MASTER THESIS | MASTER'S THESIS

Titel | Title

Impacts of Geopolitical Risk on Inflation Expectations in the Eurozone

–

Estimating Deep Conditional Transformation Models with Neural Networks

verfasst von | submitted by

Luca Phileas Vetter, B.A.

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Advanced International Studies (M.A.I.S.)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 992 940
Universitätslehrgang lt. Studienblatt | Degree programme as it appears on the student record sheet:	Internationale Studien | International Studies
Betreut von | Supervisor:	Professor Martin Feldkircher

diplomatische akademie wien
Vienna School of International Studies
École des Hautes Études Internationales de Vienne

Abstract

This thesis aims to investigate how geopolitical risk in ECB speeches impacts professional inflation expectations in the Euro area. To that end, deep conditional transformation models based on neural networks are applied, which provide a flexible and probabilistically coherent framework for estimating the entire conditional distribution of inflation expectations. The empirical analysis incorporates macroeconomic uncertainty indices alongside textual representations of ECB speeches, including sentiment scores, geopolitically filtered FinBERT classifications, and high-dimensional transformer-based text embeddings. DCTMs show latent semantic components in ECB language are associated with meaningful shifts in the predicted inflation distribution. The findings underscore the methodological advantages of distributional forecasting, and the use of deep models based on neural networks for macroeconomic applications of textual data.

Abstrakt

Diese Masterarbeit untersucht, wie geopolitische Risiken in Reden der Europäischen Zentralbank (EZB) die Inflationserwartungen professioneller Prognostiker im Euroraum beeinflussen. Zu diesem Zweck werden sogenannte Deep Conditional Transformation Models (DCTMs) auf Basis neuronaler Netzwerke angewendet, die einen flexiblen und probabilistisch kohärenten Rahmen zur Schätzung der gesamten bedingten Verteilung von Inflationserwartungen bieten. Die empirische Analyse kombiniert makroökonomische Unsicherheitsindizes mit textuellen Repräsentationen von EZB-Reden, darunter Sentiment-Scores, geopolitisch gefilterte FinBERT-Klassifikationen sowie hochdimensionale, auf Transformer-Modellen basierende Text-Embeddings. Die DCTMs zeigen, dass latente semantische Komponenten in der Sprache der EZB mit signifikanten Verschiebungen in der prognostizierten Inflationsverteilung einhergehen. Die Ergebnisse unterstreichen die methodischen Vorteile einer verteilungssensitiven Prognose sowie den Nutzen tiefer Modelle auf Basis neuronaler Netzwerke für makroökonomische Anwendungen textbasierter Daten.

Impacts of Geopolitical Risk on Inflation Expectations in the Eurozone

Estimating Deep Conditional Transformation Models with Neural Networks

Luca Phileas Vetter[1]

[1]Diplomatic Academy Vienna, Austria

June 10, 2025

Abstract

This thesis aims to investigate how geopolitical risk in ECB speeches impacts professional inflation expectations in the Euro area. To that end, deep conditional transformation models based on neural networks are applied, which provide a flexible and probabilistically coherent framework for estimating the entire conditional distribution of inflation expectations. The empirical analysis incorporates macroeconomic uncertainty indices alongside textual representations of ECB speeches, including sentiment scores, geopolitically filtered FinBERT classifications, and high-dimensional transformer-based text embeddings. DCTMs show latent semantic components in ECB language are associated with meaningful shifts in the predicted inflation distribution. The findings underscore the methodological advantages of distributional forecasting and the use of deep models based on neural networks for macroeconomic applications of textual data.

JEL Classification: E31, E52, E58, C55, C45, C53

Contents

Word Count : 21.796 (excl. References)

List of Figures

List of Tables

* All figures and tables presented in the text and appendix without citation labels represent the author's own calculations and visualizations.

1 INTRODUCTION

Policymakers and central bankers alike have in the past shown a certain *insouciance* to geopolitical risks and their impacts on macroeconomic policy and the economy overall. Since the full-scale Russian invasion of Ukraine in 2022 and the precipitous rise in headline inflation, the awareness regarding the role that geopolitics plays in inflation is much more acute. And so, on the 10th of February 2025, the European Parliament debated, together with the President of the European Central Bank Christine Lagarde, the impact of global conflicts on price stability. In the Parliaments recommendations to the ECB, it "invites the ECB to draft a Geopolitics plan 2025-2030 in order to better understand the implications of war and conflict on price stability and treat all potential sources of external shocks equally" (European Parliament 2025). The Rapporteur from the European Parliament Anouk van Brog stresses that "it is time for the ECB to anticipate conflicts and their economic impact. By systematically incorporating geopolitical risks into its economic models, the ECB should be able to respond better to future inflation shocks and protect our purchasing power. The ECB must not only respond to crises but actively anticipate them" (ibid.). Despite the recent proliferation of geopolitical crises and their subsequent effects on the economy, practitioners and academics alike have often solely focused on the impact of geopolitical risk on inflation, and nary have considered the effects on inflation expectations. Hitherto, the econometric models have failed to incorporate, what orthodox economic theory would suggest is, *inter alia* responsible for price setting, through channels like wage-bargaining and price setting mechanisms.

Geopolitical fragmentation and tensions have recently overtaken inflation as primary risk to economic growth among central banks and sovereign wealth funds, according to a survey by the World Economic Forum (2024). The Russo-Ukrainian War, the conflict in Gaza and the associated supply chain disruptions in the Middle East, as well as geopolitical fragmentation in an increasingly multipolar world, have significant effects on the work of central banks. Whereas monetary policy has previously primarily relied on *prima facie* determinants of economic performance, like inflation and unemployment, as basis for their forecasts, CBs increasingly also look at global geopolitical developments for their forecasting.

The growing relevance of geopolitical tensions for central banks becomes apparent when looking at their communication, that increasingly also mentions these non-monetary news in their speeches.

By employing large language models and natural language processing, this thesis aims to capture the relationship between heightened central bank sensitivity to geopolitical risk and inflation expectations. To that end, several methodological approaches from the literature have been considered. A discussion of these approaches, the key issues and comparative advantages of DCTMs highlight their intrinsic benefit. Given, that this is the first ever application of DCTMs in central bank communication research overall, this thesis aims to serve as an impetus for future utilization of the methodology.

In general, as this thesis aims to bridge the interdisciplinary gap at the nexus between geopolitics and macroeconomic policymaking, the literature considered herein is gathered holistically and without prejudice to the respective academic field or foci, given the multidimensional threats that shape geopolitical risk, from commodity price volatility, wars, elections and economic policy uncertainty, et cetera. A special report in the ECBs Financial Stability Review by Dieckelmann et al. (2024) is one of few investigations that specifically looks at the relationship between CBs and the GPR index, although employing econometric models measuring "impulse responses to a 1 standard deviation global geopolitical risk shock based on a Bayesian vector autoregression model[,] identified using a Cholesky decomposition with geopolitical risk ordered first" (ibid.). This paper illustrates the short- and medium-term impacts of geopolitical risk on the harmonized index of consumer prices (HICP), which is the primary indicator measuring inflation in Europe. Based on this, and the work by Caldara, Conlisk, et al. (2023) it becomes apparent that "geopolitical risks are also associated with higher inflation uncertainty and the risk of significant inflation increases " (ibid.). These findings have instigated the research presented herein, as the foundation for an investigation that has shown early on, there remains a significant need to scrutinize the effect of central bank communication concerning geopolitical risk on inflation expectations.

More recently however, papers have found a more ambivalent relationship between geopolitical risk and inflation, highlighting the specific circumstances around the amplitude and intensity of any given geopolitical risk on inflation. Bringone et al. "find that the primary transmission channel of such shocks is associated with heightened uncertainty, which significantly escalates only with substantially large GPR shocks (i.e., above 4 standard deviations). This increase in uncertainty prompts precautionary saving behaviors, exerting a strong impact on consumption and reducing activity. The response of inflation is more subdued, reflecting both diminished demand and heightened uncertainty, which influence prices in opposing directions" (Brignone, Gambetti, and Ricci 2024). The economists from the Bank of England "argue that geopolitical risk (GPR) can transmit via two separate and intrinsically different channels: (i) a deflationary macro channel, and (ii) an inflationary energy channel. [They] then use a Bayesian vector autoregression (BVAR) framework to evaluate these channels empirically. [The] estimates suggest that GPR shocks can place downward or upward pressure on advanced economy price levels depending on which of the two channels the shock propagates through." (ibid.)

In a nutshell, different geopolitical risks exert deflationary and inflationary pressures on a macroeconomic level and specific determinants contribute to the associated effect. However, these findings are based on Bayesian Vector Autoregressive Models, which do not take into account the sentiment of CBs to any given geopolitical shock. Therein lies the novelty of this paper, as we estimate the immediate effect of CB response to geopolitical risks on inflation expectations, which precipitate shifts therein directly.

As such, this timely debate exemplifies the growing awareness concerning the impacts of geopolitical risk on the mandate of the European Central Bank, but sofar the dynamic interplay between CBs and geopolitical risk are severely under-researched. This thesis aims to reflect on how central bank communication regarding geopolitical risk can, among other determinants, impact inflation expectations. To that end, the newly released `deeptrafo` package is employed, whereby deep conditional transformation models can be estimated. This methodology is particularly apt to investigate the aforementioned relationship because we can reconcile natural language processing and large language models from ECB speeches with quantitative data, including the geopolitical risk index, economic policy uncertainty and VIX, to estimate the entire conditional distribution function of inflation expectations in the Eurozone.

DCTMs are uniquely qualified for this assessment, because they can capture nonlinear effects and distributional shifts. Instead of modeling the mean like in linear regressions, the DCTM incorporates the entire probability distribution function of inflation expectations. There are a plethora of advantages to this novel approach, over more traditional methods. Linear models primary limitations to applications herein, in addition to the aforementioned ones, include their performance in capturing heteroskedasicity and changing tail risks and issues regarding higher order relationships between markets and geopolitical risks. Standard deep learning models are associated with other drawbacks. Chiefly among which are that they do not estimate probability distributions, while only offering point predictions, as they lack interpretability in quantifying risk shifts. Moreover, research suggests that linear models underestimate the effect of geopolitical risks as certain factors only *kick in* with sufficiently large geopolitical shocks.

In section 2, the current research environment will be analyzed, both in the realm of macroeconomic implications of geopolitical risk and the inclusion thereof in current inflation expectation measurements. Specifically, subsection 2.1 will investigate the empirical relationship between geopolitical risk and inflation, as a discussion on the impacts on inflation expectations are superfluous if there were to be no actual effect on the real economy.

Section 2.4 considers the role of CBs at the nexus between monetary policy and geopolitical risk. Namely, the paper will analyze where geopolitical risk enters the transmission mechanism and what the ECB currently does to incorporate the associated risks in their econometric models and policies. Furthermore, this thesis will assess the current system of inflation forecasting and their interplay with geopolitical risk. This analysis will include the survey of professional forecasts, market-based inflation expectations through inflation-linked swaps (ILS) and consumer expectation surveys, and their respective propensity to respond to changes in geopolitical risk.

Section 3 focuses on the methodology of this thesis. This includes a discussion of the measurement of geopolitical risk, the natural language model proposed herein, how the model was specified and why the novel approach based on deep conditional transformation models is best equipped to estimate the impact of geopolitical risk on inflation expectations. Section 5 will present the results of the findings. To test the efficacy of the statistical model, robustness tests are performed in section 6, relying primarily on log-likelihood models and calibration plots. Furthermore, the distributional shifts will be exemplified, both with predicted and empirical distributions, and further complemented with linear quantile regressions and Kernel-density plots. Thereupon, section 7 discusses the limitations of the statistical analysis and methodology. Lastly, section 8 will conclude the thesis and summarize the main findings and arguments presented herein.

2 LITERATURE REVIEW

This section will focus on the literature hitherto. The chapter is further subdivided between an analysis of geopolitical risk and inflation with a focus on supply chain disruptions, consumer expectations and conflict-induced energy market implications. The second part of the chapter focuses on central banks and geopolitical risk. Here, the thesis will primarily concentrate on the perception and response of central banks to geopolitical fragmentation, particularly how geopolitical risk enters the transmission mechanism and how central banks address the risks of upward dis-anchoring of short-term inflation expectations in response to geopolitical risk shocks.

This literature review is critical for the *a prior* methodological decisions, that have guided the design of the operationalization of the DCTMs. It will answer the question why this thesis only focuses on short-term inflation expectations in the Euro Area, why geopolitical risk matters for inflation expectations in the first place, and particularly how.

As previously discussed, there remain both inflationary and deflationary pressures associated with geopolitical risk, dependent on their activation provided that the geopolitical risk shock is large enough. This also implies that the conclusions drawn from the papers herein largely depend on the geopolitical risks and shocks included in the model, which henceforth may also include those that, upon preliminary inspection, contradict each other.

2.1 Geopolitical Risk and Inflation

This section delves into the direct effects of geopolitical risk on realized inflation, which constitutes a critical foundation for the subsequent analysis of its propagation into inflation expectations. Establishing a robust empirical and theoretical link between geopolitical tensions and inflation is essential, as without such a link, exploring the downstream effects on expectations among firms, households, and professionals would be analytically unsubstantiated. Fortunately, a growing body of literature confirms that geopolitical risk does, in fact, exert statistically and economically significant pressure on inflation dynamics.

Seminal in this field is the work by Caldara, Conlisk, et al. (2023) who, along with their earlier methodological contribution (Caldara and Iacoviello 2022), developed the Geopolitical Risk Index

(GPR) used in this thesis. Their empirical analysis finds consistent evidence that rising geopolitical tensions are inflationary in nature, particularly through their impact on energy prices, risk premia, and global uncertainty. Expanding on this, Brignone, Gambetti, and Ricci (2024) introduce a non-linear framework in which the inflationary impact of geopolitical shocks depends crucially on the magnitude of the initial shock. They show that geopolitical uncertainty only becomes macroeconomically relevant when a critical threshold is surpassed, beyond which indicators such as the VIX surge, amplifying financial market volatility and feeding into inflation through multiple channels. The model they develop identifies dual propagation modes: a baseline mode for moderate shocks, and an intensified mode that activates under large-scale geopolitical disruptions, a stylized version of which is visualized in Figure 1.

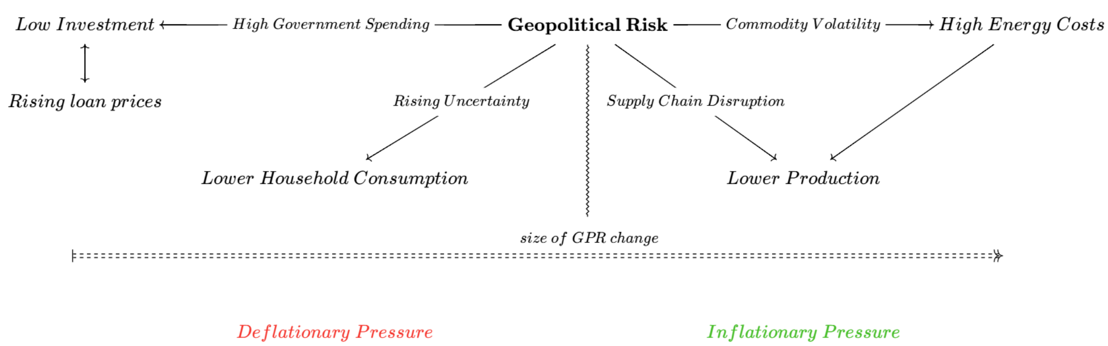


Figure 1: Inflation Impacts of Geopolitical Risk

The figure offers a stylized representation of the various transmission channels through which geopolitical risk affects inflation dynamics. Central to the figure is the conceptualization of geopolitical risk (GPR) as a shock variable with non-linear effects on macroeconomic outcomes, particularly price levels. The figure is arranged along a horizontal axis indicating the magnitude of GPR shocks, with a clear bifurcation in outcomes based on whether the shock remains moderate or reaches a critical threshold. On the left-hand side, smaller or less persistent geopolitical disruptions tend to activate demand-side mechanisms that exert deflationary pressure. These include heightened uncertainty, rising loan prices, and a resulting decline in both household consumption and investment. In such scenarios, geopolitical risk primarily suppresses economic activity, echoing findings in the literature by Klein, Skeppås, and Söderström (2025), Hundal (2023), and Kyriazis, Economou, and Stergiou (2023), who document how elevated uncertainty leads to precautionary behavior by firms and consumers, weakening aggregate demand and placing downward pressure on inflation.

However, as the magnitude of the geopolitical shock increases, additional and qualitatively distinct channels begin to dominate, represented on the right-hand side of the figure. In particular, large-scale GPR shocks trigger supply-side constraints such as disruptions in global supply chains, increased commodity price volatility, and spikes in energy costs. These dynamics reduce production capacity and increase the marginal cost of goods, thereby generating persistent upward pressure on prices. The inflationary consequences of such mechanisms are well-documented in the empirical work of Caldara, Conlisk, et al. (2023), who emphasize the inflationary role of energy markets and geopolitical tensions, and by Brignone, Gambetti, and Ricci (2024), whose threshold-based model of GPR confirms that macroeconomic effects intensify non-linearly beyond certain levels of geopolitical stress. The figure visually captures this insight by distinguishing between a baseline regime of moderate shocks and an augmented regime where additional inflationary forces are activated once geopolitical risks surpass a critical magnitude.

Furthermore, this framework aligns with the research of Marangoz (2025) and Bouri et al. (2023), who explore how geopolitical risk transmits inflationary pressures across regions, particularly in energy-importing economies like those in the Eurozone. Their findings highlight the heterogeneity in regional inflation responses depending on the exposure to global supply chains and commodity dependencies. The figure thus supports the broader thesis of this study: that geopolitical risk is not a monolithic phenomenon, but one whose inflationary effects depend crucially on scale, persistence, and the nature of the underlying channels.

Importantly, this conceptual model provides a foundation for analyzing how different types of GPR shocks may influence inflation expectations among households, firms, and financial markets. In doing so, it also sets the stage for evaluating monetary policy responses, especially in light of recent findings by Nuño, Renner, and Scheidegger (2024), who demonstrate that traditional Taylor rule frameworks may be insufficient under persistent supply-side inflation. The recognition that supply chain disruptions and energy shocks can lead to endogenous shifts in the natural interest rate adds further urgency to understanding the structural features of geopolitical inflation, as visualized in the figure.

The uncertainty channel, which is a principal component in the aforementioned vicissitudes of inflationary pressures associated with geopolitical shocks, is further examined by Hundal (2023), Klein, Skeppås, and Söderström (2025), and Kyriazis, Economou, and Stergiou (2023), all of whom emphasize the role of elevated geopolitical uncertainty in constraining investment and consumption, disrupting supply chains, and triggering precautionary price setting behavior among firms. Their findings suggest that uncertainty itself, not just observable disruptions, can exert upward pressure on inflation through expectations and supply-side rigidities. Meanwhile, Marangoz (2025) adopts a time-varying parameter VAR (TVP-VAR) framework to assess how geopolitical risk transmits through inflation spillovers within the Euro area. His findings align with those of Bouri et al. (2023), who similarly focus on regional heterogeneity in response to global risk shocks and underscore the heightened vulnerability of the Eurozone to inflation imported through energy markets and cross-border supply chains.

