1)	(2)	(3)	(4)	(5)	(6)
	$r(t)$	D_1	D_2	D_3	D_4	S_4
GPRT	−7.28e-05**	−1.08e-05	−6.17e-05*	−6.88e-06**	2.03e-06***	2.05e-08
	(3.70e-05)	(8.48e-06)	(3.41e-05)	(3.50e-06)	(7.70e-07)	(1.25e-08)
AR(1)	0.204**	−0.627***	0.362***	0.822***	0.939***	0.516*
	(0.0895)	(0.0397)	(0.0412)	(0.0179)	(0.0201)	(0.266)
Constant	0.0103**	0.00114	0.00625*	0.000723*	−0.000203**	−2.06e-06
	(0.00492)	(0.000885)	(0.00358)	(0.000435)	(8.62e-05)	(1.38e-06)
R-squared	0.045	0.387	0.053	0.676	0.894	0.213

Newey-West (1984) standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Where *r(t)* stands for NASDAQ's returns, D_j for the details of the wavelet decomposition, with low j resembling high-frequency/low-persistency trends and high j representing the low-frequency/high-persistency one. S_4 is the long-run smooth.

As I confirm later in *Table (5)*, the impact on returns will not last longer than a few months, that is, the pricing of the probability of disaster's increase is not going to be permanent. This would be consistent with rational investors as long as we assume the probability of disaster to decrease and reach the same level as before the test. However, the interpretation of the instrument is not complacent with this view. In fact, every test, signaling an improvement in military technology, should increase the probability of disaster by a part that cannot fully revert.

Nevertheless, returns do come back within a few months to their initial path thus imposing a behavioral interpretation. This evidence is consistent with two possible hypotheses. First, investors start pricing rationally the higher probability of disaster reflected in lower prices. Then, in accordance with behavioral theories, the result of heuristics such as salience and recency biases (Bordalo et al., 2012) kicks in leaving outside of investors' information set the update of a few months ago related to the probability of disaster.

The second possibility, consistent also with Cakici et al. (2020) findings, begins with an initial overreaction of investors, a well-documented phenomenon linked with price reversal (De Bondt, 1989). Investors overreact, then realize that they went too far leading to a rational partial price reversal. However, the price reversal also goes too far leading to full inversion, again due to recency and salience biases. Instead, the overreaction would be ascribed to the availability bias (Kahneman and Tversky, 1974) fuelled by the media coverage. This interpretation becomes even more reasonable once read in the light of studies such as Sustein and Zeckhauser (2011) where they show how overreaction is fostered by dreadful risks, such as a missile test, that encourage emotional reactions of anxiety.

Recalling that also the instrument may suffer (if alone) from measurement error, I focus for the rest of the study on the structural equation; however, for completeness, I report below the reduced forms. Results are consistent with what was expected given the main table above. Negative impacts on the high-frequency details and positive on the low-frequency ones. Overall, negative on returns at time *(t)*.

Table 4: Regression Results for the Reduced form

	(1)	(2)	(3)	(4)
	$r(t)$	D_2	D_3	D_4
$Z(t)$	-0.0001766^{*}	$-8.63\text{e-}05^{*}$	$-1.67\text{e-}05^{*}$	$4.24\text{e-}06^{**}$
	(0.0000945)	(4.88e-05)	(9.37e-06)	(1.67e-06)
AR(1)	0.2007833^{**}	0.352^{***}	0.820^{***}	0.933^{***}
	(0.0888734)	(0.0387)	(0.0232)	(0.0190)
Constant	0.0035275	0.000221	8.13e-05	$-3.34\text{e-}06$
	(0.0015329)	(0.000264)	(0.000156)	(3.11e-05)
R-squared	0.049	0.127	0.673	0.895

Newey-West (1984) standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Where *r(t)* stands for NASDAQ's returns, D_j for the details of the wavelet decomposition, with low j resembling high-frequency/low-persistency trends and high j representing the low-frequency/high-persistency one.

Lastly, I bring further evidence, as already signaled by the positive reversion of the low-frequency detail, that the effect reverts and vanishes within a few months. Below, in *Table (5)*, I report the projections of the IV (equation *(5)*) on cumulative returns on the NASDAQ. Depending on the specifications (i.e., whether log or whether adding more *AR(k)* parts), we see the impact fading and becoming non-significant within three months. Coefficients start as significant and they decrease in absolute value and lose statistical relevance with time.