Taken together, these contributions suggest that the relationship between geopolitical risk and inflation is complex, non-linear, and context-sensitive. They also affirm the analytical strategy pursued in this thesis: that understanding the inflationary consequences of geopolitical shocks is a prerequisite for comprehending how such shocks shape inflation expectations and, by extension, how they should inform optimal monetary policy responses.

2.2 Supply Chain Effects of Geopolitical Fragmentation

One of the main channels ostensibly responsible for geopolitical risk-induced inflation are supply chain disruptions. Transitory supply chain disruptions, as were seen during the Covid-19 pandemic when factories in China and other producers upstream of the value chain have closed, can drive inflation significantly, but don't *per se* merit tighter monetary policy. This is because the transmission mechanism of interest rates is long and varied, and when supply shocks are only temporary, tighter monetary policy could easily overshoot the target inflation rate. But, this relationship is too often considered axiomatic, without reflection for optimal policy responses when supply shocks are heterogeneous or persistent, which can both be features of geopolitical supply chain disruptions. Hence, this paper takes a closer look at the vicissitudes of supply chain disruptions in the context of geopolitical risk, because it has a significant effect on the inflation expectation by households, firms and professionals alike, as well as real inflation. To that end, this thesis *inter alia* draws from the work by Bednarski et al. (2025) who perform a literature review of geopolitical disruptions in global supply chains. Their identification of themes, gaps and causal interpretations guide the analysis herein. This includes the research by Acemoglu and Tahbaz-Salehi (2024), who look a the macroeconomic effects of supply chain disruptions. They identify fragilities in firm-level networks with significant macroeconomic implications. Asadollah et al. (2024) and Bjørnland (2022) look at the effect of supply chain disruptions on inflation specifically, whose findings are also pertinent to the research conducted herein. Lastly, Nuño, Renner, and Scheidegger (2024) build a New Keynesian model to study optimal monetary policy response with persistent supply shocks. They show that the taylor rule fails to stabilize long-term inflation due to endogenous shifts in the natural interest rates. They also analyze optimal policy responses under discretion and commitment, including through the incorporation of the zero lower bound and find evidence to support preemptive easing.

Beyond the immediate inflationary pressures, geopolitical fragmentation introduces structural changes to global supply chains that can create long-lasting frictions in trade and production. Rising tensions between major economies often lead to strategic decoupling, reshoring, or friend-shoring, all of which reduce the efficiency gains historically achieved through globalized supply chains (Ademuyiwa and Siklos 2019). This fragmentation undermines comparative advantages, increases input costs, and reduces the elasticity of supply, particularly in sectors dependent on critical raw materials or highly specialized intermediate goods. Over time, these disruptions can lead to a new inflationary regime characterized by higher volatility and persistence, especially if firms internalize the risks of geopolitical instability through precautionary hoarding or redundant sourcing strategies. Such shifts can impair productivity growth and raise the natural rate of interest, complicating the conduct of monetary policy. Therefore, understanding the second-round effects of geopolitical fragmentation is essential not only for forecasting inflation but also for reassessing the monetary transmission mechanism under global supply chain reconfiguration.

2.3 Conflict-induced Energy Market Implications

Geopolitical crises have often been associated with global commodity and energy price implications. The causal relationship between the ubiquity of geopolitical tensions and the clustering thereof in regions heavily endowed with resources exceeds the scope of this paper. Because geopolitical risk primarily propagates in headline inflation, and as S. Davis (2015) has found also to a lesser extent core inflation, and short-term inflation expectations, energy prices are extremely responsive to geopolitical shocks and therefore bear analysis herein. As shown previously, geopolitical risk shocks exacerbate both deflationary and inflationary pressures, of which the outcome on the real economy is not merely determined by the activation of non-linear effects as evidenced before, but also by the energy market implications of the specific GPR shock. For example, Pinchett (2024, p. 23) finds in his empirical analysis that "a rise in geopolitical risk is always contractionary for economic activity, irrespective of its nature. However, a rise in the GPR index is deflationary when associated with geopolitical macro shocks and inflationary when associated with geopolitical energy shocks. The magnitudes of these effects are estimated to be economically meaningful from a policymaking perspective".

Consequently, analyzing the specific type and region from which the geopolitical shocks are emanating is particularly pertinent for CB responses, as there is no analogous implications that can simply be inferred from the values of GPR change. So far, the research suggests the fulcrum for this qualitative assessment of GPR shocks seem to rest in the energy market implications, at least for the Euro area, which mostly consists of net energy importers from what are now considered fragile supply chains. However, to further complicate the achievement of the price stability mandate by the ECB, there are "considerable quantitative differences in the dynamic effects of oil supply shocks on energy and core inflation as well as inflation expectations and consumer confidence across Euro area countries". Baumeister 2023, p. 18. The author relates the cross-country heterogeneity in the relative strength of transmission channels to structural differences based on energy mix, industry composition, and competitive environment.

More importantly for this paper, are however the effects on the SPF itself and the forecasts they produce. Bjørnland (2022) finds that, on average, a 10% increase in oil prices due to geopolitical tensions or supply constraints will reduce GDP in the Euro area by 0.5%, all other things held equal. This would *ipso facto* mean that a 50% increase in oil prices could reduce the GDP by 2.5 percent, or more when oil price volatility is also high. Consequently, oil supply shocks are not just associated with higher inflation through cost channels and second round effect of wage-bargaining mechanisms, but they also depress economic activity which generates a trade-off for central banks between inflation and output stabilization (Baumeister 2023, p. 14). In a special survey among

SPF participants, researchers have found that 3 out of 4 respondents saw changes in accuracy of the point forecasts, as some indicators have been less reliable in signalling inflation and their relationship to other macroeconomic variables had changed stemming from the war in Ukraine. Interestingly, non-Ukraine geopolitical factors still affected models and long-term forecasts from 20-30% of respondents (Allayioti et al. 2024, pp. 11–12). This follows the logic by Hasselgren et al. (2025) who use a GARCH-MIDAS model to investigate the predictive properties of disagreement about the oil price from the SPF and how they translate to larger oil return volatility, which links to a broader narrative, that a large part of inflation expectation inaccuracies from the SPF stems from false assumptions in the oil price developments.

Now, it is *prima facie* apparent, that higher energy costs reduce overall demand and as oil prices directly feed through to prices due to higher production costs, this will mechanically also impact inflation (Bjørnland 2022). But the optimal response strategy for central banks or inflation expectations is less self-evident, than this cost channel. Of course, the general translation mechanism to inflation expectations still applies, as production cost channel increases indirectly pass-through via price setting and wage-bargaining mechanisms (ibid.), but it is critical to bear in mind that geopolitical risk shocks are not always associated with oil price increases and inflationary pressures. In fact, Anttonen and Lehmus (2024, pp. 9–10) find that the outbreak of the Israel-Hamas War exercises more deflationary pressure in the Euro area and coincides with a global oil price reduction, due to US production increases and an OPEC decision on output ceilings. This illustrates that oil prices can be volatile and are the sum of a plethora of factors outside of what can be individually attributable to isolated geopolitical shocks.

The crucial determinant upon which the impact of energy market fluctuations on inflation and inflation expectation hinges is persistence. Bjørnland (2022) argues "the degree of sensitivity depends on the underlying source of oil market shocks [and] unexpected large global economic activity shocks, such as the early millennium oil price surge of 2003-08, elicits a persistent response in both expected and actual inflation [while transitory] shocks to oil supply, consumption demand, or inventory demand, both expected and actual inflation initially increase but then gradually revert back to zero". As such, qualitative assessments of geopolitical shocks first need to consider its magnitude, the role of the region or polity for energy supply and the persistence of the shock to assess if non-nonlinearities in stock uncertainty activate and the volatility in commodity prices translates to prolonged energy inflation through the cost channel, the consequences of which are still contentious. Bjørnland (2022) argues, that because stabilizing inflation is difficult during periods of high oil price volatility and there remains a substantial share of variance in inflation explained by the oil price shocks, even when central banks are hawkishly responsive, that "central banks need to be swift in their response to curb inflation, so as to prevent inflation expectations from increasing ".

The thesis relies on a wide breadth of literature that has investigated the relationship between oil rents, energy markets and geopolitical shocks. Yang et al. (2023) use TVP-SV-VAR models to assess the impact of geopolitical shock on oil prices, which is further corroborated by Foglia, Palomba, and Tedeschi (2023) who finds that impulse responses to geopolitical risk shocks are heterogeneous, country-specific and dependent on proximity to the shock.

They also show that GPR shocks affect the energy sector more than food and metals, which is why the section focuses on commodities like oil and gas more than precious metals and foodstuff. Gregorio (2012) and Drechsel et al. (2024) look at monetary policy responses to geopolitical risk shocks, given the multidimensional contagion of commodity price spikes in the transmission mechanism. Corsello, Gomellini, and Pellegrino (2023) provide historical context with the monetary policy response to energy price shocks in the aftermath of the Yom Kippur War and the Iranian Revolution in the 1970s. They highlight the importance of an independent institutional framework for anchored inflation expectations in times of energy price shocks, which remains critical hitherto, because the risk of upward dis-anchoring of inflation expectations, may also affect the way the SPF assesses inflation expectations, given the experience of costly re-anchoring during that time.

Ademuyiwa and Siklos (2019) investigate the impact of GPR on the real economy and international trade, finding that their quantitative estimations underscore the role of commodity and trade shocks as responses to heightened GPR in the G20. Pinchett (2024) analyzes the role of energy markets for inflation, which is critical to understand to make assumptions of the role of energy markets on real and expected inflation by consumers. In the same vain, Bjørnland (2022) investigates the recent energy price increases in Europe and what macroeconomic consequences they have. Qin and Zhang (2024) identify the transmission mechanism by which geopolitical risk precipitates in energy price increases. A similar study using quantile regressions is conducted by Mo, Nie, and Zhao (2024).

Furthermore, Alsagr and van Hemmen (2020) analyze how oil rents in times of geopolitical risk affect the banking sector, which show in a fixed effect model, that geopolitical risk weakens banking sector profitability and weakens the moderation role of oil rents. Baumeister (2023) looks at the pass-through of oil supply shocks in the Euro area and focuses on both structural changes in the supply side of global oil markets and the demand-side policy tools to stabilize markets and counter inflationary pressures in the aftermath of the Russian invasion of Ukraine.

2.4 Central Banks and Geopolitical Risk

This section will focus on central bank responses to geopolitical risk, primarily by means of monetary policy and communication. Given, that this thesis solely focuses on the Euro area, this chapter will concentrate on the ECB specifically. Fundamentally, in the wake of geopolitical shocks, central banks face a dilemma: whether to adopt a hawkish stance aimed at preempting inflationary pressures or to "see through" the shock, treating it as transitory and beyond the control of monetary policy. Advocates of a hawkish response argue that geopolitical disruptions, particularly those affecting energy and commodity markets, can lead to persistent second-round effects on inflation expectations, wages, and price-setting behavior. If unaddressed, these dynamics may dis-anchor expectations and threaten credibility, especially in jurisdictions where inflation targets are symmetric and forward-looking. In contrast, proponents of a more accommodative approach contend that geopolitical shocks primarily constitute negative supply shocks, which central banks are ill-equipped to counter without exacerbating output losses. From this perspective, tightening policy in response to externally driven, temporary price increases risks undermining growth and amplifying volatility without materially influencing the inflation trajectory. The appropriate stance ultimately depends on the perceived persistence of the shock, the transmission channels involved, and the institutional mandate of the central bank, but central bankers and the academic literature still seem somewhat divided on the optimal policy response to geopolitical shocks, which are *a priori* difficult to assess for the purposes of the price stability mandate.

Apart from shaping monetary policy decisions, geopolitical risk has increasingly permeated central bank communication, becoming a recurrent theme in speeches, reports, and press conferences. As central banks seek to manage expectations and maintain credibility in an environment characterized by elevated uncertainty, their language itself becomes a tool of monetary transmission. Accordingly, the present thesis incorporates natural language processing (NLP) and sentiment analysis techniques to evaluate the salience and tone of geopolitical risk references in European Central Bank (ECB) communications. This methodological approach is directly motivated by the growing academic recognition of the non-monetary informational content embedded in central bank discourse.

Cieslak and Schrimpf (2019) make a compelling contribution to this area by demonstrating that non-monetary news, such as geopolitical developments, financial instability, or fiscal policy shocks, embedded in central bank communication can exert significant effects on financial markets. Using high-frequency data and a decomposition approach, they find that asset prices respond not only to forward guidance and monetary policy surprises but also to qualitative statements regarding broader macroeconomic conditions. This implies that market participants parse central bank language carefully, attributing weight even to indirect or non-policy-specific commentary. In the context of geopolitical risk, this finding reinforces the analytical value of examining how frequently and in what tone geopolitical topics are addressed in central bank speeches, as done in this thesis. If such language influences market expectations and risk premia, then it forms a de facto component of the monetary transmission mechanism, even if no immediate policy action accompanies it.

Complementing this perspective is the earlier work by Hayo, Kutan, and Neuenkirch (2012), who examine the impact of U.S. Federal Reserve communication on commodity price volatility, thereby extending the relevance of central bank language beyond interest rate expectations. Their findings suggest that central bank statements can serve as informational shocks, especially in markets sensitive to uncertainty, such as those for energy or metals. Crucially, the study finds that specific forms of communication, particularly those that provide clarity or acknowledge geopolitical developments, can either amplify or mitigate market volatility depending on the perceived credibility and informativeness of the message. In the European context, this underscores the importance of how the ECB communicates geopolitical concerns, especially during crises involving trade wars, military conflict, or energy insecurity.

Together, these studies support the argument that central bank communication on geopolitical matters is not merely rhetorical or reactive, but has tangible effects on financial variables and inflation expectations. Consequently, analyzing this communication using NLP and sentiment analysis, as undertaken in this thesis, is not only methodologically justified but economically meaningful.

The frequency, context, and sentiment of references to geopolitical risk in ECB speeches provide an additional lens through which to assess the central bank's perception of macroeconomic threats, and its strategy in managing public and market sentiment in response.

Furthermore, the ECB has produced a report regarding the EU's Open Strategic Autonomy from a Central Banking Perspective by Ioannou et al. (2023), which also sheds light on the growing challenges CBs face in an era of elevated geopolitical risk. Other relevant papers include a paper by Plakandaras and Papadimitriou (2019) on the effect of geopolitical uncertainty in forecasting financial markets. Jarociński and Karadi (2020) discuss the role of information shocks and monetary policy surprises, which may trigger geopolitical shocks themselves. Bouri et al. (2023) discuss geopolitical risk and inflation spillovers across European and North American economies, and Morgan (2019) performs a quantitative study on geopolitical awareness and monetary policy.

2.5 Geopolitical Risk in the Transmission Mechanism

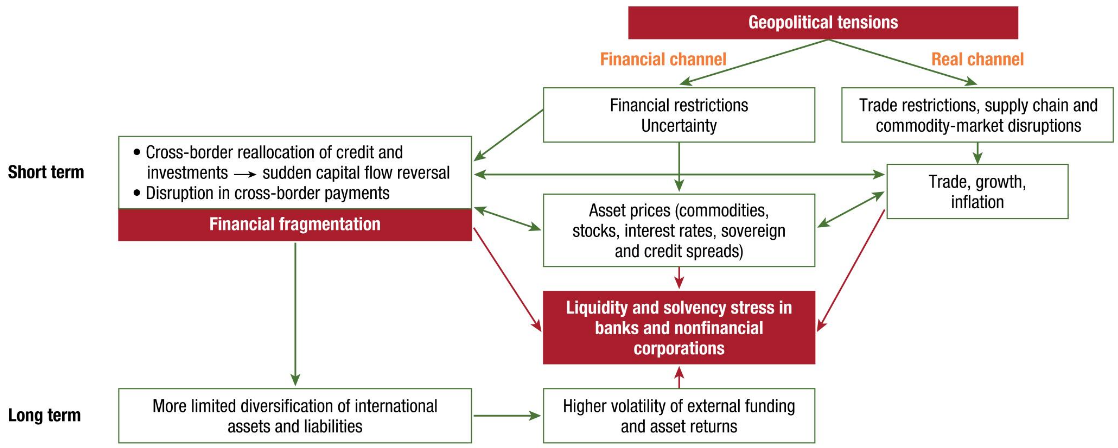
This chapter focuses on the impacts of geopolitical risk in the transmission mechanism. To that end, the channels through which geopolitical risk propagates outside of inflation needs to be understood. Figure 2 by the International Monetary Fund (2023), captures the transmission both on the financial and the real economy side. Given, that the literature on geopolitical risk transmission in monetary policy in Europe is scant, this thesis will broaden the horizon and draw from a broad set of literature and infer from the conclusions the overall architecture of the transmission channel pathways for geopolitical risk. To that end, this paper will build on the work by Güntner and Henßler (2021), who use high-frequency external instruments to separate the signaling channel of monetary policy from exogenous interest rate changes to show that supposedly contractionary monetary policy shocks effects on U.S. banks' balance sheets, asset markets, and economic activity hinge on the level of geopolitical risk at the time of the FOMC announcement. Moreover, this paper takes the work by Qin and Zhang (2024) into account, that looks at the transmission of geopolitical risk on commodities, whose effect on the transmission mechanism are more broadly known and understood.

Geopolitical risk has increasingly been recognized as a critical modifier of the monetary policy transmission mechanism, influencing both the real economy and financial markets in ways that challenge standard assumptions. As Figure 2 illustrates, geopolitical risk propagates through multiple channels, ranging from heightened uncertainty and risk premia in financial markets to disruptions in global trade, supply chains, and commodity flows. These channels, while not always inflationary in the short term, alter the transmission of interest rate changes by shifting the behavior of households, firms, and financial intermediaries. In high-risk environments, precautionary savings tend to rise, investment becomes more sensitive to policy uncertainty, and capital reallocation across borders may accelerate. Consequently, the same monetary impulse, whether expansionary or contractionary, may yield significantly different macroeconomic outcomes depending on the prevailing geopolitical context.

This thesis draws on evidence from both European and non-European contexts to map out these conditional dynamics. The work of Güntner and Henßler (2021) provides compelling evidence that the effectiveness of monetary policy announcements in the United States hinges on the level of geopolitical risk. Specifically, they demonstrate that contractionary monetary policy shocks can, paradoxically, have expansionary effects on asset prices and bank balance sheets when geopolitical uncertainty is elevated, suggesting a dampening or reversal of the conventional transmission mechanism. Similarly, the findings by Qin and Zhang (2024) highlight the amplification of geopolitical shocks through commodity markets, particularly energy, which serve as a central node in the inflation channel. These insights are essential for understanding the European context, where reliance on energy imports and exposure to global political instability introduce pronounced vulnerabilities. By incorporating these conditional transmission effects into the modeling of inflation expectations, this thesis aims to offer a more nuanced view of how geopolitical risk reconfigures the channels through which monetary policy operates, calling into question the universality of standard policy prescriptions during turbulent geopolitical periods.

The IMF's schematic in Figure 2 underscores the dual-channel nature of geopolitical risk transmission: one operating through the financial channel, and the other through the real channel. This distinction is crucial in the context of this thesis, which explores how geopolitical signals embedded in ECB communication may shape inflation expectations via these very mechanisms.

On the financial side, geopolitical tensions increase uncertainty and lead to financial restrictions, such as capital controls or sanctions. These frictions trigger financial fragmentation by disrupting cross-border capital flows and reallocation of investment. In the short run, this can cause sudden re-



Source: IMF staff.
Note: The figure shows the two key channels of transmission, financial and real, through which geopolitical tensions could contribute to financial fragmentation and exacerbate macro-financial stability risks. In addition to these channels, macro-financial stability could also be affected if geopolitical tensions increase cybersecurity risks, compliance, legal and reputational risks for entities, risks associated with money laundering and financing of terrorism, or climate-related risks because of lack of international coordination to mitigate climate change.

Figure 2: (International Monetary Fund 2023, p. 87)

versals in capital movement and dislocations in payment systems, while in the long run, it constrains the diversification of international portfolios. Financial fragmentation limits the efficacy of interest rate policy by reducing arbitrage, impairing transmission across jurisdictions, and increasing the volatility of external funding conditions. ECB communication that acknowledges or downplays such dynamics may significantly affect how professional forecasters interpret monetary policy credibility and inflation risks.

Simultaneously, geopolitical shocks influence asset prices, notably commodities, interest rates, and sovereign credit spreads, which are directly relevant for forward-looking inflation expectations. As shown in this thesis, the DCTM framework is capable of capturing such distributional effects when embedding vectors encode underlying geopolitical sentiment. The liquidity and solvency stress that may arise in banks and non-financial corporations further amplifies systemic risk, which in turn feeds into inflation forecasts through expectations of policy loosening or fiscal backstopping.

On the real side, geopolitical tensions lead to trade restrictions, supply chain disruptions, and commodity market volatility, all of which were observed during the energy crisis following the Russian invasion of Ukraine. These real-side frictions feed into the macroeconomy through reduced output, higher input costs, and ultimately inflationary pressures, which monetary policy must address. ECB speeches that reflect concerns about these frictions may be interpreted as either hawkish signals, if inflation control is prioritized or dovish signals if growth risks dominate, depending on the semantic structure and context of the language used.

What the IMF diagram makes clear, and what this thesis emphasizes, is that geopolitical risk is not an exogenous backdrop but an active force that reconfigures the architecture of the transmission mechanism itself. It alters both the pathways through which policy affects the economy and the expectations that economic agents form about future states of the world. In doing so, it complicates the central bank's signaling function and renders the task of managing inflation expectations more delicate. The deep learning approach embedded in the DCTMs allows for such structural breaks and nonlinearities to be captured, offering a more resilient framework for modeling expectations under conditions of elevated uncertainty and global fragmentation.

Ultimately, the integration of textual signals into the DCTM framework complements the IMF's conceptual model by offering an empirical lens through which to observe how central bank communication interacts with the geopolitical risk transmission structure. It affirms that geopolitical tensions do not merely add noise, they reshape the signal itself.

2.6 De-Anchoring Inflation Expectations

Orthodox economic theory suggests that inflation expectations can have effects on real inflation if market participants make behavioral decisions based on perceived and expected inflation (Afunts, Cato, and Schmidt 2024, p. 1) (Backer et al. 2023, p. 3). Households may adjust their demands in wage negotiations and firms set their prices in anticipation of short and longer-term inflation expectation. Consequently, it is critical that inflation expectations are well-anchored, and that the CB sets credible targets, to avoid spiraling behaviors in response to inflation surprises, that require tight monetary policy to reign in, which comes costly to central banks.

Re-anchoring inflation expectations has been necessary in the United States in the 1970s and 1980s, and the experience of fighting, what the Fed Chairman Burns called *inflation psychology*, has unequivocally shown how difficult and expensive correcting entrenched inflation expectations can become *vis-á-vis* GDP growth and unemployment (ibid., p. 3).

The ECB and scholars alike are therefore extremely vigilant, when the inflation expectation diverges from the, since 2021 symmetric, 2% target over the medium term (Neri et al. 2022, p. 8), to prevent the dis-anchoring from inflation targets. Research has shown that consumers value geopolitical risk in the context of their inflation expectations. This section will build on these findings to discuss how geopolitical risk propagates in the short-term inflation expectations, how likely pass-through to longer and long-term inflation expectations are and how other market participants estimate the likelihood of upward dis-anchoring in response to geopolitical shocks.

There are three main sources by which inflation expectations are gauged; namely market-based measurements like inflation-linked swaps (hereafter: ILS), professional forecasters and surveys among households and firms. Of course, all these measurements are associated with some drawbacks, which make capturing the actual inflation expectations difficult. For example, the ILS prices include risk premia, which make an isolated assessment of market-based inflation expectations difficult (Backer et al. 2023, p. 6). Surveys among consumers are notoriously skewed to the right, as market participants overestimate both perceived and expected inflation for psychological and sociodemographic reasons, which also relate to the volatility of headline inflation that consumers and firms are more attuned to, and that are also more likely to change in the aftermath of geopolitical shocks. Lastly, professional forecasters, at least the SPF by the ECB, doesn't include long-term inflation expectations over 5 years, and professional forecasters have shown to be susceptible to volatile commodity price changes affecting their model, which also coincides with geopolitical shocks.

The thesis will focus on the work by Afunts, Cato, and Schmidt (2024), who use the online household survey from the German Bundesbank (BOP-HH) as a natural experiment in an event study to compare the inflation expectations of the same survey round before and after the begin of the invasion. Moreover, the paper will also draw from the work by Coibion et al. (2025) who ask geopolitical risk specific items in several rounds of the Consumer Expectation Survey by the ECB.

This chapter will therefore focus on how these measurements of inflation expectations diverge in the risk assessment for shock dis-anchoring in response to geopolitical risks and how these propagate in the pass-through from short to longer-term inflation expectations. This chapter will show that while geopolitical risk is not just implicating inflation, it also has an effect on inflation expectations, as shown by Neri et al. (2022) and Backer et al. (2023). Furthermore, the event study conducted by Afunts, Cato, and Schmidt (2024), which assess responses from household surveys conducted by the German Bundesbank immediately before and after the start of the full-scale Russian invasion serve as a natural experiment for changes in the inflation expectations, which also holds true for professional forecasters, as analyzed by Dräger, Gründler, and Potrafke (2022).

In their study, treatment effects of the Russian invasion on inflation expectation are measured by a survey immediately before and after the begin of the war We conducted the survey between 22 February 2022 and 1 March 2022, allowing us to compute treatment effects of the Russian invasion of Ukraine on 24 February 2022 by comparing mean expectations of experts surveyed before and after the invasion. (ibid., p. 3). The survey focused on German economic experts, who update their inflation expectations more rationally than households and firms and significantly shape the policy space and public debate on inflation (ibid., p. 2). Of course, the survey estimations are not perfectly analogous to the BOP-HH, disected in section herein, or the SPF itself, given the variance in the survey design which diverge in inflation expectation intervals, as well as socio-demographic and geographic distribution of survey participants, but the overarching notion remains true, that professionals and consumers alike update their inflation expectations specifically because of geopolitical risk shocks, which make the research endeavor explored in this thesis all the more pertinent.

A comprehensive discussion of inflation expectations must also consider foundational theoretical challenges to their measurement and interpretation, notably Ricardian equivalence and the Lucas critique. Ricardian equivalence, originating with Barro (1974), posits that rational agents internalize the government's intertemporal budget constraint, thereby rendering fiscal policy neutral: if a government increases deficit-financed spending, forward-looking agents anticipate future taxes and adjust their consumption and savings accordingly. In the context of this thesis, which models forward-looking inflation expectations using both macroeconomic variables and central bank communication, Ricardian equivalence underscores the importance of how agents interpret the fiscal-monetary nexus, especially in periods of geopolitical stress. For example, if fiscal expansions in response to geopolitical shocks are perceived as transitory and expected to be offset by future tightening, inflation expectations may remain anchored, regardless of current spending.

The Lucas (1976) critique further complicates the analysis by emphasizing that expectations are not invariant to policy regimes: any econometric model that fails to account for the endogenous formation of expectations under changing policy frameworks will produce unstable and potentially misleading forecasts. This insight is particularly relevant to the methodological choice of using deep conditional transformation models in this thesis, as DCTMs can capture nonlinearities and potential regime shifts in the expectation formation process. Moreover, it supports the inclusion of ECB communication as a dynamic explanatory factor, since agents respond not only to realized macroeconomic outcomes but also to evolving perceptions of policy intent.

3 METHODOLOGY

The methodology section will explain how the statistical model was developed, and why the operationalization of central bank speeches, both as a text embedding with an RNN and as a structural variable in a sentiment analysis, can provide an additional benefit to the literature, that has hitherto mainly focused on employing variations of vector auto-regressive models to approach the impact of geopolitical risk on inflation and other macroeconomic indicators. Critically, due to the novelty of the `deeptrafo` package, I will use the opportunity to shed some light on the intricacies that emanate from the reconciliation of DCTMs approach with deep neural networks and bidirectional encoder representations from transformers, in this instance FinBERT, that require separate `Python` environments to work in `Rstudio`.

Due to the breadth of the methods harmonized in the model herein, and the diverse range of issues that arise when using DCTMs in this context, several iterations of the model have been built. The baseline model is the most parsimonious. It only includes structural macroeconomic variables relevant to the discussion herein. It will be used as foundational model guiding the methodology and robustness strategy to avoid repetition. Where the other models diverge from the parsimonious baseline model, they will be described in more detail. This section will focus on why the DCTM approach was chosen and the advantages of its application. Thereafter, will be an in-depth analysis of the CB speeches and natural language processing with a qualitative assessment of both the forward-looking sentiment and the keywords-in-context (hereafter: KWIC) analysis of geopolitical risk in ECB speeches. Following that will be the description of the model specification, which includes a matrix describing the differentiated approach between the various models proposed herein. A more holistic characterization of the DCTMs can be found in section 4, and the empirical strategy for the robustness, which ranges from shift models, bootstrap testing and partial dependence plots, to linear quantile regressions, Kernel-density point predictions, and the production of conditional distribution functions can be found after the results, as they are discussed in chapter 6.

To quickly reiterate, the research question of this paper aims to investigate how references of geopolitical risks in central bank communication and macroeconomic uncertainty impacts the distribution of inflation expectations in the Euro Area. So far, geopolitical risk has been primarily examined in the context of real inflation and central banks have primarily been treated as respondents to geopolitical risk with macroeconomic policy. In anyway, CBs have since the beginning of the inclusion of geopolitical risk in macroeconomic research moved from tacit agents passively affected by geopolitical risks, to more active players that aim to predict or prevent inflationary pressures stemming from geopolitical risk shocks. This more active role of CBs is most acutely apparent when looking at the references of geopolitical risk over time, where the full-scale Russian invasion of Ukraine in 2022 can be considered a *caesura* in the threats geopolitical risks pose for the mandate of price stability in the Euro area.

Central Bank communication, as we recall, is part of the monetary policy itself as it has an direct statistically and economically significant effect on the volatility of commodity prices (Hayo, Kutan, and Neuenkirch 2012, p. 248). Furthermore, non-monetary news in CB communication has shown to have an effect on asset price behavior (Cieslak and Schrimpf 2019, pp. 31–32) and informal communication of central bankers is used to inform financial markets of interest rate changes and so also has an impact on those markets (Hayo, Kutan, and Neuenkirch 2012, p. 249). All of these factors have further contributed to the use of the ECB speeches as corpus for the statistical analysis that aims to be econometrically rigorous while also showcasing analytical acumen for the phenomenon at the heart of this thesis.

A priori the use of a deep conditional transformation model may seem counterintuitive, given the overt dominance of methodological strategies based on some form of a VAR, impulse response function based on Cholesky decomposition or even wavelet-based approaches in the GPR literature hitherto, but it boasts a variety of advantages that have led to an empirical strategy tailored to their effective use. For example, the DCTM can reconcile the direct insertion of text embeddings from the corpus into the same model with structural parametric variables. Moreover, as the model can capture nonlinearities, the estimations capture how geopolitical risks interact with VIX and CB sentiment more accurately than linear models. Moreover, because it models the full distribution, not just the mean, the entire range of possible inflation expectations can be assessed. Incorporating deep learning for complex patterns is advantageous because neural networks automatically learn risk factors' effects on inflation uncertainty.

Again, the empirical objective of this thesis is to investigate how macroeconomic uncertainty and central bank communication co-determine the conditional distribution of inflation expectations in the euro area. To this end, the analytical framework must accommodate not only mean effects but also the possibility of asymmetries, fat tails, and distributional shifts, which represent phenomena that are empirically salient in monetary policy forecasting.

In light of these requirements, I employ the Deep Conditional Transformation Model (DCTM) developed by Baumann, Hothorn, and Rügamer (2021) and adapted to an R package by Kook et al. (2024) , which extends the classical transformation model to a flexible, semiparametric setting via neural architectures. The DCTM paradigm serves as a methodological *via media* between structural econometric modeling and fully nonparametric machine learning as it retains interpretability through its explicit use of transformation functions, while allowing for highly nonlinear and interaction-rich functional forms through deep neural substructures.

From a statistical standpoint, DCTMs estimate a monotonic transformation of a latent response variable that is assumed to follow a canonical distribution, such as the standard normal in this instance, whose transformation is conditioned on observed covariates. In contrast to linear regression or quantile regression, which offer only local summaries (e.g., the conditional mean or specified quantiles), the DCTM permits inference over the entire conditional distribution of the target variable. This feature is indispensable when the academic discussion and the policy interest lies in understanding tail risks or distributional skewness in inflation expectations. For example when computing the probability mass assigned to outcomes exceeding inflation targets, or scenarios consistent with deflationary dynamics.

A distinctive feature of this study is the incorporation of textual data from the ECB speech corpus as a dynamic predictor of inflation expectations. While traditional models often treat text-derived variables as exogenous sentiment indices or dummy classifications, which were used in some of the alternative DCTM configurations, the present framework endogenizes the communicative content of monetary authorities by modeling speech as a structured temporal signal.

To this end, speech transcripts are processed into sequential embeddings and passed through a Recurrent Neural Network (RNN), more precisely a Long Short-Term Memory (LSTM) architecture, which learns latent representations of the evolving semantic and emotional content of central bank communication. The final hidden state of the RNN serves as a learned, high-dimensional embedding that is directly input into the interaction component of the DCTM. The text embeddings are not in the parsimonious baseline model which only includes the structural macroeconomic parameters, but in a secondary derivative model. The vicissitudes between the model which includes purely parametric, continuous variables, and those with text endogeneity will be further elaborated in the model specification section. Suffice to say, the inclusion of RNN-based text embeddings within a transformation model represents, in effect, a fusion of econometric reasoning and machine-learned abstraction. This hybridization facilitates a nuanced analysis by plotting how the inflation expectation distribution shifts, depending on how the ECB communicated, i.e. in a more dovish or geopolitically neutral tone, *ceteris paribus*.

By uniting the theoretical rigor of transformation models with the empirical expressiveness of deep learning, the DCTM affords a uniquely suitable framework for interrogating the conditional dynamics of inflation expectations. The model not only aligns with the distributional nature of the research question but also treats language as a first-order economic signal. A deliberate methodological commitment that is both timely and theoretically coherent within the evolving literature on central bank communication.

The methodological ambitions of this thesis, which models the conditional distribution of inflation expectations using a deep semi-parametric approach, while simultaneously incorporating high-dimensional textual data, imposed a number of formidable technical and computational challenges. The implementation rested upon the integration of two distinct yet interdependent computational frameworks: a distributional modeling architecture based in R, and a natural language processing pipeline based in Python. The coordination of these tools necessitated the deployment of multiple virtual environments, facilitated by the Anaconda distribution, in order to maintain internal consistency across divergent dependency trees and ensure reproducibility under nontrivial constraints, which are further dissected in section 3.4

Overall, the research question will take into account the survey of professional forecasters by the ECB on a quarterly basis. The thesis aims to understand how geopolitical risk effects inflation expectations at the one-year-ahead horizon. Using DCTMs enables the amalgamation of diverse data types including quantitative and text-as-data. As such, this paper will consider the geopolitical risk index, as well as the economic policy uncertainty index, which can also account for election uncertainty implications that can also have geopolitical consequences, the VIX to account for confidence shifts in financial markets, and topic modeling and sentiment analysis from a corpus of ECB speeches, both regarding the frequency of geopolitical risk references and the sentiment towards geopolitical risk.

The focus on the Eurozone was a deliberate choice, because of data availability and relevance. The Euro area has been under immense pressure in the aftermath or the Russo-Ukrainian War and the War in Gaza, as well as other conflicts and terror acts and so the issue of geopolitical risk contagion in the market is of great importance to the ECB. Also by focusing on one region, deeper and more granular conclusions can be drawn, that will shed some light in highly complex environment. A preliminary assessment on the quantitative relationship between geopolitical risk indicators and sentiment to inflation expectations has been performed in Table 4. The benchmark OLS regression results provide preliminary insights into the relationship between various uncertainty and sentiment indicators and the dependent variable. Statistically significant predictors at the 5% level include log(GPR) ($p = 0.005$), (log)EPU ($p = 0.016$), and (log)WUI ($p = 0.019$), suggesting that geopolitical risk, economic policy uncertainty, and world uncertainty are meaningful drivers. Notably, (log)GPR and (log)EPU have positive coefficients, indicating a direct relationship, while (log)WUI is negatively associated with the outcome. Other variables such as VIX, EMV, and sentiment-related indices do not show statistical significance in this baseline specification, pointing to limited explanatory power in this setup.

3.1 Macroeconomic Structural Variables

In order to capture the evolving landscape of macroeconomic uncertainty and its influence on inflation expectations, this study incorporates a range of structural indicators that quantify risk, volatility, and policy ambiguity across economic and financial domains. These variables serve not merely as controls but as theoretically grounded channels through which agents form expectations about future inflation. They are selected for their empirical prominence, robustness across datasets, and complementary nature in characterizing different dimensions of uncertainty.

The guiding rationale for their inclusion lies in the theoretical literature on expectation formation, particularly within the framework of noisy information models and rational inattention. In such models, economic agents update their beliefs in response to signals about policy, markets, and geopolitical conditions. Structural uncertainty indices serve as measurable proxies for the informational environment in which expectations are formed. By conditioning the full distribution of inflation expectations on these indices, the model is equipped to detect whether tail risks or asymmetric inflation fears co-move with macro-financial volatility, a relationship of considerable importance to central bank credibility and forward guidance efficacy.

Geopolitical Risk (GPR) The Geopolitical Risk Index (GPR), constructed by Caldara and Iacoviello (2022), quantifies newspaper-based coverage of geopolitical tensions, threats of war, and military action. GPR captures both realized and latent geopolitical shocks, reflecting agents' concern with global uncertainty that may influence commodity prices, risk premia, and inflationary pass-through channels. Within this framework, GPR functions not just as a backdrop to inflation expectations but as a variable potentially shifting the skew and dispersion of the expectation distribution.

Economic Policy Uncertainty (EPU) The Economic Policy Uncertainty Index (EPU), developed by Baker, Bloom, and S. J. Davis (2016), captures the frequency of newspaper articles discussing economic policy uncertainty. EPU embodies perceived indecisiveness or unpredictability in fiscal, monetary, or regulatory policy, which may prompt precautionary savings, delay in investment, or shifts in inflation risk premia. Its inclusion in the model tests whether expectations react not just to policy content, but to the perceived reliability and coherence of that policy.