Table 5: Regression Results on cumulative returns

	(1) $cr(t)$	(2) $cr(t{+}1)$	(3) $cr(t{+}2)$	(4) $cr(t{+}3)$
GPRT	-0.000139^{***}	$-8.82\text{e-}05^{***}$	$-6.70\text{e-}05$	$-4.23\text{e-}05$
	(3.79e-05)	(2.85e-05)	(7.29e-05)	(5.26e-05)
AR(1)	0.133^{*}	0.153^{**}	0.164^{**}	0.172^{**}
	(0.0683)	(0.0651)	(0.0654)	(0.0675)
Constant	0.0176^{***}	0.0118^{***}	0.00940	0.00715
	(0.00526)	(0.00412)	(0.00774)	(0.00567)
R-squared	0.025	0.051	0.070	0.036

Newey-West (1984) standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: Where *cr(t+h)* stands for NASDAQ's cumulative returns.

Robustness checks

First of all, I do the whole analysis by substituting NASDAQ with SP500 and the value-weighted index of CRSP. Here, I report the SP500 results (*Table (6)*), in the *Appendix* the counterpart for the CRSP index can be found. As we can see, coefficients are significant and they maintain the same size as in the NASDAQ case. Negative in the D_2 and D_3, positive in D_4 with the same order of magnitude. Results are thus robust to changing market indexes.

Table 6: Regression Results for the SP500 decomposition

	(1) $SP500 - D_2$	(2) $SP500 - D_3$	(3) $SP500 - D_4$
GPRT	$-6.25\text{e-}05^{***}$	$-1.47\text{e-}05^{**}$	$2.97\text{e-}06^{***}$
	(2.26e-05)	(6.48e-06)	(6.63e-07)
AR(1)	0.336^{***}	0.828^{***}	0.943^{***}
	(0.0407)	(0.0219)	(0.0203)
Constant	0.00586^{**}	0.00145^{*}	-0.000280^{**}
	(0.00228)	(0.000790)	(0.000111)
R-squared	0.101	0.686	0.888

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: D_j stands for the details of the wavelet decomposition of SP500, with low j resembling high-frequency/low-persistency trends and high j representing low-frequency/high-persistency ones.

Moreover, results are robust to changing words for the interaction, that is, results do not rely on the choice of 'Nuclear war'. In *Table (7)*, I report results using 'Korea' as an interaction on the decomposition of NASDAQ's returns. Coefficient remain of the same sign and order of magnitude as above.

Table 7: Regression Results changing the IV wording

	(1)	(2)	(3)
	D_2	D_3	D_4
GPRT	−6.04e-05*	−5.92e-06	5.47e-06***
	(3.32e-05)	(6.87e-06)	(1.18e-06)
AR(1)	0.362***	0.818***	0.953***
	(0.0421)	(0.0174)	(0.0222)
Constant	0.00612*	0.000613	−0.000576***
	(0.00339)	(0.000779)	(0.000131)
R-squared	0.056	0.676	0.880

Newey-West standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: D_j stands for the details of the wavelet decomposition of NASDAQ.

To show the robustness of the structural equation I report in *Table (8)* the same results as in *Table (3)* by changing/adding controls or by changing the Anderson and Hsiao *AR(k)* part. In particular, here I paste the results obtained by adding EPU USA as a control. Results are relatively unvaried signaling that the specification is robust to changes in the structural equations.

Table 8: Regression Results with a different control

	(1)	(2)	(3)
	D_2	D_3	D_4
GPRT	−6.02e-05*	−7.82e-06	2.22e-06***
	(3.35e-05)	(3.82e-06)	(9.07e-07)
AR(1)	0.361***	0.831***	0.916***
	(0.0465)	(0.0163)	(0.0273)
EPU USA	−9.96e-07	4.35e-06	1.72e-07
	(8.94e-06)	(2.72e-06)	($4.13x10^{-7}$)
Constant	0.00624**	0.000199	−0.000246**
	(0.00313)	(0.000408)	(0.000113)
R-squared	0.056	0.676	0.880

Newey-West standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: D_j stands for the details of the wavelet decomposition of NASDAQ.