Equity Market Volatility (VIX) Often dubbed the "fear" index, the VIX measures expected volatility in U.S. equity markets derived from option prices. While not Eurozone-specific, it serves as a globally watched barometer of financial market stress and risk aversion. The logic of including the VIX rests on the increasingly interconnected nature of global financial systems, where episodes of market turbulence can rapidly transmit to inflation dynamics via credit spreads, capital flows, and commodity prices. It is produced by the Chicago Board Options Exchange (CBOE Global Markets 2023).

Equity Market Volatility Index (EMV) The Equity Market Volatility (EMV) tracker focuses more specifically on equity-market-related uncertainty as expressed in financial press coverage. It distinguishes itself from the VIX by being narrative-based rather than derived from market pricing, offering a complementary lens on perceived versus priced volatility. Inclusion of both EMV and VIX allows the model to parse sentiment-driven uncertainty from market-implied expectations. It is also produced by Baker, Bloom, and S. J. Davis (2023)

World Uncertainty Index (WUI) The World Uncertainty Index (WUI), compiled by Ahir, Bloom, and Furceri (2018), aggregates textual uncertainty signals across more than 140 countries, offering a panoramic view of global uncertainty shocks. It has the advantage of being internationally harmonized and forward-looking, potentially capturing channels that are not country-specific but exert inflationary pressures through global demand, supply chains, or commodity pricing.

All structural variables are integrated into the DCTM as linear shift covariates, capturing their direct effect on the location of the inflation expectation distribution. Some, such as GPR, may also be modeled via deep neural interaction components to capture nonlinearities or threshold effects. All variables are standardized prior to model fitting to ensure comparability and numerical stability. By including these structural variables, the model gains the ability to dissect how different forms of macro-financial uncertainty map not only to the level but to the shape, including the variance, skew, and kurtosis, of inflation expectations. This enhances the interpretability of model outputs and offers a richer policy-relevant understanding of inflation expectation dynamics in uncertain times.

3.2 Natural Language Processing of ECB Speeches

To investigate how the ECB communicates geopolitical risk and how such communication may influence inflation expectations, this thesis employs a structured natural language processing (NLP) framework tailored to central bank discourse. The focus lies on extracting sentiment and thematic structures from ECB speeches, with a view toward integrating these features into a predictive model of inflation expectations derived from the Survey of Professional Forecasters (SPF). The analysis begins with the construction of a speech corpus comprising ECB policy addresses and official statements over a defined time horizon. Each document is preprocessed through standard NLP techniques, including tokenization, lowercasing, and removal of non-informative content. To preserve context and semantic nuance, the corpus is embedded using FinBERT, a domain-specific variant of BERT pretrained on financial texts. This model generates dense, contextualized representations of textual units at the sentence or paragraph level, which are particularly suited to capturing the subtleties of central bank language.

To uncover latent thematic structures within the corpus, topic modeling is performed using a Bayesian nonparametric approach based on the Hierarchical Dirichlet Process (HDP). Unlike conventional models such as Latent Dirichlet Allocation (LDA), HDP does not require the number of topics to be specified *ex ante*. Instead, it infers the optimal number from the data. Estimations are conducted via Reversible Jump Markov Chain Monte Carlo (RJMCMC), which allows the model to transition between topic spaces of varying dimensionality. This approach is particularly appropriate in a policy context, where the prominence and nature of geopolitical concerns can shift substantially over time.

Sentiment analysis is carried out using a dual-layered method. First, FinBERT is employed to classify sentences as positive, neutral, or negative with respect to sentiment, based on its deep contextual understanding. Second, a lexicon-based approach is implemented with preliminary reference to the Loughran-McDonald sentiment dictionary, which provides a benchmark grounded in financial text analysis. To enhance interpretability and qualitative insight, Key Word in Context (KWIC) analysis is applied to sentiment-bearing phrases, especially around terms like "war," "sanctions," or "energy security," which are indicative of geopolitical risk framing. The outputs of the topic modeling and sentiment analysis processes, namely topic prevalence scores and sentiment metrics, are then aggregated and prepared for integration into the empirical stage of the thesis. Specifically, they serve as covariates in a control DCTM, where the target variable is the SPF inflation expectation. This allows the investigation to go beyond point forecasting and explore how shifts in the ECB's tone and thematic emphasis condition the entire distribution of inflation expectations, capturing potentially nonlinear or asymmetric effects.

Overall, this NLP approach combines the semantic sophistication of large language models with the interpretive depth of lexicon-based techniques and the probabilistic rigor of Bayesian topic inference. It enables a systematic assessment of how central bank communication about geopolitical risk may shape market and expert perceptions of future inflation under conditions of uncertainty, while also taking into account sentiment polarity of forward-looking statements

However, producing the components for the natural language pipeline to that effect came with considerable challenges which includes both computational and conceptual issues. Some initial difficulties like model input length errors due to FinBERTs 512-token limit could more easily be resolved by truncating or excluding overly long sentences. Some documents didn't contain geopolitical risk-related references at all, which is not in itself entirely surprising, given both the novelty and recently growing awareness regarding the role of geopolitical risk in macroeconomic policymaking, and the primary objective of central banks, which is only indirectly linked to geopolitical fragmentation overall. Hence, the corpus was filtered based on keyword regrex and `nchar > 5` was ensured before passing to FinBERT. Nonetheless, when aggregating the CB speeches by quarter to build a harmonized sentiment and gpr score, it became apparent that some quarters didn't contain gpr references entirely. Those `NA` values consequently reduced the overall number of quarters included in the model incorporating NLP-based parametric values, i.e. DCTM 3, by 11 quarters from the initial 105 quarters that were included based on observational data.

The conceptual issues primarily related to overspecification and noisy predictions. The aggregation of sentence-by-sentence applications to the FinBERT pipeline created variance from classification confidence and somewhat weak signals due to sparse keyword matches, which were supposed to be mitigated with averaging and minimum-length filters, which as later to be seen was only moderately successful. These weak signals also emanated from overlapping topics and the fact that even geopolitically tagged sentences weren't clearly associated with positive or negative sentiment. This also makes sense, as random qualitative assessments of speeches have shown that members of the ECB community are more likely to circumscribe geopolitical fragmentation to the extent relevant for their work, rather than attribute normative values to them, which (used to be) considered outside their *prima facie* mandate.

The most pressing issue that endangered the utility of the approach overall concerned fine-tuning the model. An initial aggregation of GPR scores by quarter showed no clear trajectory of growing relevance, as would be expected by looking at the geopolitical risk index by Caldara and Iacoviello (2022) and what would be anticipated by the literature hitherto. A closer look at what the model identified as geopolitical references made it abundantly clear that the ubiquity of relevant terminology made distinctions attributable to the phenomenon explored herein very difficult. This was particularly difficult to ascertain with regards to stylistic devices that are commonly used in the respective speeches. Metaphors, similes and synecdoches are difficult to train in a FinBERT model, in a way that is neither to restrictive which would omit relevant sentences, but also not too lenient, which amplifies the issue of noise and overspecification, for the quantitative and deep estimations that follow thereafter.

This is particularly true for the geopolitical risk score, rather than the forward-looking sentiment indicator, as the former was not only used in the DCTM 3 as a parametric variable, but also lay the foundation for the corpus used as text embeddings as an endogenous, non-parametric variable, to which the DCTM is foremost employed and specifically qualified. Using the entire ECB corpus has shown to be ineffective to discern how particularly hawkish or dovish language surrounding geopolitical risk could effect the inflation expectations by the SPF. Hence, DCTM 2 only included sentences that were proven to include geopolitical risk references based on keywords-in-context (KWIC) methods. The logic behind the measure was threefold. Firstly, the goal was to generate a cleaner signal that removes noise. Secondly, the idea rested upon the aspiration to create a model that has a stronger focus, which is difficult when using deep neural networks that are based on implicit learning. Lastly, this operationalization ensures semantic alignment between the text and the target variable with regards to the hypothesis.

The effort aligns with the goals and functions of the DCTM 2, which is based on embeddings and RNN layers that transform into dense representations for the transformation function to model the

CDF. It is designed as an end-to-end supervised learning model, as the DCTM learns the relevance of the text from the target, hence-why the input text needs variation and mapping from the target. This necessitates pre-filtering for possible relevance. In laymen's terms: the DCTM needs to be fed geopolitically-relevant statements because otherwise it will not be able to discern the relevant signal that impacts the target variable for this exercise.

The difficulty that arose when cleaning the KWIC-GPR model from the aforementioned issues with a focus on the stylistic devices, can be exemplified when conducting specific qualitative assessments of the tables produced by the model, which were critical for an iterative process that fine-tuned the model with manual, learned experience, as other known statistical or methodological approaches were insufficiently successful. And so, while tedious and prone to errors, a manual restriction from FinBERT-based solutions was conducted to reduce some of these issues, to which ultimately no strategy is immune.

It is important to note that the following examples are chosen to illustrate insufficiencies produced by the model, and so the sentences are reconstructed after tokenization, lemmitization, stop-word removal, *et cetera*. The KWIC model was specified to include 6 tokens pre and post keyword. For example in speech no. 317 from the 1st of March 2021 the model recognized a geopolitical reference in the statement: "country's Wirtschaftswunder after the Second World War It was critical to Germany's swift", which is an erroneous selection for the aim of this paper. But text no. 45 from the 20th of September 2024 discusses "pandemic since the 1920s the worst conflict in Europe since the 1940s". Excluding phrases like "conflict of interest", which is ubiquitously used in CB language is much easier, than these edge cases. The model struggles tremendously in distinguishing between actual geopolitical references and those that are only tangential, because the relevant topic is contextually very nuanced. Even conceptually it is difficult to derive a model that correctly identifies and includes statements regarding the worst invasion since the second world war, without also including statements referring to the strongest output growth since WWII.

Ex ante irregularities like these create uncertainty and make it difficult for the deep conditional transformation models to attribute shifts in the inflation expectations to speech patterns and sentiments, because their derivative signals are neither strong nor well-isolated, despite a strong theoretical argument that sentiment polarity in geopolitical risk references can effect the economic assumptions upon which the inflation expectations pronounced in the SPF are based. The table 1 contains a random sample of the kind of gpr-related sentences that are initially included in the estimations by the model. A visualization of the FinBERT-based sentiment and gpr scores with inflation expectation movement over time is included in the Appendix A. Specifically, Figure 3 shows the raw data as produced by the FinBERT model, and Figure 4 illustrates the final inclusion of KWIC-GPR cleaned sentences with manual restriction enforcement.

Document ID	GPR-Related Sentence
text219	... our sovereignty EMU has a strong **geopolitical** dimension crystallising its members commitment to ...
text135	... pandemic the brutal Russian invasion of **Ukraine** and the ensuing energy crisis have ...
text219	... the public sectors Even before the **invasion** of Ukraine the attainment of the ...
text1743	... swept Europe after the First World **War** were a prominent collective test whose ...
text55	... volatility but they too have a **geopolitical** impact as they can strengthen the ...
text768	... different exchange rate regimes since the **war** It was also in other words ...
text1403	... of earlier and more gradual financial **sanctions** within the macroeconomic surveillance framework Discretion ...
text120	... Blog 24 February Before the war **Russia** and Ukraine accounted for about 25 ...
text877	... the refugee crisis or to counter **terrorism** Therefore the fact that Europe is ...
text122	... triggered by Russia's unjustified war against **Ukraine** For a given inflation outlook the ...

Table 1: KWIC-GPR cleaned NLP model output - Selected Sentences

3.3 Model Specification

Let $Y \in \mathbb{R}$ denote the 12-month-ahead inflation expectation. The Deep Conditional Transformation Model (DCTM) expresses the conditional distribution function $F_{Y|X}(y \mid x)$ as a transformation of a known latent distribution function F_Z (assumed to be standard normal), through a strictly increasing transformation function $h(y \mid x)$:

$$F_{Y|X}(y \mid X) = \Phi\left(h_\theta(y) - \eta(X)\right)$$

where:

- Y denotes the 12-month-ahead inflation expectation;
- X is a vector of covariates including Geopolitical Risk (GPR), Economic Policy Uncertainty (EPU), VIX, and Central Bank Speech Sentiment (CBS);
- $\Phi(\cdot)$ is the standard normal cumulative distribution function;

- $h_\theta(y)$ is a monotonic transformation of the response variable parameterized by θ;
- $\eta(X)$ is a structured predictor, which may include both linear components and deep neural network transformations (e.g., ReLU-activated layers).
- $\mathcal{H}_{\text{mono}}$ denotes the space of monotonic transformation functions. The function $h(y \mid x)$ is parameterized as:

$$h(y \mid x) = a(y)^\top \theta(x),$$

where $a(y) \in \mathbb{R}^K$ is a vector of basis functions (e.g., Bernstein polynomials), and $\theta(x)$ is a vector of transformation coefficients that depend on the covariates x. We decompose $\theta(x)$ into structured (linear) and deep (nonlinear) components:

$$\theta(x) = \theta_0 + M_{\text{struct}}(x_s) + f_{\text{deep}}(x_d),$$

where:

- θ_0 is an intercept term,
- x_s denotes structured predictors (e.g., GPR, EPU),
- x_d denotes deep predictors (e.g., text embeddings or sentiment indices),
- $M_{\text{struct}}(x_s) = \sum_j \theta_j x_{s,j}$ is a linear shift in the transformation,
- $f_{\text{deep}}(x_d)$ is a neural network transformation learned jointly with the model.

Model 1: Baseline DCTM with Macroeconomic Structural Variables

The simplest model includes only standardized macroeconomic variables as structured predictors. In this specification, $\eta(X)$ consists exclusively of structured linear predictors.

$$x_s = (\text{GPR}, \text{EPU}, \text{EMV}, \text{VIX}, \text{WUI}), \quad x_d = \emptyset.$$

$$h(y \mid x) = a(y)^\top (\theta_0 + \theta_{\text{GPR}} \cdot \text{GPR} + \theta_{\text{EPU}} \cdot \text{EPU} + \cdots + \theta_{\text{WUI}} \cdot \text{WUI}).$$

Model 2: DCTM with Text Embeddings via Neural Network

This model enriches the structure by incorporating ECB speech embeddings $T \in \mathbb{R}^d$ via a deep neural network, hence $\eta(X)$ represents both structured macroeconomic covariates and nonlinear deep neural representations of ECB speech embeddings.

$$x_s = (\text{EPU}, \text{EMV}, \text{VIX}, \text{WUI}), \quad x_d = T,$$

$$h(y \mid x) = a(y)^\top (\theta_0 + M_{\text{struct}}(x_s) + f_{\text{text}}(T; \phi)).$$

Model 3: DCTM with FinBERT Sentiment and GPR Scores

This version uses interpretable scalar features $S \in \mathbb{R}^p$ from FinBERT where $\eta(X)$ is driven by structured macro variables and shallow neural representations of FinBERT-derived sentiment indices.

$$x_s = (\text{EPU}, \text{EMV}, \text{VIX}, \text{WUI}), \quad x_d = S,$$

$$h(y \mid x) = a(y)^\top \left(\theta_0 + M_{\text{struct}}(x_s) + f_{\text{sent}}(S; \psi)\right).$$

The deep component f_{text} is implemented as a multilayer perceptron with ReLU activation functions.

Estimation and Inference

In all models, F_Z is assumed to be the standard normal distribution and $a(y)$ is based on a Bernstein polynomial basis of order K. Estimation is conducted via maximum likelihood, using gradient descent, early stopping, and dropout where necessary to regularize the deep components.

3.4 Technical Implementation and Computational Architecture

As previously mentioned, the implementation of the operationalization rested upon the integration of two distinct yet interdependent computational frameworks: a distributional modeling architecture based in `R`, and a natural language processing pipeline based in `Python`. The coordination of these tools necessitated the deployment of multiple virtual environments, facilitated by the Anaconda distribution, in order to maintain internal consistency across divergent dependency trees and ensure reproducibility under nontrivial constraints.

At the heart of the econometric specification lies the Deep Conditional Transformation Model (DCTM), estimated via the `deeptrafo` package in R. This model architecture builds upon the classical transformation model framework by allowing a monotonic, data-driven transformation of a latent base distribution conditioned on both linear covariates and nonlinear features captured through deep neural network components. The implementation of the DCTM in R, however, revealed the limitations of interfacing deep learning libraries within traditional statistical environments. The model is underpinned by the `Keras` and `TensorFlow` back-ends, which, although powerful, exhibit significant sensitivity to versioning and dependency resolution. Small discrepancies between the R wrapper versions of `Keras/TensorFlow` and their native `Python` counterparts often produced opaque errors, ranging from coercion failures to silent tensor misalignments. In particular, the interplay between reticulate, the R-to-Python interface, and Anaconda environments required meticulous manual synchronization, including explicit version pinning for libraries such as `tensorflow==2.9,` `keras==2.9,` and legacy protocol buffer dependencies such as `protobuf<4`. This ultimately made the employment of the DCTM and FinBERT possible. It is advisable to keep track of the packages hardwired into the different R-environments via Anaconda to avoid NLL errors from overwritten package libraries and to keep entirely segregated R projects for R-deep, R-reticulate, R-sentiment, et cetera. A list of specific package version compatibilities in each Python environment is available from the corresponding author upon request.

In parallel, the text-based component of the analysis was implemented using `Python` and the Hugging Face `transformers` library. The FinBERT model, a financial-domain fine-tuning of BERT, was employed to extract contextual embeddings and sentiment scores from European Central Bank (ECB) speeches. These embeddings were not merely treated as exogenous regressors, but were conceptualized as endogenous signals of central bank tone, possessing the potential to shift the shape of the inflation expectation distribution. Given the computational burden and hardware dependencies associated with transformer-based models, the entire textual pipeline was executed outside the `R` environment and the outputs subsequently serialized into structured `.csv` matrices for integration into the `R`-based modeling framework. The necessity of bridging environments in this way reflects the disciplinary boundaries between natural language processing and econometrics, and speaks to the methodological interdisciplinarity of the approach adopted herein.

The confluence of these systems; R for statistical estimation, Python for deep linguistic representation, and Anaconda for environment encapsulation, constituted a fragile but ultimately tractable architecture. Yet this fragility manifested itself in several instructive ways. Training routines for the DCTM were acutely sensitive to memory usage, especially when high-order Bernstein basis expansions (e.g., order 30 or higher) were used to approximate transformation functions. Dropout regularization, batch size adjustments, and early stopping heuristics were employed not for theoretical elegance, but as pragmatic responses to numerical instability. Similarly, the inability to produce robust predictions at the tails of the inflation distribution without manual calibration of the target variable's scale, due to the presence of outliers or extreme values, illuminated the sometimes stark gap between model theory and empirical viability.

What emerges from this technical inquiry is a broader reflection on the epistemology of modern empirical research. The sheer power of deep distributional models and transformer-based embeddings lies not in their elegance alone, but in their capacity to absorb heterogeneous data modalities ranging from text and indices to volatility measures and yield inferentially-rich structures. Yet their implementation is replete with operational fragility. The need for explicit dependency management, layered environments, and cross-language orchestration is not an incidental nuisance but a fundamental feature of cutting-edge applied work. That these models can be made to function at all across such fault lines is the primary feature of this research endeavor, rather than just the empirical results, which are statistically and economically less significantly powerful than orthodox methodological approaches would have yielded.