Finally, in *Table (9)*, I report the results obtained by doing the whole analysis on NASDAQ decomposition with a different wavelet family. In particular, I move from the Daubechies least-asymmetric to the Coiflet wavelet. This is to rule out data mining related to the choice of the wavelet family. Results are stable.

Table 9: Regression Results: Coiflet Wavelet Decomposition

	(1) coif_D_2	(2) coif_D_3	(3) coif_D_4
GPRT	$-6.43e-05^{**}$	$-6.88e-06^{***}$	$2.12e-06^{**}$
	(3.27e-05)	(2.07e-06)	(1.02e-06)
AR(1)	0.379^{***}	0.824^{***}	0.929^{***}
	(0.0385)	(0.0165)	(0.0174)
Constant	0.00647^{*}	0.000743^{***}	-0.000194^{*}
	(0.00338)	(0.000278)	(0.000107)
R-squared	0.072	0.684	0.899

Newey-West (1984) standard errors in parentheses

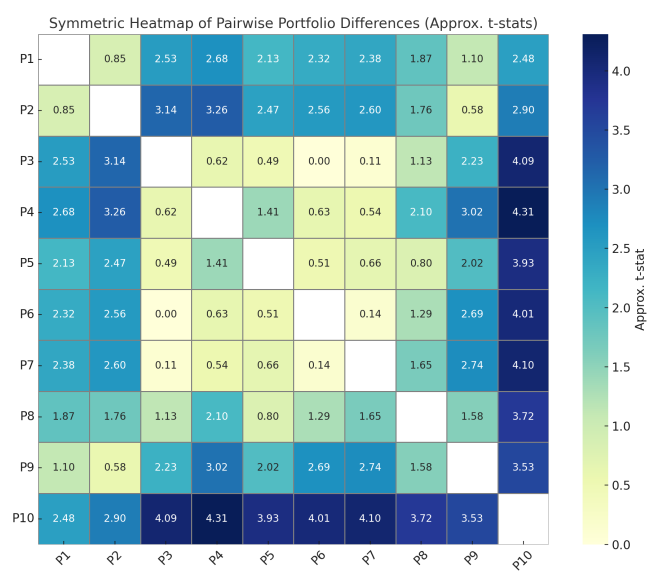
*** p<0.01, ** p<0.05, * p<0.1

Note: D_j denotes the details of the wavelet decomposition estimated using the Coiflet family for NASDAQ returns.

4.2 *Cross-section Results*

In this sub-section, I report the results referred to the main representative candidate (i.e., when not specified, that is 'the factor') for the geopolitical risk factor following Pastor and Stambaugh (2003), that is, the yearly sorted high minus low beta portfolio coming from a three-factor control regression (i.e., yearly formation, sorting on $\beta_i^{GPRT,3}$). Moreover, as per the robustness check, I carry the same analysis also on GPR factors built with different sorting frequencies and/or different controls. Finally, I split the sample into two halves and see if I get results consistent with the full sample.

First of all, in the *Figure* below, I show that there is significant variation among deciles given the high variation in betas. In particular, this table displays the t-statistics of the test on the difference between all pairs of portfolio returns based on the beta sorting (i.e., $\beta_i^{GPRT,3}$). As expected, there is significant variation in the first two columns and the bottom two rows, that is, high or low beta deciles are significantly different from others. The relative absence of statistical significance among the central deciles can be easily explained by the fact that those are the deciles whose betas are closer to zero, that is, deciles of stocks that are not considerably sensitive to geopolitical risk.



Note: This table reports the t-statistics of the differences among decile portfolio returns of the value-weighted, three controls and yearly sorted geopolitical risk factor.