4 DEEP CONDITIONAL TRANSFORMATION MODELS

The DCTM is based on the `deeptrafo` package by Kook et al. (2024). It serves as the core econometric framework of the thesis: the Deep Conditional Transformation Model (DCTM). The DCTM is a flexible, distributional modeling approach designed to estimate the entire conditional distribution of a response variable, in this case, inflation expectations from the Survey of Professional Forecasters (SPF), given a set of predictor variables derived from both textual and macro-financial sources.

Traditional econometric models, such as OLS or quantile regression, estimate only specific moments (e.g., the mean or median) of a response variable. However, in the context of expectation formation under uncertainty, forecasters are often concerned with the full distribution of inflation outcomes, including tail risks and skewness, rather than central tendencies alone. Conditional transformation models address this by estimating a monotonic transformation of the response variable such that the transformed outcome follows a standard distribution, which is usually either uniform or normal, conditional on covariates. The DCTM extends this idea by allowing the transformation function itself to be approximated by a neural network, enabling flexible, nonlinear interactions between covariates and the shape of the inflation expectation distribution. All the machine learning and neural network models have been built with reference to Goodfellow, Bengio, and Courville (2016).

The model estimation proceeds by minimizing the negative log-likelihood of the conditional distribution over the observed data, regularized to avoid overfitting. The model is trained using mini-batch gradient descent on a time-aligned panel of ECB speeches, macro indices, and SPF forecast data. Sentiment scores are extracted from ECB speeches using a FinBERT model and lexicon-based methods. These scores capture the tone and emotional valence of ECB communication around geopolitical events. The scores are aggregated over speech windows to align with SPF forecast rounds. Additional context is incorporated via topic prevalence from hierarchical topic modeling, which identifies recurring geopolitical themes.

These features allow the model to evaluate how changes in ECB tone and topic emphasis affect not only average inflation expectations, but also the uncertainty and skewness of those expectations. To contextualize the textual sentiment data within broader geopolitical and economic dynamics, the model also includes the Geopolitical Risk Index (GPR), which captures the frequency of newspaper articles reporting geopolitical tensions, wars, and terror threats, the Economic Policy Uncertainty Index (EPU), which reflects uncertainty about fiscal, regulatory, and monetary policy based on keyword frequency in major news sources, and the CBOE Volatility Index (VIX), a market-based measure of implied volatility in equity markets, often interpreted as a proxy for global financial stress or "risk-off" sentiment. These indices serve as exogenous controls that help isolate the unique contribution of ECB communication to inflation expectations, net of global and domestic uncertainty shocks.

The Deep Conditional Transformation Model provides a robust and interpretable platform for understanding how ECB communication, when contextualized by sentiment analysis and broader measures of geopolitical and financial uncertainty, affects inflation expectations. By modeling the full conditional distribution of SPF forecasts, the DCTM captures not only expected levels of inflation but also the uncertainty and perceived risks surrounding those expectations, thereby offering deeper insight into the expectations channel of monetary transmission under geopolitical stress.

Whereas Model 1 adopts a classical specification that assumes a linear and additive influence of structured macroeconomic covariates on the distribution of inflation expectations, Model 2 represents a substantial extension in both conceptual architecture and technical implementation. Specifically, Model 2 harnesses the flexibility of the Deep Conditional Transformation Model (DCTM) to incorporate high-dimensional, non-parametric representations of textual information, namely sentence embeddings derived from central bank communication, directly into the modeling of the conditional distribution of inflation expectations.

This shift in approach is not merely one of dimensionality, but one of epistemology. Model 1 rests on the premise that macroeconomic variables such as geopolitical risk indices and financial volatility proxies exert interpretable, linear shifts on the location and shape of the conditional distribution. By contrast, Model 2 allows for the conditional distribution itself to be learned as a flexible function of latent semantic features extracted from the central bank's textual output. These embeddings, produced via pretrained FinBERT models and aggregated quarterly, encode nuanced linguistic variation that would otherwise remain opaque to linear modeling strategies.

The theoretical innovation of `deeptrafo` lies in its architecture, which generalizes classical transformation models by allowing the parameters governing the distributional transformation to depend on both structured and unstructured covariates. In this framework, the transformation function

$$F_{Y|X}(y \mid X) = \Phi\left(h_\theta(y) - \eta(X_s, f_{\text{text}}(T; \phi))\right)$$

determines how the base distribution (e.g., a standard normal) is warped into the conditional distribution of the observed variable, and is parameterized not only by macro variables but also by deep neural networks acting on text-derived embeddings. This capacity to accommodate high-dimensional, non-linearly interacting predictors makes the DCTM uniquely suited for incorporating textual information.

The practical implementation of Model 2 required substantial pre-processing and technical scaffolding. First, the relevant textual content was extracted using a keyword-in-context (KWIC) filtering strategy centered on geopolitical risk terminology. Each document segment was then passed through the FinBERT language model to obtain 768-dimensional sentence embeddings. These embeddings were aggregated by quarter and joined with macroeconomic controls. Due to the high dimensionality of the resulting feature space, direct inclusion into an R formula object exceeded the parsers capacity and triggered recursion-depth errors.

To circumvent these issues, I constructed the transformation formula incrementally by partitioning the embeddings into manageable subsets and assigning each to a separate neural pathway within the model. These sub-networks were subsequently integrated into the shift function via named deep components in the `deeptrafo` framework. The resulting model retained the capacity to learn arbitrarily complex and non-monotonic interactions between language and inflation expectations while remaining estimable within the computational constraints of the R environment.

In sum, Model 2 advances beyond the limitations of parametric specification by capturing the dynamic and latent informational content of central bank communication in a distributional forecasting context. It transforms raw textual information, often considered analytically elusive, into a structurally integral component of probabilistic inference. This methodological expansion not only enhances the model's descriptive capacity but also lays the groundwork for more granular interpretive diagnostics, such as counterfactual text experiments or local explanation methods, the latter of which exceeds the scope of this paper.

5 RESULTS

Based on the literature, the following assumptions were made *ex ante* regarding the empirical results, upon which their interpretation is based. This preliminary assessment of expected results illustrate the mechanical properties of change depictions in the DCTMs more clearly, given that the employment of deep neural network architecture makes them non-obvious from the outset. If negative sentiment shifts the GPR CDF to the right, markets perceive central banks' geopolitical concerns as increasing risk. If positive sentiment reduces tail risk, that could indicate that markets trust central banks to mitigate geopolitical uncertainty and if the variance increases, it would indicate that uncertainty rises when CBs discuss geopolitics negatively. Correspondingly, if DCTMs show geopolitical risk shifts the entire inflation expectation distribution upward, then the ECB may need tighter monetary policy or bolder reactions to counteract inflation fears from geopolitical instability. Secondly, if only the upper tail of the inflation distribution expands, then inflation tail risks increase, signaling higher uncertainty rather than widespread inflation pressure. Lastly, if CB speech sentiment weakens the link between GPR and inflation expectations, it would indicate that central bank credibility matters and so if policymakers frame geopolitical shocks as manageable, inflation expectations remain anchored.

The empirical findings from the estimation and evaluation of the DCTMs applied to inflation expectations paints a more diverse and heterogeneous picture than assumed in the anticipatory interpretation, thereby supporting the overarching notion in the literature that overwhelmingly ascertains both deflationary and inflationary pressures to geopolitical risk. The analysis is grounded in quarterly macroeconomic and textual data, with a focus on assessing the predictive capacity and distributional insights of the proposed model specifications and special attention that is paid to the calibration of the models and to the structure of the predicted conditional distributions, as visualized via cumulative distribution functions (CDFs).

The baseline DCTM, which is the most parsimonious model, captures key macroeconomic determinants of professional forecasters' twelve-month-ahead inflation expectations, using inputs such as the GPR, EPU, VIX, and the WUI. As shown in the main calibration plot in figure 9, the predicted values from the baseline model align closely with the observed inflation expectations reported in the ECB's Survey of Professional Forecasters (SPF). The scatter plot of predicted versus observed values, plotted against the 45-degree reference line, illustrates a dense clustering along the diagonal, indicating a high degree of model calibration. This suggests that the DCTM is capable of recovering the central tendency of the conditional distribution of inflation expectations with notable precision.

The shift coefficients derived from DCTMs offer valuable insight into how each predictor influences the conditional distribution of 12-month-ahead inflation expectations. In the DCTM framework, these coefficients originate from the transformation function and represent additive, linear shifts in the distribution, in effect, they determine how covariates move the location of the cumulative distribution function (CDF) to the left or right. This interpretability makes the linear shift coefficients especially useful as a complement to the more flexible but opaque nonlinear transformation layers encoded by the deep learning components of the model. The shift coefficients for the respective DCTM models can be found in Table 2.

In the case of DCTM 1, which includes traditional macroeconomic uncertainty indicators, the shift coefficients highlight a strong, positive influence of the sentiment index (+2.558) on the distribution of inflation expectations. This indicates that a more optimistic or hawkish tone in ECB forward-looking language pushes the conditional distribution of expected inflation to the right, consistent with higher forecasted outcomes. In contrast, geopolitical risk (GPR) and the World Uncertainty Index (WUI) both exhibit negative coefficients (-0.182 and -0.270, respectively), implying that increased global or political instability has a dampening effect on inflation expectations. Economic Policy Uncertainty (EPU) has a strong positive coefficient (+2.126), reflecting the market's tendency to associate policy ambiguity with higher inflation risks or tighter monetary conditions. Other variables like the VIX and EMV have marginal or moderate effects, with VIX being essentially neutral and EMV showing a small positive shift (+0.370), likely due to media narratives around economic instability.

Building on this baseline, DCTM 2 introduces a novel input modality: text embeddings. Instead of relying on scalar sentiment scores or keyword-matching heuristics, this model incorporates high-dimensional numerical representations derived from ECB speeches using pretrained language models. These embeddings are capable of encoding complex semantic patterns in monetary policy communication. In the `deeptrafo` framework, such high-dimensional textual inputs are processed by a deep neural network embedded within the transformation model. The architecture allows the deep network to flexibly learn nonlinear transformations of the text data that interact with the conditional distribution of the target variable.

Unlike standard parametric models that impose strict additive or linearity assumptions, the `deeptrafo` architecture enables the model to adjust the entire shape of the conditional distribution based on features extracted from text. This makes it particularly well-suited for working with unstructured language data, where effects may be complex, nonlinear, and distributional rather than mean-shifting.

In DCTM 2, which extends the baseline model by replacing scalar sentiment indicators with these high-dimensional text embeddings derived from ECB speeches, a different picture emerges. Here, the shift coefficient for the latent text embedding vector, represented as a single compressed dimension, is notably large (+1.259), suggesting that semantic content captured by deep neural representation significantly influences the expectation distribution. Compared to DCTM 1, the effect of GPR becomes more negative (-0.478), and WUI also becomes more negative (-0.567), while EPU remains strongly positive (+1.829). These shifts likely reflect the model's ability to internalize more nuanced language signals through the embedding layer, which indirectly reweights the influence of macro variables. The inclusion of the text embedding appears to absorb some of the information carried by the simpler sentiment index and GPR score variables in other models, potentially leading to a reconfiguration of their explanatory power.

The contrast between DCTM 1 and DCTM 2 illustrates the trade-off between transparency and expressiveness in model design. DCTM 1 and 3 are more interpretable, especially for policy communication and literature alignment, but they rely on aggregated or simplified textual signals. DCTM 2, in incorporating text embeddings, captures deeper linguistic nuance and interaction effects at the cost of reduced linear traceability. Nevertheless, the fact that the embedding variable receives a strong positive shift coefficient in DCTM 2 confirms that the textual signal is substantively meaningful and predictive, just not easily disentangled. This highlights the usefulness of hybrid models that combine linear interpretability with deep learning flexibility, allowing researchers to explain core effects while preserving the capacity to detect complex relationships.

The performance of DCTM 2 is illustrated by two key outputs: the calibration plot and the conditional CDFs. The calibration plot in Figure 10 demonstrates that predicted means based on the text embeddings align well with actual observed inflation expectations. The spread of predictions remains centered around the identity line, suggesting that the model is able to extract meaningful predictive information directly from the text embeddings. The absence of systematic bias or heteroskedasticity in this plot further indicates that the deep neural component is well-tuned to the embedding space.

Most importantly, the conditional CDFs derived from DCTM 2 reveal that the text embeddings are not merely noise features but have a measurable effect on the predicted shape of the inflation distribution. In the annexed Figure 13 CDFs are plotted for two scenarios: one with high embedding effect and another with low embedding effect. These scenarios correspond to ECB speeches whose embeddings lie at the upper and lower quartiles along key latent dimensions identified by the neural network.

The result is a visible and interpretable shift in the CDFs. The distribution associated with the high embedding effect is stochastically larger than the one with low embedding effect. This implies that textual information embedded in ECB speeches is systematically associated with shifts in the perceived likelihood of higher inflation. In other words, the model has learned to associate certain latent semantic configurations in central bank language with upward or downward movements in inflation expectations. This effect was not observed in DCTM 3, where sentiment scores and GPR sentence-level averages produced flat or overlapping CDFs. The fact that DCTM 2 captures this variation while DCTM 3 does not points to the representational advantages of deep embeddings over scalar sentiment indicators.

The practical real-world interpretation derived from the CDF plotted in figure 13 provides observational shifts. For speeches with low geopolitical content, the CDF rises more steeply and earlier. This implies that the model assigns a higher probability to lower inflation expectations in such contexts. For speeches with high geopolitical emphasis, the CDF curve shifts rightward, indicating a higher probability mass in the upper tail, that is, a greater model-estimated likelihood of elevated inflation expectations. In a nutshell this signifies that when ECB communication contains more intense or prominent references to geopolitical risks, i.e military conflict, global instability, or diplomatic crises, professional forecasters interpret this language as a signal of increased inflationary pressures over the following year.

These findings do not negate the possibility that some geopolitical risk references can induce deflationary pressures in the Euro area, but it shows that overwhelmingly inflationary association of geopolitical risk on inflation expectations in the SPF. This could be indicative of a plethora of underlying phenomena including the weight geopolitical instability has on energy markets and supply chains, which apply upwards pressure on the cost channel. Other possibilities include the constrained or reactive policy toolkit by the ECB, which raises uncertainty into inflation trajectories, as the ECB cannot actively change geopolitical fragmentation or risks derived thereupon. As previously discussed, the ECB is moving from tacit acceptance of geopolitical circumstances towards their anticipation, but that does not imply changing behavior including through more hawkish language and tighter monetary policy, as they may remain sub-optimal in environments dominated by transitory supply chain disruption that are more likely to be seen through by CBs. Lastly, we know that households and firms base their inflation expectations on current inflation and past inflation experiences, and while experts, both financial and non-financial, as surveyed in the SPF are less susceptible to these fallacies doesn't mean they are immune, nor does it mean that historical inflation during high geopolitical risk isn't a factor that guides future inflation expectations, particularly when we know that orthodox economic theory suggests inflation expectations affect real inflation through channels like the wage spiral.

The partial dependence plot for the first principal component (PCA 1) of the ECB speech text embeddings in DCTM 2 in Figure 15 reveals a U-shaped relationship with the predicted mean of the conditional inflation expectation distribution. This shape suggests a non-monotonic and nonlinear effect, where both high and low values along PCA 1 are associated with elevated predicted inflation expectations, while intermediate values correspond to a local minimum in the model's output.

This U-shaped structure implies that two distinct types of ECB speech content, those located at opposite ends of the principal semantic axis, are both interpreted by the model as indicative of heightened inflation expectations. Meanwhile, more neutral or moderate textual content (i.e., values of PCA 1 near zero) is associated with lower expected inflation. Because PCA 1 captures the dominant direction of variance across all embedding vectors, it represents the most semantically influential dimension of ECB communication as perceived by the language model.

One possible interpretation is that speeches with strong rhetorical emphasis, whether assertively dovish or hawkish, register at the tails of PCA 1 and are perceived by forecasters, or by the model learning from forecaster responses, as reflecting greater inflationary uncertainty or volatility. For instance, speeches dominated by crisis language, policy urgency, or geopolitical references may map to one extreme, while highly technical or structurally reformist language may occupy the other. In both cases, the language diverges from the typical, measured tone of ECB communication, potentially signalling regime shifts or policy inflection points and hence prompting higher inflation expectations.

The central trough in the PDP corresponds to moderate, institutional ECB language, more procedural and less emotionally or politically charged, which is associated with more anchored expectations. This fits well with the theoretical expectation that policy predictability and communication stability help to anchor inflation expectations, whereas unusual or extreme shifts in tone signal risk and ambiguity, which the model learns to associate with the upper tails of the inflation distribution.

Importantly, this nonlinearity reinforces the value of the deep learning component in the DCTM architecture. Traditional linear models or additive sentiment indices would likely have failed to capture this bidirectional effect. The U-shape in the PDP offers strong evidence that the model has internalized complex semantic relationships and that embeddings provide predictive content that is not reducible to monotonic effects and that the deep network is sensitive to textual polarity or extremity in central bank communication, assigning greater inflationary signal to more emphatic or atypically expressed messages.

In terms of robustness, the curvature of the PDP also lends credence to the conditional CDF shifts observed in the same model. The effect is not driven by a uniform trend across the embedding space but rather by departures, on either side, from a normatively stable discourse. Thus, the PDP provides an essential diagnostic window into how the neural architecture maps linguistic signals to macroeconomic expectations, revealing a richer and more nuanced relationship than a simple linear or monotonic function would suggest.

In sum, DCTM 2 stands out as the model that best leverages unstructured textual data. Through its deep neural architecture, it learns meaningful distributional effects of text without relying on hand-crafted sentiment scores. The shift in conditional CDFs across embedding conditions provides compelling evidence that central bank communication, when encoded as embeddings, can influence not just the mean but the entire distribution of inflation expectations. These findings underline the empirical value of combining conditional transformation models with modern NLP representations for macroeconomic forecasting.

Extending the model, DCTM 3 introduces additional text-derived covariates, specifically sentiment scores and GPR-specific sentiment extracted from ECB communication. Despite the theoretical relevance of these indicators, their empirical contribution appears limited in the context of the full model. The predicted values in DCTM 3, as illustrated by figure 11 where the scatter of predicted versus actual values displays no material correlation with the diagonal benchmark. The R-squared value, although not reported in isolation, can be inferred to remain small given the visual fit of the model.

The conditional cumulative distribution functions (CDFs) derived from the DCTMs further substantiate the model's predictive behavior. In the annexed figures, conditional CDFs are plotted across representative quarters or feature values. In the case of DCTM 3, these include distributions conditional of high and low GPR speech sentiment. The distributions do not maintain their logical form, which should increase monotonically over the support of the response variable, as the variation across conditions is marginal. For both sentiment and GPR sentiment, the average CDFs across quantile-split groups (top and bottom quartiles) are nearly indistinguishable, as seen in figure 14. This visual result implies that, conditional on the macroeconomic structure, the sentiment-based features contribute minimal additional predictive information.

Moreover, the absence of a pronounced shift in the CDFs across the textual covariates suggests that the model's learned conditional distributions are primarily shaped by macroeconomic fundamentals rather than linguistic signals from ECB communication. This is especially noteworthy given that the sentiment measures were carefully constructed using FinBERT-based classification of GPR-relevant sentences. The negligible difference between high and low GPR or sentiment cases, both in terms of predicted means and distributional shape, highlights the limits of text-derived features when integrated into a heavily macro-informed predictive framework.