This table shows that variation among deciles that are heavily relying on this source of uncertainty is significant. Moreover, we can see that the t-stat of the 'high minus low' portfolio (i.e., *dec10 minus dec1*) is 95% significant reaching a comforting level of 2.48. Similar results are obtained when we look at yearly sorted factors whose betas have been generated by a different regression (unreported). In particular, to provide some examples, for the value-weighted factor coming out of the five-factor model (i.e., $\beta_i^{GPRT,5}$), I obtained a t-stat of 2.32, whereas for the one coming out of the same three-factor controls but sorted only on *significant* betas the result is 2.53. Surprisingly, also the equally-weighted portfolios remain significant with a t-stat of 2.21 and 1.98 for the three and five controls respectively. Overall, all those 'long high beta and short low beta' portfolios have a positive and significant equity premium that ranges from a monthly 0.69% (i.e., a substantial yearly 8.3%) of our main candidate factor (i.e., the time-series mean return of the 'high minus low' portfolio) to a more moderate 0.33% of the equally weighted one.

Being assured of enough variation among deciles and significantly non-zeros equity premia, it remains to verify whether the geopolitical risk factor is priced in the cross-section. Regressing every decile on the Fama-French five factors model plus Momentum, I obtain all significant alphas with t-statistics ranging from 3.03 to a remarkable 6.08, of the sixth and eighth deciles respectively. I perform the GRS test to see whether those alphas are jointly different from zero and I can reject the null with 99% confidence. Splitting the sample in two halves, results are unchanged with the t-stat varying from 1.90 to 5.64 for the sixth and first deciles respectively in the first part, whereas for the second half the lowest t-stat reaches 2.30.

Then, I analyse the factor alongside other substitutes of the main candidate and I show that the result of the former does not hinge upon the estimation criteria given robust results also for the latter. I regress the 'High minus low' geopolitical factor again on Fama-French five factors plus Momentum. Results are shown in *Table (10)*. Confidence levels do not fall below 90% for all the six different factors. The alphas seem significant and are always strictly positive. There appears to be evidence of a priced factor.

Table 10: Alphas

GPR factors	(1) Vw_Yr	(2) Vw_Mr	(3) Vw_Yr_ff3	(4) Vw_Mr_ff3	(5) Vw_Yr_ff5	(6) Vw_Mr_ff5
Alpha	0.00671*	0.00817*	0.00586*	0.00461**	0.00533**	0.00489**
	(0.00396)	(0.00451)	(0.00351)	(0.00231)	(0.00260)	(0.00236)

Newey-West (1984) standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Where Vw means value-weighted, Yr yearly sorted, Mr monthly sorted, ff3 and ff5 indicate the number of controls in the beta estimation phase. Alphas are from regressions of proposed factors on the Fama-French five factors model plus Momentum.

In addition to the above check, consistent with Pastor and Stambaugh (2003), I perform the same analysis also on Fama-French three factors and on CAPM focusing on three and five control estimation procedures. In *Table (11)*, I again obtain significant and positive alphas when using the three factors model as a benchmark. The same applies to the CAPM case, where alphas remain positive and generally significant (*Table (12)*).

Table 11: Alphas

GPR Factors	(1) Vw_Yr_ff3	(2) Vw_Mr_ff3	(3) Vw_Yr_ff5	(4) Vw_Mr_ff5
Alpha	0.00673**	0.00475**	0.00582**	0.00485*
	(0.00319)	(0.00228)	(0.00282)	(0.00255)

Newey-West (1984) standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Where Vw means value-weighted, Yr yearly sorted, Mr monthly sorted, ff3 and ff5 indicate the number of controls in the beta estimation phase. Alphas are derived from regressions of proposed factors on the Fama-French three-factor model.

Table 12: Alphas

GPR Factors	(1) Vw_Yr_ff3	(2) Vw_Mr_ff3	(3) Vw_Yr_ff5	(4) Vw_Mr_ff5
Alpha	0.00678**	0.00493*	0.00575*	0.00478*
	(0.00339)	(0.00276)	(0.00303)	(0.00279)

Newey-West (1984) standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Where Vw means value-weighted, Yr yearly sorted, Mr monthly sorted, ff3 and ff5 indicate the number of controls in the beta estimation phase. Alphas from the regressions of proposed factors on the CAPM.

Finally, as robustness check, I show in *Table (13)* the results coming from regressing the significant beta factor (i.e., sorting on significant $\beta_i^{GPRT,3}$, yearly formation) on the Fama and French five-factor model plus Momentum, the three-factor model, and the CAPM. Also in this case, alphas are positive and significant.