Overall, the empirical findings confirm that the DCTM framework is well-suited to modeling the conditional distribution of inflation expectations. The models provide not only calibrated point predictions but also coherent conditional distributions that can be analyzed and visualized through CDFs. However, the marginal explanatory power of text-based features in this context raises important questions regarding their robustness and interpretability, which are explored in the following sections.

To conclude, the shift coefficients reinforce key findings from the broader empirical analysis. Forward-looking ECB communication, whether captured through sentiment indices or embeddings, consistently increases inflation expectations, while geopolitical and global uncertainty reduce them. The fact that these effects persist, and even strengthen, under different DCTM architectures underscores their overall robustness and the central role of monetary communication in shaping macroeconomic forecasts.

6 ROBUSTNESS

To ensure the reliability and interpretability of the modeling framework, this chapter presents a series of robustness tests designed to evaluate both the statistical and substantive performance of the NLP-driven predictive model. The focus is on assessing how well the DCTM captures the distributional characteristics of SPF inflation expectations using features derived from ECB geopolitical speech sentiment and topic prevalence. Because building the deep models has proven to be far more complex than anticipated, the robustness strategy presented here is an illustrative effort to show how the validity and reliability of the model could be further ascertained in statistically meaningful and methodologically sound manner. Deep computational models based on neural networks are known to incur difficulties in the interpretability of their findings and this exercise is no different. A trade-off exists between the employment of hidden layers and the direct incorporation of text embeddings that may capture non-linear dynamics and produce probabilistic distribution functions versus frequentist or Bayesian statistics in traditional approaches that yield more directly interpretable results, that can also be corroborated more easily with alternative identification strategies. Their utility is however limited, given the presumption that central bank communication on geopolitical risk is inherently linked to aleatoric uncertainty, and not epistemic uncertainty, which is underscored by the success of the DCTM 2 in comparison with the next iteration in DCTM 3.

The first robustness check involves an assessment of the model's in-sample log-likelihood. Because the DCTM estimates the full conditional distribution of the target variable rather than a point forecast, log-likelihood provides a natural metric for evaluating the accuracy of the estimated conditional density functions. A higher log-likelihood indicates better fit, meaning the predicted distributions are more sharply aligned with the actual realizations of SPF inflation expectations. To provide a baseline for comparison, the log-likelihood of the DCTM is evaluated against that of simpler benchmark models, such as a normal a OLS regression that can be found in Table 4.

The summary statistics presented in Table 3 provide a comparative evaluation of the DCTMs and an OLS benchmark in terms of root mean squared error (RMSE), log-likelihood, and Akaike Information Criterion (AIC). These metrics offer distinct yet complementary perspectives on model performance: RMSE captures predictive accuracy of the conditional mean, log-likelihood reflects overall distributional fit, and AIC penalizes complexity to reward parsimony.

Among the models, DCTM 3, which incorporates FinBERT-derived sentiment scores and GPR-specific sentiment, demonstrates the best overall balance of predictive performance and model fit, with the lowest AIC (312.99) and a substantially lower RMSE (1.8925) relative to DCTM 1 and DCTM 2. Interestingly, DCTM 2, which integrates high-dimensional transformer-based text embeddings, performs worse than DCTM 3 in terms of RMSE and AIC, despite achieving a more favorable log-likelihood. This result suggests that while embeddings may improve the model's ability to capture fine-grained distributional features, they do not necessarily enhance mean prediction accuracy. The higher AIC of DCTM 2 (565.27) also reflects the complexity penalty associated with using over-parameterized neural representations, due to the high number of embeddings used, which significantly amplified the number of κ -variables in the calculations.

By contrast, DCTM 1, relying solely on macroeconomic uncertainty indices, performs worst across all metrics, reinforcing the added value of textual inputs. The OLS model, while not probabilistic, provides a surprisingly low RMSE (0.5261) and the lowest AIC (175.13), highlighting its effectiveness in approximating central tendency. However, its log-likelihood (-78.5869) remains inferior to DCTM 2, indicating that OLS may offer strong point prediction but lacks capacity to model the full conditional distribution. Taken together, these results confirm that incorporating textual data, particularly structured sentiment signals, substantially enhances the performance of conditional forecasting models, and that while deep architectures like DCTMs can match or exceed traditional models in flexibility, their utility must be weighed against increased complexity.

Additional modeling strategies were employed to validate the findings from the DCTMs and assess the sensitivity of distributional insights to alternative methodologies. Specifically, linear quantile regression and nonparametric kernel conditional density estimation were implemented. These approaches serve as valuable complements to the DCTMs, offering alternative ways to model the conditional distribution of inflation expectations without relying on neural network architectures or probabilistic transformation layers. Quantile regression estimates the effect of covariates across different points of the outcome distribution and is particularly suited to capturing distributional heterogeneity with minimal functional assumptions. Kernel density estimation, in turn, provides a fully nonparametric method for estimating conditional densities based solely on observed data structure and local bandwidth smoothing. While both approaches offer detailed distributional perspectives, they do not yield cumulative distribution functions (CDFs) in the same form as the DCTMs, and are therefore not directly comparable in terms of probabilistic calibration or transformation modeling. Their purpose is instead to serve as interpretive and methodological cross-checks that help assess the plausibility and consistency of DCTM-derived conclusions.

The results of the linear quantile regressions provide valuable insight into how the relationship between the explanatory variables and inflation expectations varies across the conditional distribution of the dependent variable. As shown in Table 5, the coefficient on GPR increases consistently from 0.365 at the 25th percentile to 0.368 at the 75th percentile, suggesting that geopolitical risk exerts a stronger upward influence on higher inflation expectations than on the lower tail. This pattern is corroborated by the quantile coefficient path plot, illustrated in Figure 5, where GPR exhibits a positive and slightly upward-sloping trajectory across quantiles. In contrast, the effect of EPU remains relatively flat and slightly negative throughout the distribution, indicating that general economic policy uncertainty does not exhibit significant quantile-dependent behavior in this specification.

The `speech gpr score` variable shows its strongest effect at the median quantile (0.355), with a moderate decline at the 75th percentile (0.169), pointing to a nonlinear influence that peaks at the center of the distribution. This may indicate that GPR-relevant speech content from the ECB affects forecasters' inflation expectations most strongly when expectations are moderate, while extreme expectations are less sensitive to communication content, although this trajectory is quantitatively weak and relatively uniform over quantiles. Meanwhile, the sentiment index shows only marginal variation across quantiles, with a mild negative effect at the lower tail (-0.084 at the 25th percentile) and a small positive effect at the median (0.031), suggesting its influence is weakly differentiated across the distribution. Taken together, the quantile regression results reinforce the finding that GPR and, to a lesser extent, `speech gpr score`, shape not just the central tendency but also the dispersion of inflation expectations, a key motivation for applying distribution-sensitive models such as DCTMs.

The kernel density estimates presented in Figure 8 illustrate how the distribution of inflation expectations shifts conditionally on the level of geopolitical risk (GPR), offering a nonparametric counterpart to the distributional findings derived from the DCTM framework. The estimated conditional density for high-GPR observations (red solid line) exhibits a clear rightward shift compared to the low-GPR distribution (blue dashed line), indicating that elevated geopolitical tensions are associated with a higher probability of inflation expectations falling in the upper range of the support. Conversely, the low-GPR distribution is more sharply peaked and centered around lower inflation expectations, reflecting a more anchored outlook in times of relative geopolitical calm. This pattern aligns with theoretical expectations from the literature, which suggest that geopolitical shocks, through channels such as commodity price volatility, trade disruptions, and uncertainty, can dis-anchor expectations and increase tail risk, particularly in the upward direction.

These kernel-based findings are also consistent with the conditional cumulative distribution functions (CDFs) produced by DCTM 2, where a similar rightward shift in the distribution was observed under high GPR embedding conditions. Figure 7 plots the a stylized graph of the expected outcome for various GPR scenarios in point predictions based on a Kernel Distributions, as later corroborated by Figure 8, and the CDF from DCTM 2 in the high text embeddings scenario as represented in Figure 13. The DCTM results further captured the distributional widening and tail behavior in a probabilistically coherent manner, reinforcing the robustness of the nonparametric kernel estimates. Together, these results support the hypothesis that geopolitical risk is a significant determinant not only of the central tendency but of the shape and spread of inflation expectations, thus validating the use of distribution-sensitive models and justifying the modeling choice to incorporate textual signals into the conditional transformation framework.

A second validation step is the use of calibration plots to compare predicted and empirical distributions. These plots test whether the quantiles predicted by the model are consistent with the observed empirical frequency of inflation expectations. In practice, this involves plotting the cumulative distribution function (CDF) implied by the DCTM against the empirical CDF of the SPF data. If the model is well-calibrated, the calibration curve should closely follow the 45-degree line. Deviations from this diagonal indicate over- or under-estimation of specific quantiles, highlighting where the model may systematically misrepresent tail risks or central tendencies in inflation expectations. Quantile-quantile (Q–Q) plots are also employed as a complementary diagnostic, comparing specific percentiles (e.g., the 10th, 50th, and 90th) of the predicted distributions with those of the actual SPF data across time. A well-fitting model will produce a Q–Q plot that approximates a straight line through the origin with slope one. The Q-Q plots were excluded from Annex A for the sake of brevity, but can be supplemented upon request to the corresponding author. Beyond the aforementioned strategy that was ultimately pursued, robustness was considered to be further assessed through a wide array of complementary methods including quantified degrees of distributional similarity estimations with statistical distance measures such as the Wasserstein distance or Kullback–Leibler (KL) divergence or the examination of stability measurements over time, through a series of rolling-window regressions, all of which were omitted for the sake of brevity, but could be conducted in future research. Most chiefly of which would be the employment of these methods for counterfactual analysis, which also could not be conducted herein. However, Appendix B contains a descriptive strategy for a possible counterfactual analysis, as an extension to the work presented in this thesis.

An additional robustness check deemed pertinent to the credibility of the models derived in this thesis suggested comparing model performance when using different sentiment extraction techniques. Specifically, the DCTM is estimated using sentiment features generated from (i) the FinBERT language model, and (ii) checked with lexicon-based approaches using the Loughran–McDonald dictionary as cross-reference. Differences in model fit and calibration are assessed to determine whether deep contextual embeddings offer superior predictive value, or whether simpler, rule-based approaches suffice for this application. The extensive considerations that were discussed in the qualitative assessment section of the NLP processing chapter were ultimately deemed sufficient for the interpretability-versus-accuracy trade-off that is necessary when using FinBERT in favor of richer sentiment representation but without the transparency and auditability of lexicon-based approaches. The robustness checks presented in this chapter provide evidence that the DCTM, when fed with NLP-derived sentiment and topic indicators from ECB geopolitical communication, produces well-calibrated, distributionally accurate forecasts of SPF inflation expectations. Log-likelihood comparisons, and calibration diagnostics point to a model that is statistically sound and substantively relevant. These findings alone support the use of central bank speech analysis as a viable input into expectations modeling under geopolitical uncertainty.

7 DISCUSSION

The results presented in the empirical section demonstrate that deep conditional transformation models (DCTMs) are a flexible and powerful framework for modeling the full distribution of inflation expectations. However, as with any empirical strategy, their application in this context raises important methodological, interpretive, and practical considerations. The predictive performance of the DCTMs, particularly the baseline and macro-focused versions, confirms the suitability of transformation models for capturing distributional information. The calibration plots indicate a high degree of alignment between predicted and observed SPF inflation expectations, particularly in the DCTM 1 and DCTM 2 specifications. However, the structure of the predicted cumulative distribution functions (CDFs) varies substantially across model configurations and reveals some key limitations that merit discussion.

One particularly noteworthy issue emerges from the analysis of DCTM 1, which includes only macroeconomic uncertainty indices as predictors. While the calibration plot shows reasonable fit, the corresponding CDF displays an unusual shape, characterized by a sharp vertical drop at the upper tail of the support. This drop suggests that the model assigns a disproportionately high probability to inflation expectations not exceeding a certain threshold, approximately 4.8 percent. This is not a technical error but a faithful reflection of the empirical distribution of the target variable which never exceeds 4.8 percent in the entire sample. This modeling constraint, however, introduces a conceptual disconnect: although SPF expectations remain anchored below 5 percent, actual realized inflation in the euro area peaked at over 10 percent in October 2022 (Hoffmann et al. 2025, p. 11), driven by energy shocks and geopolitical events. This would raise a speculative argument that the SPF itself is not a neutral reflection of expected inflation among experts, but itself is materially normative to the anchoring of the inflation expectation target and more broadly the symmetric inflation target at 2 % over the medium term. This argument does not imply that the SPF is not objective or neutral, but it connotes to an environment that is institutionally or subconsciously geared to left-hand skews to prevent a vicious cycle or the SPF acting as a self-fulfilling prophecy.

More broadly, this asymmetry reveals an important limitation in the data-generating process rather than the model *per se*. The model cannot learn patterns that are not present in the training data; hence, it cannot predict tail inflation risks that professional forecasters themselves failed to anticipate. This reflects a broader issue in expectation modeling: survey expectations often exhibit lagged and muted reactions to macroeconomic shocks, especially unprecedented ones. The DCTM is technically capable of modeling fat-tailed or skewed distributions, but only to the extent that such features are represented in the empirical support. This remains true irrespective whether this failure to predict high inflation scenarios is the consequence of intentional safeguarding against premonitory properties of the SPF, or simply a failure to predict inflationary repercussions from geopolitical risks, which are subject to high-frequency variation and thereby create significant volatility that cannot be captured at even the shortest time horizons for inflation expectations.

Furthermore, while DCTM 2 successfully integrates text embeddings and demonstrates distributional variation based on the latent content of ECB speeches, DCTM 3, which uses aggregated sentiment and GPR scores, shows minimal contribution from the textual features. The CDFs across sentiment or GPR score quantiles are flat and nearly indistinguishable, indicating that these scalar NLP features carry little predictive signal when included alongside powerful macroeconomic controls. This may be due to the oversimplification of complex textual content into single-index values or a mismatch between the textual signal and the structure of SPF expectations. But it also reflects a deeper challenge in using NLP-derived covariates: while transformer embeddings preserve semantic richness, scalar scores often collapse this information and dilute its predictive utility. From a methodological standpoint, `deeptrafo` provides a flexible infrastructure for modeling such interactions, especially through its ability to learn nonlinear transformations and distributional effects via neural network components. However, the model's performance is contingent on the quality, coverage, and granularity of the input data. Several design choices, such as the decision to treat the response variable as discrete due to its limited support, introduce both interpretive clarity and statistical rigidity. For example, the CDFs are only evaluated at observed support points, limiting their smoothness and introducing staircase-like behavior in some plots.

There are also trade-offs involved in model complexity. While deep architectures can learn high-dimensional interactions, they are more prone to overfitting and require careful regularization. In this thesis, model fitting was constrained by data availability, particularly the relatively short time series of SPF data. This limits the model's ability to generalize or detect rare tail events, such as the inflation spike of 2022. But the granularity of the data available was also severely limited by the fact that the SPF is only available on a quarterly basis, which inhibited the temporal resolution of model training, the alignment of textual data or macroeconomic shocks that occur at higher frequency, and sensitivity to real-world volatility. This issue was somewhat subdued by the use of deep models to recover latent temporal shifts, meaning that even after aggregation on a quarterly basis, DCTMs with deep neural nets were used to learn nonlinear and interaction-based effects of embeddings and macroeconomic variables. This ensures that the model can *pull out* meaningful relationships from aggregated features that may be lost in linear averaging. Moreover, the empirical strategy for the textual data opted for the inclusion of all ECB speeches, that were either averaged based on their features like in the sentiment analysis or the embeddings or based on selected representative examples, as seen in the KWIC application of CDF illustration. This approach allowed the model to retain some linguistic variance even under coarse time aggregation, that stemmed from semantic flattening via averaging. Alternative modeling strategies could include bootstrapped simulations to enrich the tail of the response distribution, Bayesian versions of transformation models to incorporate prior beliefs about inflation risks, or hybrid models that combine macroeconomic forecasting with textual attention mechanisms to better capture forward guidance. Another promising direction is the use of quantile transformation forests or distributional boosting algorithms that can model heteroskedasticity and tail risk explicitly.

Finally, while the empirical analysis focused on ECB communication and macroeconomic uncertainty, there remains scope to incorporate financial market expectations (e.g., from inflation swaps), cross-country inflation data, or high-frequency monetary surprises to enhance the responsiveness of the model. Importantly, none of the models used in this thesis captured the latent potential for inflation to exceed the historical SPF range, highlighting the persistent gap between realized macro volatility and expectation-based forecasts. The use of inflation-linked swaps, while associated with other methodological challenges because they include risk premia in addition to the inflation expectations, retains a much more detailed picture of inflation expectations because of their high-frequency that is far more responsive to geopolitical risk shocks. This strategy and the inclusion of mean and variance data for the structural macroeconomic variables would have enhanced the model and make the statistical analysis more powerful, but time limitations and most importantly, difficulties in the data gathering process for frequent ILS data prevented the adoption of that strategy herein. Future research could include this methodology, which could serve as an *argumentum a fortiori* to the findings presented here. While the empirical performance of the DCTMs is encouraging in terms of both calibration and distributional interpretability, the robustness of the models remains a subject of concern. Several important econometric and modeling techniques commonly used in inflation forecasting are not incorporated in the current framework, and their omission limits the ability to benchmark or statistically validate the findings under alternative specifications.

First, the thesis does not formally evaluate model performance using predictive accuracy tests such as the Diebold-Mariano test. These tests are specifically designed to assess whether forecast errors from competing models are statistically significantly different. Their inclusion would have enabled a more rigorous comparison of the DCTMs against baseline econometric models such as autoregressive (AR) benchmarks or macroeconomic VARs. In the absence of such formal comparisons, the evaluation of DCTM performance remains largely qualitative and visual, relying on calibration plots and conditional distribution shapes rather than inferential forecast error metrics.

Second, the static nature of the models presented here means that time-variation in relationships between predictors and inflation expectations is not explicitly modeled. While the DCTM architecture can capture complex nonlinearities and interaction effects, it does not allow for evolving parameter estimates over time. This limitation could be addressed through the incorporation of state-space models and Kalman filtering, which would permit recursive updating of latent state variables and dynamic adjustment of coefficients in response to new information. A Kalman-filtered version of the DCTM could track shifting expectation regimes, such as the structural break introduced by the post-pandemic inflation surge. Such temporal flexibility is essential in settings where expectations are known to adapt over time in response to policy shifts or global events. Moreover, the models assume independence across time and do not incorporate serial correlation or persistence in inflation expectations directly. While the use of lagged macro variables offers a partial remedy, a proper time-series framework, such as a recurrent neural network or temporal convolutional architecture, could model dependencies across sequential periods more directly and robustly. From a methodological perspective, the thesis also does not incorporate probabilistic scoring rules such as the Continuous Ranked Probability Score (CRPS), which would have provided a quantitative measure of how well the predicted conditional distributions match the empirical distributions of observed inflation outcomes. CRPS or log score evaluations are standard in the forecast comparison literature and would have enriched the model validation process beyond RMSE and visual fit.