Table 13: Alphas

Vw_Yr_ff3_sig	(1) FF5+Mom	(2) FF3	(3) CAPM
Alpha	0.0102**	0.0110**	0.0108*
	(0.00520)	(0.00547)	(0.00566)

Newey-West (1984) standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Where Vw means value-weighted, Yr yearly sorted, and *sig* implies sorting on significant betas. *ff3* indicates the number of controls in the beta estimation phase. Alphas are from regressions of the proposed factors on the five-factor model plus Momentum (FF5+Mom), the three-factor model (FF3), and the CAPM, respectively.

Overall, there seems to be enough evidence to support the hypothesis that geopolitical risk is priced in the cross-section of stock returns. Alphas are positive and significant. Notably, they remain significant for the geopolitical factor regardless of whether tested on the Fama-French five-factor model with Momentum, the three-factor model, or the CAPM. The equity premium of the factor is positive and significantly different from zero across the all thirty years sample.

5 Conclusion

Geopolitical risk has reminded us with recent events what a major and impactful source of uncertainty it can be. A blooming literature, starting with Caldara and Iacoviello (2018, 2022), is trying to study this relatively novel risk through the research's lenses, mainly attempting to improve returns forecastability. The approach and contribution that this study chooses to assume is instead twofold. On one hand, I try to assess the time series impact of geopolitical risk on an aggregate index such as the NASDAQ and I show how this effect is limited in time in contrast with EMH prediction. On the other hand, this study verifies whether geopolitical risk is priced in the cross-section of the US stock market by building and testing a new factor.

This two-leg analysis has the advantage of providing a comprehensive overview of the geopolitical impact on stock returns both in a time-series and cross-sectional perspective. One investigation can be considered further evidence of the other. The existence of a trace in time of GPR's influence should leave trails in the individual stocks cross-section and the opposite applies as well. That is, cross-sectional evidence of GPR's may be time-varying thus resulting in a time-series footprint. In this paper, I indeed find evidence of both thus enforcing and strengthening the relevance of GPR's uncertainty vertically in time and quantifying it horizontally up to a sizable 0.69% monthly return.

Firstly, for the time-series leg, this paper uses wavelet decomposition combined with an instrumental variable approach. The former has mainly been used in the literature for forecasting purposes (Rua, 2011) so the interpretative framework I provide of wavelet is also novel to this study. Instead, the latter tries to address measurement error and possible endogeneity of geopolitical risk on NASDAQ's returns.

In particular, the instrument employed is the interaction between Google searches in the US of topics such as 'Nuclear war' and a dummy being one every time North Korea shoots a missile into the sea. Many missile attempts have the goal of showing power and new technologies. Once one range has been achieved, from there the reachable distance can only monotonically increase. Therefore, we should expect this to impact directly the probability of a geopolitical disaster. This probability cannot fully revert back, no matter what will happen later. Once North Korea achieves the technology to strike your position, the probability moves from zero to positive. This should be negatively priced in; in particular, the information hypothesis teaches that this negative effect should, at least in part, be permanently priced in returns given the, at least partial, permanent change in the probability.

Nevertheless, despite our prior, we observe a transitory effect on returns. This is not in accord with a rational story though. Most likely, investors update their information leading to a negative return on the month of the strike, maybe even overreacting coherently with Cakici et al. (2022). After that, returns slowly revert back in a few months undoing the possible overreaction and canceling from prices the information gained a few months earlier. Behavioural heuristics such as availability, recency and salience may underpin this violation of the EMH as discussed in the results section.

Secondly, for the cross-sectional leg, this study chooses also to see whether there is a geopolitical risk factor that could help explain stock market returns. To my knowledge, this is the first paper that tests its existence in the US. Consequently, this paper is entering also the literature of factor models. Crucially, this is one of the few factors that does not rely on economic concepts.

Therefore, after having built the GPR factor, I show that it is priced in the cross-section of stock returns. In particular, I assess the factor to have a positive and significantly nonzero equity premium. Given the positive and significant alphas both in the main candidate factor and in the relative robustness checks, there is evidence that suggests that geopolitical risk is indeed substantial and sizable for stock excess returns thus underpinning also the time-series findings.

Overall, it seems that geopolitical risk has an impact on financial markets and that investors do not respond in a fully rational way to this type of shock. However, from a policy point of view, this seems to be paradoxically good news: markets tend to be resilient to those shocks (e.g., Korean war 1950-1953). Nevertheless, further research should be done to understand how no-flight to quality may occur making the US market even better-off during those periods of uncertainty. Studying international capital flows may thus be an interesting direction in which to extend future studies.

6 References

Anderson, T.W., Hsiao, C., 1982. Formulation and estimation of dynamic models using panel data. Journal of Econometrics 18. 47-82

Baker, S., Bloom, N., Davis, S.J., 2016. Measuring Economic Policy Uncertainty, The Quarterly Journal of Economics. Vol 131. 4 1593–1636.

Barro, R., 2006. Rare disasters and asset markets in the twentieth century. The Quarter Journal of Economics 121, no. 3: 823-866.

Born, B. and Pfeifer, J. 2014. Policy risk and the business cycle. Journal of Monetary Economics, Vol.68, 68–85.

Cakici, N., Demir, E., Long, H., Zaremba, A., 2020. When bad news is good news: Geopolitical risk and the cross-section of emerging market stock returns. J. Fin. Stability 58. 100964.

Caldara, D., Iacoviello, M., 2018. "Measuring geopolitical risk.". Int. Financ. Discuss. Pap., No. 1222, Febr. 2018. Board Gov. Fed. Reserve Syst. https://doi.org/10.17016/IFDP.2018.1222.

Caldara, D., M. Iacoviello, and A. Markiewitz. 2018. "Country-Specific Geopolitical Risk." Mimeo.

Caldara, D., Iacoviello, M., 2022. Measuring geopolitical risk. American Economic Review 112 (4), 1194-1225.

Carhart, Mark M., 1997, On persistence in mutual fund performance, Journal of Finance 52, 57–82.

Cochrane, J., 2011. Presidential address: Discount rates, Journal of Finance 66, 1047–1108.

Das, D., Kannadhasan, M., Bhattacharyya, M., 2019. "Do the emerging stock markets react to international economic policy uncertainty, geopolitical risk and financial stress alike?". North Am. J. Econ. Financ. 48, 1–19.

De Bondt, W.F.M., 1989. "Stock Price Reversals and Overreaction to News Events: A Survey of Theory and Evidence.". In: Guimaraes, ˜ R.M.C., Kingsman, B.G., Taylor, S.J. (Eds.), A Reappraisal of the Efficiency of Financial Markets. NATO ASI, vol 54. Springer, Berlin, Heidelberg.

Diamonte, R.L., Liew, J.M., Stevens, R.L., 1996. "Political risk in emerging and developed markets.". Financ. Anal. J. 52 (3), 71–76.

Dimic, N., Orlov, V., Piljak, V., 2015. "The political risk factor in emerging. Front., Dev. Stock Mark. " Financ. Res. Lett. 15, 239–245.

Dimic, N., Orlov, V., Piljak, V., 2016. "The effect of political risk on currency carry trades.". Financ. Res. Lett. 19, 75–78.

Drakos, K., 2004. "Terrorism-induced structural shifts in financial risk: airline stocks in the aftermath of the september 11th terror attacks.". Eur. J. Political Econ. 20 (2),435–446.

Fama, E., French, K., 1993. Common Risk Factors in the Returns on Stocks and Bonds. J. Financial Econ. 33. 3–56.

Fama, E., French, K., 1996. Multifactor explanations of asset pricing anomalies. J. Finance. 51, 55–84.

Fama, E., French, K., 2010. Luck versus skill in the cross-section of mutual fund returns. J. Finance. 65, 1915–1947.

Fama, E., French, K., 2012. Size, value, and momentum in international stock returns. J. Finance. Econ. 105, 457–472.

Fama, E., French, K., 2015. A five-factor asset pricing model. J.F.E. 116, 1–22.

Fama, E., French, K., 2016. Dissecting anomalies with a five-factor model. Rev. Financial Stud. 29, 69–103.

Fama, E., French, K., 2017. International tests of a five-factor asset pricing model. J. Finance Econ. 123, 441–463.

Fama, E., French, K., 2018. Choosing factors. J. Financ. Econ. 128, 234-252

Fama, E., MacBeth, J., 1973. Risk, return, and equilibrium: empirical tests. J. Polit. Econ. 81 (3), 607–636.