As discussed, another methodological limitation lies in the choice of aggregation. Due to the quarterly frequency of the SPF, all textual and macroeconomic features were collapsed to the same temporal resolution, potentially smoothing away relevant high-frequency variation. This issue could be addressed through mixed-frequency modeling techniques, such as MIDAS regressions, which allow higher-frequency predictors to inform lower-frequency targets without discarding temporal richness. Alternatively, hierarchical models could exploit the multiple observations within each quarter to inform the distribution of expectations at the quarter level. In terms of model architecture, the DCTMs used here are restricted to a single-output setting. They predict one distributional outcome 12-month-ahead inflation expectations, without modeling the joint distribution of other economic variables such as GDP growth, unemployment, or interest rates. A multivariate extension of the DCTM framework, potentially with copula-based transformations or structured priors, could provide a richer probabilistic description of macroeconomic expectations.

Finally, while the thesis uses advanced NLP methods to extract sentiment and embeddings from ECB speeches, it does not explore attention-based mechanisms that could learn which parts of the text are most informative for predicting inflation expectations. Transformer models with integrated attention could be used not only for text embedding, but also to highlight specific linguistic features or rhetorical structures that are most strongly associated with shifts in expectations. This would offer a more interpretable link between textual communication and economic forecasting. In sum, while the DCTM framework offers a powerful platform for distributional forecasting, several methodological extensions and robustness checks remain unexplored. The inclusion of dynamic state-space modeling, formal predictive accuracy testing, probabilistic scoring metrics, and attention-based NLP architectures would significantly strengthen the robustness and interpretability of the empirical findings. These remain promising directions for future research which could include richer data sources, improved text representation strategies, and models explicitly designed to capture tail behavior and rare events.

8 CONCLUSION

The dynamics of inflation expectations have recently occupied a central place in macroeconomic research and policy analysis. As both a determinant and reflection of future economic conditions, inflation expectations influence consumption, investment, and wage-setting behavior, and thereby shape the transmission of monetary policy itself. In times of heightened geopolitical instability, policy uncertainty, and inflationary shocks, as witnessed in the aftermath of the COVID-19 pandemic and the Russian invasion of Ukraine, the accurate modeling of these expectations becomes not merely an academic exercise but a critical input into real-world economic governance.

This thesis addresses a timely and underexplored question: how do textual signals from central bank communication, particularly regarding geopolitical risk, shape professional forecasts of future inflation? In doing so, it brings together strands of literature on monetary policy communication, textual analysis, and probabilistic forecasting to assess whether central bank language contains predictive content that can be systematically modeled. Prior research has established that forward guidance and monetary narratives can influence expectations, yet the evidence on whether unstructured text, especially regarding geopolitical themes, alters the shape or distribution of inflation expectations remains limited. The literature reviewed in this thesis would suggest that central bank communication acts as a signal channel through which agents update beliefs. Accordingly, one might expect that hawkish or conflict-laden language would translate into a rightward shift in the inflation expectation distribution, either via an anticipatory response to supply-side shocks or through perceived constraints on monetary accommodation.

Methodologically, this thesis introduces the application of Deep Conditional Transformation Models (DCTMs) to the domain of inflation expectation forecasting. Unlike traditional models that estimate point forecasts or conditional means, DCTMs provide a full probabilistic description of the target variable conditional on covariates. Their flexibility stems from the use of transformation functions and the integration of deep neural networks, which allow the model to capture nonlinearities, high-dimensional interactions, and distributional shifts. This is particularly advantageous when working with complex, unstructured inputs such as text embeddings derived from central bank speeches. By learning the conditional distribution of inflation expectations, the DCTM framework moves beyond central tendencies and into the realm of risk-aware and uncertainty-sensitive forecasting.

The empirical analysis evaluates three model variants. DCTM 1 establishes a macroeconomic baseline using well-known uncertainty indices. DCTM 2 incorporates high-dimensional text embeddings derived from ECB communication, allowing the model to learn latent semantic effects. DCTM 3 uses aggregated sentiment and GPR-specific sentiment scores to test a more interpretable but lower-dimensional textual signal. The results show that the DCTM framework performs well in modeling the observed range of SPF inflation expectations. Predicted values align closely with observed outcomes, and the models provide smooth and coherent conditional CDFs. However, the performance varies across specifications. DCTM 2 emerges as the most effective in capturing distributional shifts associated with textual content. Its CDFs reveal a meaningful difference in inflation expectations depending on the semantic characteristics of the input text, suggesting that deep embeddings successfully encode relevant forward guidance signals. In contrast, DCTM 3 fails to uncover any substantial effect from scalar sentiment scores, indicating that such features may be too coarse or noisy when used alongside macroeconomic controls.

At the same time, the analysis surfaces several limitations that bear on the interpretation of the findings. Notably, DCTM 1 produces a CDF with a sharp drop at the upper tail, a feature directly attributable to the empirical support of the response variable, which never exceeds 4.8 percent. This stylized fact reflects the historically anchored nature of SPF forecasts, even during episodes of actual inflationary spikes. For example, while inflation in the euro area exceeded ten percent in October 2022, SPF forecasts remained subdued, possibly due to institutional forecasting inertia or a lag in expectation adjustment. The DCTM faithfully reproduces this observed distribution, but it also highlights a blind spot: when data do not include tail events, the model cannot generate predictive tail risks. This underscores the challenge of forecasting under structural breaks or regime shifts, particularly when expectations are slow to internalize new information.

The relevance and novelty of this thesis lie in its integration of deep learning methods within a formally interpretable probabilistic forecasting framework. While prior work has applied sentiment analysis or NLP-based regressions to macroeconomic outcomes, few studies have modeled how text alters the shape of the entire expectation distribution. This thesis demonstrates that DCTMs can serve as a bridge between the semantic richness of textual data and the quantitative demands of economic modeling. By doing so, it provides a blueprint for incorporating qualitative signals, especially those related to geopolitical risk, into probabilistic macroeconomic forecasting. The thesis also highlights the representational superiority of text embeddings over scalar sentiment scores, suggesting that future work should further exploit high-dimensional and context-aware textual features. As such, this thesis should be understood more as a proof-of-concept rather than a standalone empirical analysis of the hypothesis itself, that is primarily aimed to serve as an impetus for future research.

Looking ahead, the findings offer several important conclusions and open questions. First, central bank communication, when properly encoded, does contain predictive content for inflation expectations, though this signal may not be linear or uniformly interpretable. Second, the shape of expectation distributions is inherently tied to the information set and beliefs of forecasters, which are themselves shaped by institutional, informational, and behavioral constraints. Third, probabilistic modeling approaches such as DCTMs are particularly suited to environments of heightened uncertainty, where risk is not symmetric and expectations may be anchored or non-Gaussian. Future research could extend this framework to multivariate settings, incorporate higher-frequency expectation measures where available, or deploy attention-based models to identify which linguistic signals are most salient for macroeconomic forecasting. Moreover, the integration of regime-switching dynamics or Kalman-filtered state-space layers could allow for more adaptive learning in the face of structural change.

In a nutshell, the key takeaways from the findings in this paper are threefold. Firstly, for policymakers the most important conclusion is the fact that the way geopolitical developments are framed in official speeches can measurably shape professional expectations, not just via stated forecasts but via latent signals embedded in language. Secondly, forecasters do or should account for linguistic shifts in ECB communication, especially in times of geopolitical tensions, as their assessment can enhance prediction models for inflation or policy rate movements as their deployment is conditional on the assessment by the ECB regarding their properties. The literature has shown, that hawkish stances and tighter monetary policy may be advantageous if energy or supply shocks are persistent, and so the interpretation of geopolitical risk by CBs matters, even though the geopolitical risk itself is not a precursor to inflation. Lastly, for researches in the field of natural language processing in macroeconomics, this paper has validated the use of deep text embeddings and nonlinear models to capture implicit forward guidance in central bank language, while those effects have been missed or omitted when traditional sentiment analysis was employed.

In an era where uncertainty is both macroeconomic and geopolitical, the ability to model how expectations are formed from words, not just numbers, remains both a scientific and policy-relevant endeavor. This thesis contributes to that goal by demonstrating that language matters, and that modern tools can help us understand how it shapes our economic futures.

References

Acemoglu, Daron and Alireza Tahbaz-Salehi (2024). "The macroeconomics of supply chain disruptions". In: *Review of Economic Studies*, rdae038. ISSN: 0034-6527.

Ademuyiwa, Idris and Pierre L. Siklos, eds. (2019). *The global interplay between trade, geopolitical and commodity shocks*.

Afunts, Geghetsik, Misina Cato, and Tobias Schmidt (2024). "Inflation expectations in the wake of the war in Ukraine". In: *Journal of behavioral and experimental economics* 113. ISSN: 2214-8043. DOI: 10.1016/j.socec.2024.102303.

Ahir, Hites, Nicholas Bloom, and Davide Furceri (2018). "The World Uncertainty Index". In: *VoxEU.org*. Available at https://www.worlduncertaintyindex.com.

Allayioti, Anastasia et al. (2024). "A look back at 25 years of the ECB SPF". In: *ECB Occasional Paper* 2024/364.

Alsagr, Naif and Stefan van Hemmen (2020). "Oil rent, geopolitical risk and banking sector performance". In: *International Journal of Energy Economics and Policy* 10.5, pp. 305–314.

Anttonen, Jetro and Markku Lehmus (2024). "Impact of geopolitical surprises on euro area inflation varies case by case". In: *Bank of Finland*.

Asadollah, Omid et al. (2024). "Geopolitical risk, supply chains, and global inflation". In: *The World Economy* 47.8, pp. 3450–3486. ISSN: 0378-5920. DOI: 10.1111/twec.13585.

Backer, Bruno de et al. (2023). "Inflation expectations and monetary policy". In: *NBB economic review* 7, pp. 1–37.

Baker, Scott R., Nicholas Bloom, and Steven J. Davis (2016). "Measuring Economic Policy Uncertainty". In: *The Quarterly Journal of Economics* 131.4, pp. 1593–1636. DOI: 10.1093/qje/qjw024.

— (2023). *Equity Market Volatility Tracker*. https://www.policyuncertainty.com/EMV_monthly.html. Accessed: 2025-03-01.

Barro, Robert J. (1974). "Are Government Bonds Net Wealth?" In: *Journal of Political Economy* 82.6, pp. 1095–1117. DOI: 10.1086/260266.

Baumann, Philipp F. M., Torsten Hothorn, and David Rügamer (2021). "Deep Conditional Transformation Models". In: *Machine Learning and Knowledge Discovery in Databases. Research Track*.

Ed. by Nuria Oliver et al. Cham: Springer International Publishing, pp. 3–18. ISBN: 978-3-030-86523-8.

Baumeister, Christiane (2023). *Pandemic, War, Inflation: Oil Markets at a Crossroads?*

Bednarski, Lukasz et al. (2025). "Geopolitical disruptions in global supply chains: a state-of-the-art literature review". In: *Production Planning & Control* 36.4, pp. 536–562. ISSN: 0953-7287. DOI: 10.1080/09537287.2023.2286283.

Bjørnland, Hilde C. (2022). "The effect of rising energy prices amid geopolitical developments and supply disruptions". In: *1892-2198*. ISSN: 1892-2198.

Bouri, Elie et al. (2023). *Geopolitical risk and inflation spillovers across European and North American economies.*

Brignone, Davide, Luca Gambetti, and Martino Ricci (2024). "Geopolitical risk shocks: when the size matters". In: *ECB Working Paper Series* 2972.

Caldara, Dario, Sarah Conlisk, et al. (2023). "Do geopolitical risks raise or lower inflation". In: *Tech. rep., Federal Reserve Board*, pp. 1–32.

Caldara, Dario and Matteo Iacoviello (2022). "Measuring geopolitical risk". In: *American Economic Review* 112.4, pp. 1194–1225. ISSN: 0002-8282.

CBOE Global Markets (2023). *CBOE Volatility Index (VIX)*. https://www.cboe.com/tradable_products/vix/. Accessed: 2025-03-01.

Cieslak, Anna and Andreas Schrimpf (2019). "Non-monetary news in central bank communication". In: *Journal of International Economics* 118, pp. 293–315. ISSN: 0022-1996. DOI: 10.1016/j.jinteco.2019.01.012. URL: https://www.sciencedirect.com/science/article/pii/S002219961830268X.

Coibion, Olivier et al. (2025). *Worrying about war: geopolitical risks weigh on consumer sentiment.* URL: https://www.ecb.europa.eu/press/blog/date/2025/html/ecb.blog20250407~7023432957.en.html.

Corsello, Francesco, Matteo Gomellini, and Dario Pellegrino (2023). "Inflation and energy price shocks: lessons from the 1970s". In: *Bank of Italy Occasional Paper* 790.

Davis, Scott (2015). *The effect of commodity price shocks on underlying inflation: The role of central bank credibility.* SSRN.

Dieckelmann, Daniel et al. (2024). "Turbulent times: geopolitical risk and its impact on euro area financial stability". In: *Financial Stability Review* 1.

Dräger, Lena, Klaus Gründler, and Niklas Potrafke (2022). "Political shocks and inflation expectations: Evidence from the 2022 Russian invasion of Ukraine". In: *Journal of International Economics* 153, p. 104029. ISSN: 0022-1996.

Drechsel, Thomas et al. (2024). "Commodity shocks with diverse impacts: how can different central banks tailor their policies?" In.

European Parliament (2025). *European Parliament resolution of 11 February 2025 on European Central Bank – annual report 2024 (2024/2054(INI)*. Accessed: 2025-02-27. URL: https://www.europarl.europa.eu/news/en/press-room/20250204IPR26681/meps-ask-the-ecb-to-assess-the-impact-of-global-conflicts-on-price-stability.

Foglia, Matteo, Giulio Palomba, and Marco Tedeschi (2023). "Disentangling the geopolitical risk and its effects on commodities. Evidence from a panel of G8 countries". In: *Resources Policy* 85, p. 104056. ISSN: 0301-4207. DOI: 10.1016/j.resourpol.2023.104056. URL: https://www.sciencedirect.com/science/article/pii/S0301420723007675.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville (2016). *Deep Learning*. http://www.deeplearningbook.org. MIT Press.

Gregorio, José de (2012). "Commodity Prices, Monetary Policy, and Inflation". In: *IMF Economic Review* 60.4, pp. 600–633. ISSN: 2041-417X. DOI: 10.1057/imfer.2012.15.

Güntner, Jochen and Johannes Henßler (2021). *Ease on the cannons, tighten on the trumpets: Geopolitical risk and the transmission of monetary policy shocks*.

Hasselgren, Anton et al. (2025). "Do oil price forecast disagreement of survey of professional forecasters predict crude oil return volatility?" In: *International Journal of Forecasting* 41.1, pp. 141–152. ISSN: 0169-2070. DOI: https://doi.org/10.1016/j.ijforecast.2024.04.005. URL: https://www.sciencedirect.com/science/article/pii/S0169207024000360.

Hayo, Bernd, Ali M. Kutan, and Matthias Neuenkirch (2012). "Communication matters: US monetary policy and commodity price volatility". In: *Economics Letters* 117.1, pp. 247–249. ISSN: 0165-1765. DOI: 10.1016/j.econlet.2012.05.018. URL: https://www.sciencedirect.com/science/article/pii/S0165176512002868.

Hoffmann, Mathias et al. (2025). "A KISS for central bank communication in times of high inflation". In: *Discussion Paper* 12.

Hundal, Amaya (2023). "Forecasting real economic activity and inflation : the role of geopolitical risks and economic policy uncertainty in major economies". PhD thesis.

International Monetary Fund (2023). *Global Financial Stability Report April 2023 : Safeguarding Financial Stability amid High Inflation and Geopolitical Risks*. Global Financial Stability Report. Washington, D.C.: International Monetary Fund. ISBN: 979-84-00-23329-6. DOI: Global.

Ioannou, Demosthenes et al. (2023). "The EU's Open Strategic Autonomy from a Central Banking Perspective. Challenges to the Monetary Policy Landscape from a Changing Geopolitical Environment". In.

Jarociński, Marek and Peter Karadi (2020). "Deconstructing Monetary Policy Surprises—The Role of Information Shocks". In: *American Economic Journal: Macroeconomics* 12.2, pp. 1–43. DOI: 10.1257/mac.20180090. URL: https://www.aeaweb.org/articles?id=10.1257/mac.20180090.

Klein, Mathias, Emanuel Skeppås, and Ulf Söderström (2025). *How is Swedish inflation affected by geopolitical risk?* URL: https://www.riksbank.se/globalassets/media/rapporter/ekonomiska-kommentarer/engelska/2025/how-is-swedish-inflation-affected-by-geopolitical-risk.pdf.

Kook, Lucas et al. (2024). "Estimating Conditional Distributions with Neural Networks Using R Package deeptrafo". In: *Journal of Statistical Software* 111.10, pp. 1–36. DOI: 10.18637/jss.v111.i10. URL: https://www.jstatsoft.org/index.php/jss/article/view/v111i10.

Kyriazis, Nikolaos A., Emmanouil M. L. Economou, and Andreas Stergiou (2023). "Connectedness Among Geopolitical Risk, Inflation, Currency Values, and Exports by TVP-VAR Analysis: A Worldwide Perspective". In: *Peace economics, peace science, and public policy* 29, p. 338. ISSN: 1554-8597. DOI: 10.1515/peps-2023-0026.

Lucas, Robert E. (1976). "Econometric Policy Evaluation: A Critique". In: *Carnegie-Rochester Conference Series on Public Policy* 1, pp. 19–46.

Marangoz, Cumali (2025). "Inflation Spillovers and Geopolitical Risks: Evidence from Euro Area Countries Using TVP-VAR and Quantile Models". In: *Ekonomi Politika ve Finans Araştırmaları Dergisi* 10.1, pp. 140–159. DOI: 10.30784/epfad.1599945.

Mo, Bin, He Nie, and Rongjie Zhao (2024). "Dynamic nonlinear effects of geopolitical risks on commodities: Fresh evidence from quantile methods". In: *Energy* 288, p. 129759. ISSN: 0360-5442. DOI: 10.1016/j.energy.2023.129759. URL: https://www.sciencedirect.com/science/article/pii/S0360544223031535.

Morgan, Sedric K. (2019). “Geopolitical Awareness and Understanding of the Current Monetary Policies: A Quantitative Study”. PhD thesis.

Neri, Stefano et al. (2022). “On the anchoring of inflation expectations in the euro area”. In: *Bank of Italy Occasional Paper* 712.

Nuño, Galo, Philipp Renner, and Simon Scheidegger (2024). “Monetary policy with persistent supply shocks”. In: *CES ifo Working Papers*.

Pinchett, Marco (2024). *Geopolitical Risk and Inflation: The Role of Energy Markets*. London School of Economics and Political Science.

Plakandaras, Vasilios and Theophilos Papadimitriou (2019). “The Effects of Geopolitical Uncertainty in Forecasting Financial Markets: A Machine Learning Approach”. In: *Algorithms* 12.1, p. 1. DOI: 10.3390/a12010001. URL: https://www.proquest.com/scholarly-journals/effects-geopolitical-uncertainty-forecasting/docview/2545919009/se-2?accountid=146310.

Qin, Yun and Zitao Zhang (2024). “Transmission mechanisms of the effects of geopolitical risk on energy returns and volatility”. In: *International Review of Financial Analysis* 95, p. 103363. ISSN: 1057-5219. DOI: 10.1016/j.irfa.2024.103363. URL: https://www.sciencedirect.com/science/article/pii/S1057521924002953.