Feng, G., Giglio, S., Xiu, D., 2020. Taming the Factor Zoo: A test of new factors. J. Finance. 75. 1327-1370.

Fernandez V., 2005. The international CAPM and a wavelet-based decomposition of value at risk. Studies in Nonlinear Dynamics and Econometrics 9(4): Article 4.

Francis, J., LaFond, R., Olsson, P., Schipper, K., 2004. Costs of equity and earnings attributes, The Accounting Review 79, 967–1010.

Gibbons, Michael R.; Ross, Stephen A.; and Shanken, Jay. "A Test of the Efficiency of a Given Portfolio." Econometrica 57 (September 1989): 1121–52.

Giglio, S., Liao, Y., Xiu, D., 2018. Thousands of alpha tests, Technical report, University of Chicago.

Glick, Reuven, and Alan M. Taylor. 2010. "Collateral Damage: Trade Disruption and the Economic Impact of War." Review of Economics and Statistics 92 (1): 102–27.

Ha, J., Lee, S., So, I., 2022. The Impact of Uncertainty Shocks: Evidence from Geopolitical Swings on the Korean Peninsula. Oxford Bulletin of Economics and Statistics. 84.1. 0305-9049.

Hoque, M.E., Zaidi, M.A.S., 2020. "Global and country-specific geopolitical risk uncertainty and stock return of fragile emerging economies.". Borsa Istanb. Rev. 20 (3), 197–213.

Jegadeesh, N., and Titman, S., 1993 Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. J. Finance 48. 65–91.

Jung, SeungHo and Lee, Jongmin and Lee, Seohyun. 2021 The Impact of Geopolitical Risk on Stock Returns: Evidence from Inter-Korea Geopolitics. Bank of Korea WP 2021-10

Ma, F., Lu, F., Tao, Y., 2022. Geopolitical risk and excess stock returns predictability: New evidence from a century of data. Finance Res. Letters. 50. 103211.

Newey, W. K., West, K. D., 1987. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. Econometrica, 55(3), 703–708.

Nonejad, N., 2022. An interesting finding about the ability of geopolitical risk to forecast aggregate equity return volatility out-of-sample. Fin. Research Let. 47. 102710

Pastor, L., Stambaugh, R., 2003. Liquidity risk and expected stock returns, ˇ Journal of Political Economy 111, 642–685.

Percival, D., Walden, A., 2000. Wavelet Methods for Time Series Analysis. Cambridge University Press: Cambridge, UK.

Rua, A., 2011. A Wavelet Approach for Factoraugmented forecasting. Journal of Forecasting. 30. 666-678.

Saiz, Albert, and Uri Simonsohn. 2013. "Proxying for Unobservable Variables with Internet Document-Frequency." Journal of the European Economic Association 11 (1): 137–65.

Soliman, T., 2008. The use of Dupont analysis by market participants, The Accounting Review 83, 823–853.

Sustein, C.R., Zeckhauser, R., 2011. "Overreaction to fearsome risks.". Environmental Resour. Econ. 48, 435–449.

Taylor, S., Thompson, S., 1982. "Stalking the elusive "vividness" effect. Psychol. Rev. 89 (2), 155–181.

Tversky, A., Kahneman, D., 1973. "Availability: a heuristic for judging frequency and probability.". Cogn. Psychol. 4, 207–232.

Tversky, A., Kahneman, D., 1974. "Judgment under uncertainty: heuristics and biases.". Science 185, 1124–1131.

Valta, P., 2016. Strategic default, debt structure, and stock returns, Journal of Financial and Quantitative Analysis 51, 197–229.

Welch, I., 2008. The link between Fama-French time-series tests and Fama-Macbeth cross-sectional tests, Technical report, UCLA.

Yang, J., Yang, C., 2021. The impact of mixed-frequency geopolitical risk on stock market returns, Economic analysis and Policy Dec-21 226-240.

Zaremba, A., 2016. Risk-based explanation for the country-level size and value effects. Financ. Res. Lett. 18. 226-233.

7 Appendix

Table 14: Descriptive Statistics

	Mean	Std. Dev.	Skewness	Kurtosis
GPRT	105.49	45.59	2.197	15.968
NASDAQ ret	0.00199	0.0236	-1.335	6.605
Missile Dummy	0.177	0.382	1.691	3.862
GT: Nuclear War	11.821	9.735	5.334	40.521
GT: North Korea	8.543	13.163	4.871	29.699

Note: NASDAQ ret stands for NASDAQ's monthly returns (i.e., $r(t)$), and *GPRT* represents the Geopolitical Risk Index by Caldara and Iacoviello (2022). The *Missile Dummy* equals one when a missile is launched. *GT* refers to Google Trends search volumes for the terms "Nuclear war" and "North Korea."

Table 15: Regression Results with one more Vanishing Moment

Details	(1) D_2	(2) D_3	(3) D_4
GPRT	$-7.28\text{e-}05^{***}$	$-6.93\text{e-}06^{**}$	$2.06\text{e-}06^{**}$
	(2.27e-05)	(3.29e-06)	(8.10e-07)
AR(1)	0.370^{***}	0.823^{***}	0.941^{***}
	(0.0403)	(0.0176)	(0.0201)
Constant	0.00743^{***}	0.000753^{*}	-0.000205^{**}
	(0.00224)	(0.000403)	(8.98e-05)
R-squared	0.034	0.679	0.896

Newey-West (1984) standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: D_j stands for the details of the wavelet decomposition estimated using the same wavelet family for NASDAQ returns as in the main analysis but adding a vanishing moment to the wavelet. Coefficients remain significant and similar in magnitude.

Table 16: Regression Results with one less Vanishing Moment

Details	(1) D_2	(2) D_3	(3) D_4
GPRT	$-6.88\text{e-}05^{***}$	$-7.50\text{e-}06^{**}$	$2.16\text{e-}06^{**}$
	(2.14e-05)	(3.39e-06)	(8.55e-07)
AR(1)	0.354^{***}	0.819^{***}	0.912^{***}
	(0.0410)	(0.0187)	(0.0271)
Constant	0.00702^{***}	0.000788^{*}	-0.000216^{**}
	(0.00211)	(0.000426)	(9.18e-05)
R-squared	0.020	0.671	0.889

Newey-West (1984) standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: D_j stands for the details of the wavelet decomposition estimated using the same wavelet family for NASDAQ returns as in the main analysis equation *(8)* but subtracting also a vanishing moment to the wavelet. Coefficients remain significant and similar in magnitude.

Table 17: Regression Results using Fejér-Korovkin Wavelet Family

Details	(1) D_2	(2) D_3	(3) D_4
GPRT	$-6.15e\text{-}05^{***}$	$-8.58e\text{-}06^{***}$	$2.36e\text{-}06^{***}$
	(1.87e-05)	(3.10e-06)	(9.04e-07)
AR(1)	0.319^{***}	0.809^{***}	0.872^{***}
	(0.0396)	(0.0238)	(0.0355)
Constant	0.00629^{***}	0.000905^{**}	-0.000221^{*}
	(0.00188)	(0.000393)	(0.0001022)
R-squared	0.001	0.654	0.818

Newey-West (1984) standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: D_j stands for the details of the wavelet decomposition estimated using the Fejér-Korovkin family for NASDAQ returns. Same analysis as in equation *(8)* with a different wavelet family. Coefficients remain significant and similar in magnitude.

Table 18: Regression Results for CRSP index

Details	(1) D_2	(2) D_3	(3) D_4
GPRT	$-5.48e\text{-}05^{**}$	$-1.19e\text{-}05^{*}$	$3.24e\text{-}06^{***}$
	(2.60e-05)	(6.38e-06)	(1.12e-06)
AR(1)	0.326^{***}	0.831^{***}	0.938^{***}
	(0.0353)	(0.0207)	(0.0255)
Constant	0.00560^{**}	0.00118	-0.000312^{*}
	(0.00274)	(0.000777)	(0.000183)
R-squared	0.072	0.689	0.889

Newey-West (1984) standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: Where D_j stands for the details of the wavelet decomposition of value-weighted index of CRSP, with low j resembling high-frequency/low-persistency trends and high j representing low-frequency/high-persistency ones. Coefficients remain significant and similar in magnitude if not larger in absolute terms.