World Economic Forum (Aug. 2024). *Geopolitics replaces inflation as the top worry for central banks and sovereign wealth funds*. Accessed: 2025-05-02. URL: https://www.weforum.org/stories/2024/08/geopolitics-inflation-central-banks/.

Yang, Tianle et al. (2023). “Geopolitical risks, oil price shocks and inflation: Evidence from a TVP–SV–VAR approach”. In: *Energy Economics* 127, p. 107099. ISSN: 0140-9883. DOI: 10.1016/j.eneco.2023.107099. URL: https://www.sciencedirect.com/science/article/pii/S0140988323005972.

PLEDGE OF HONESTY

On my honour as a student of the Diplomatische Akademie Wien, I submit this work in good faith and pledge that I have neither given nor received unauthorised assistance on it.

Luca Phileas Vetter

A APPENDIX

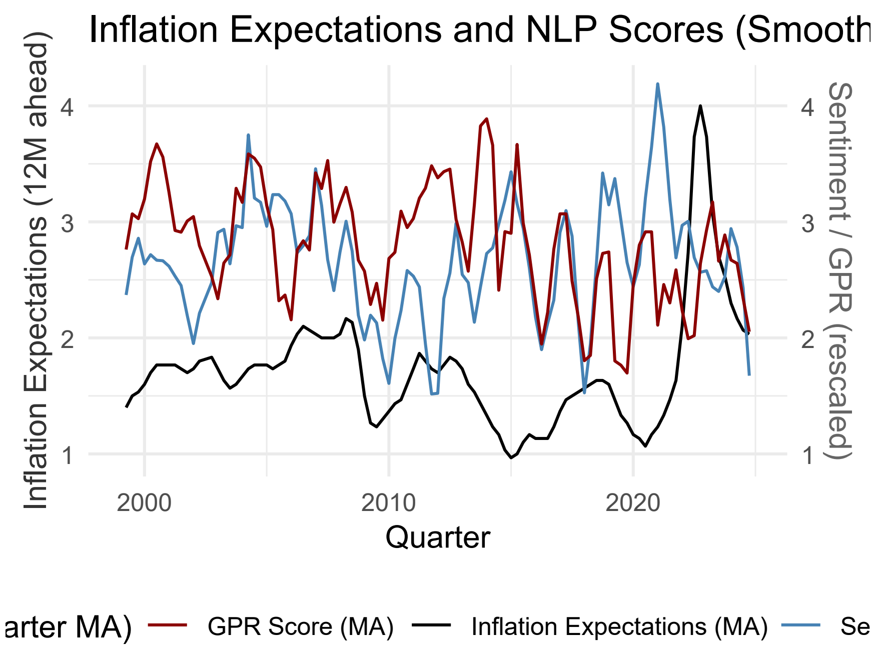


Figure 3: Raw FinBERT-based Sentiment and GPR scores

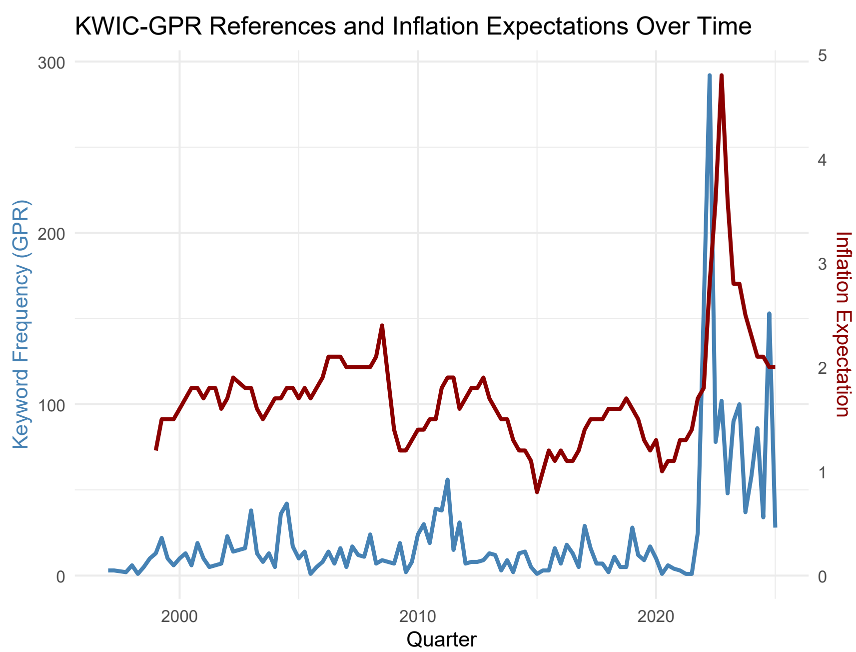


Figure 4: Cleaned KWIC-GPR related scores

Table 2: Comparison of Estimated Shift Coefficients Across DCTM Models

Variable	**DCTM 1**	**DCTM 2**	**DCTM 3**
GPR	−0.765	−0.478	−0.182
EPU	1.544	1.829	2.126
VIX	−0.589	−0.297	−0.001
EMV	−0.223	0.072	0.370
WUI	−0.853	−0.567	−0.270
Sentiment Index	—	—	2.558
Speech GPR Score	—	—	0.175
Text Embedding (101 dims)	—	—	—
Intercept	1.987	1.259	0.532

Table 3: Summary of Deep Conditional Transformation Models (DCTMs)

Model	**Feature Set**	**Textual Input**	**RMSE**	**Log-Likelihood**	**AIC**
DCTM 1	Macroeconomic uncertainty indices	None	2.5997	−1759.709	3531.42
DCTM 2	Macroeconomic indices + text embeddings	Transformer-based embeddings (mean pooled)	2.6422	−126.6371	565.27
DCTM 3	Macroeconomic indices + sentiment scores	FinBERT sentiment + GPR sentiment	1.8925	−148.4969	312.99
OLS	Macroeconomic indices + sentiment scores	FinBERT sentiment + GPR sentiment	0.5261	−78.5669	175.13

Table 4: OLS Regression Results

	Estimate	**Std. Error**	**t value**	**Pr(>|t|)**
(Intercept)	1.583	1.407	1.125	0.263
log_GPR	0.507	0.178	2.851	0.005
log_EPU	0.384	0.156	2.455	0.016
log_WUI	−0.410	0.172	−2.379	0.019
VIX	−0.001	0.014	−0.105	0.917
EMV	0.002	0.016	0.147	0.883
sentiment_index	−0.365	0.793	−0.460	0.646
speech_gpr_score	0.236	0.236	0.999	0.320

Table 5: Quantile Regression Coefficients for Inflation Expectations

	(Intercept)	log_GPR	log_EPU	sentiment_index	speech_gpr_score
25th percentile	0.053	0.365	−0.057	−0.084	0.266
50th percentile	0.901	0.210	−0.040	0.031	0.355
75th percentile	0.395	0.368	−0.027	−0.079	0.169

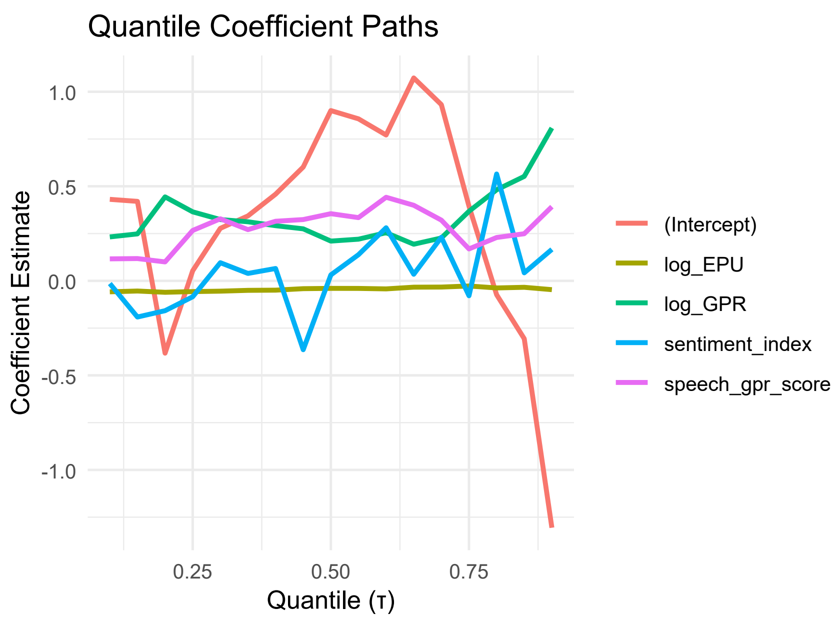


Figure 5: Disaggregated Quantile Coefficient Paths

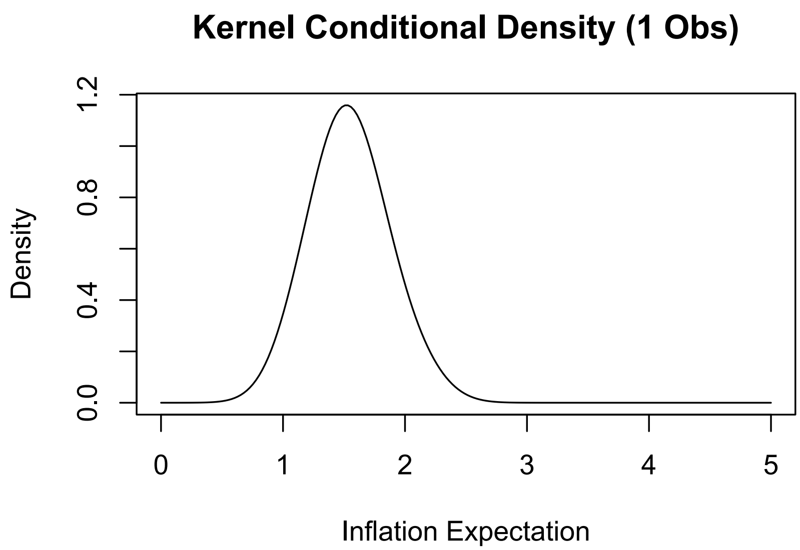


Figure 6: Point-based Kernel Conditional Density

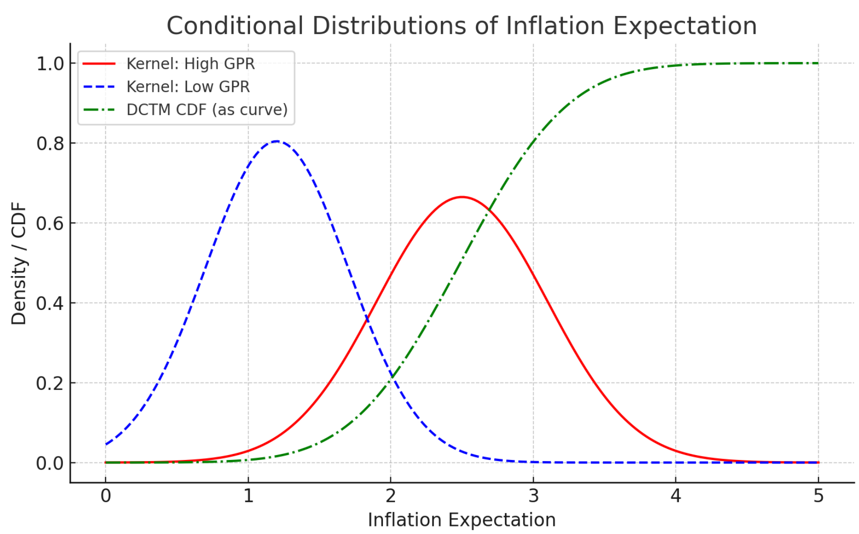


Figure 7: Expected Relationship between Kernel Distribution and CDF under GPR Scenarios

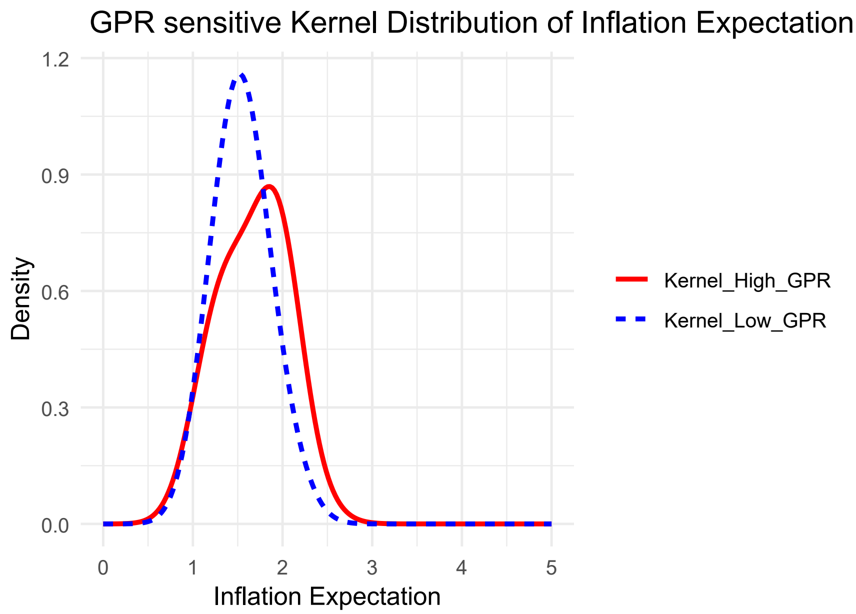


Figure 8: Kernel Distribution under various GPR scenrarios

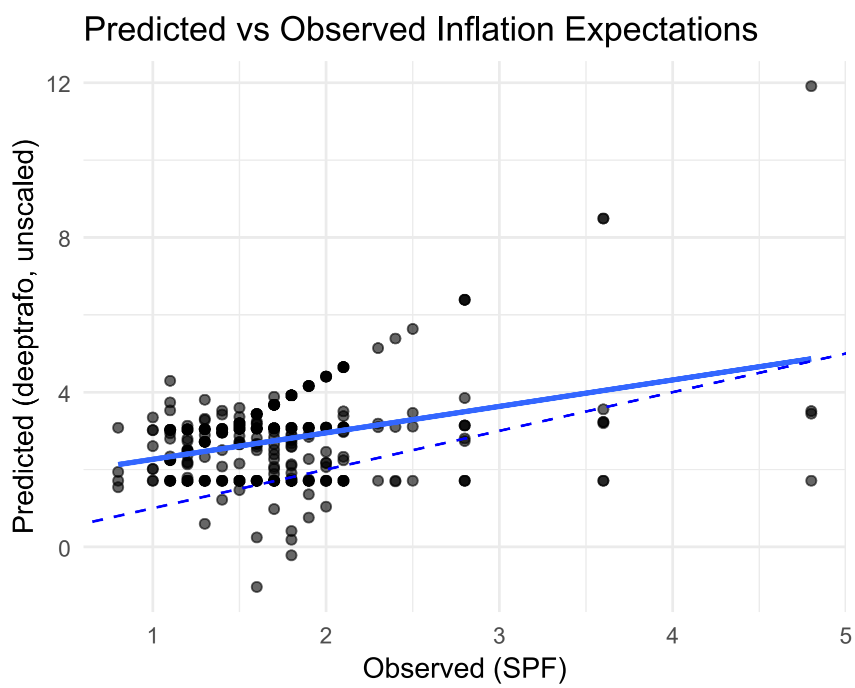


Figure 9: Calibration Plot DCTM 1

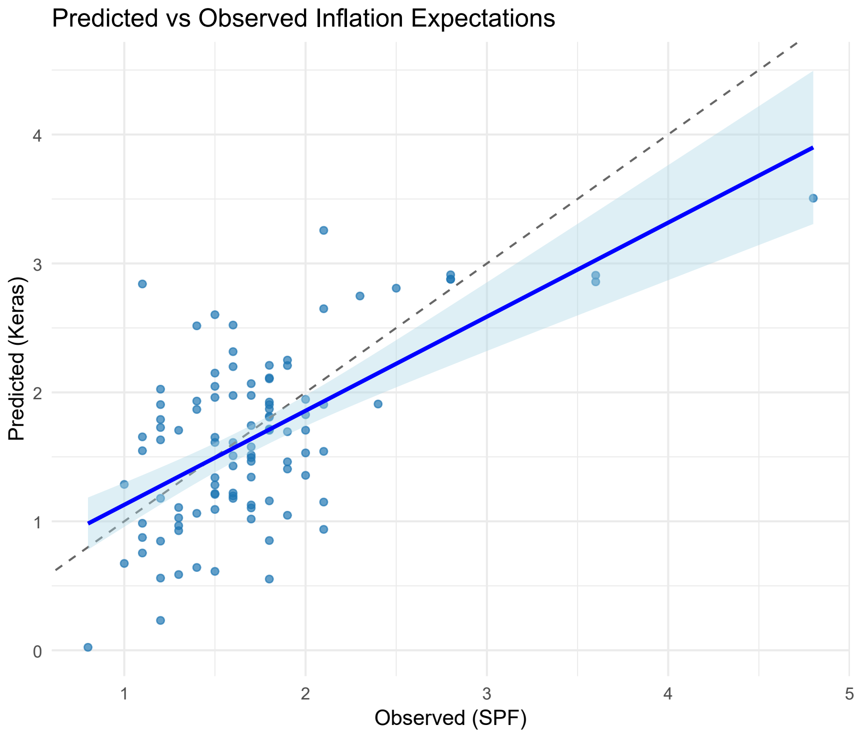


Figure 10: Calibration Plot DCTM 2

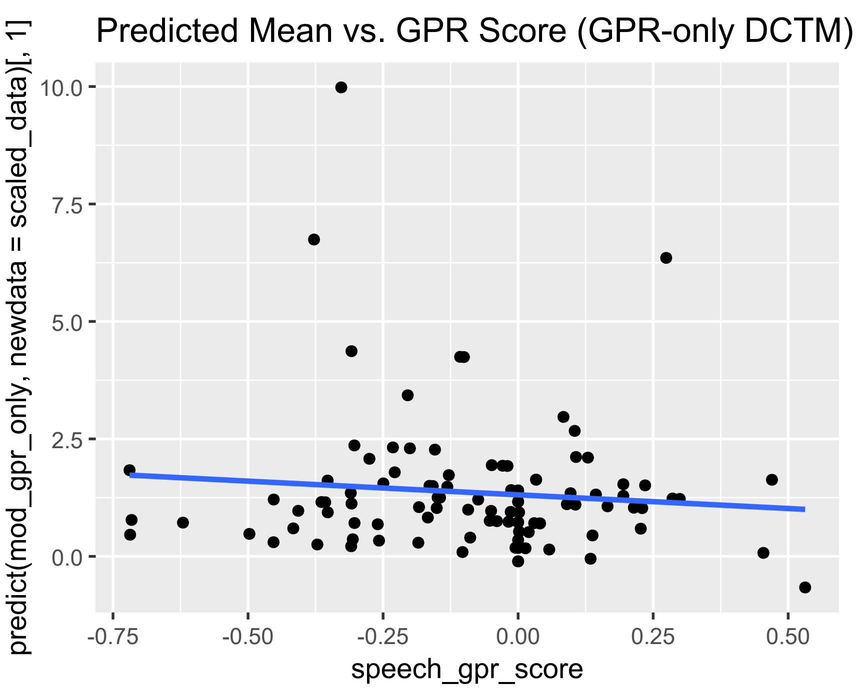


Figure 11: Calibration Plot DCTM 3

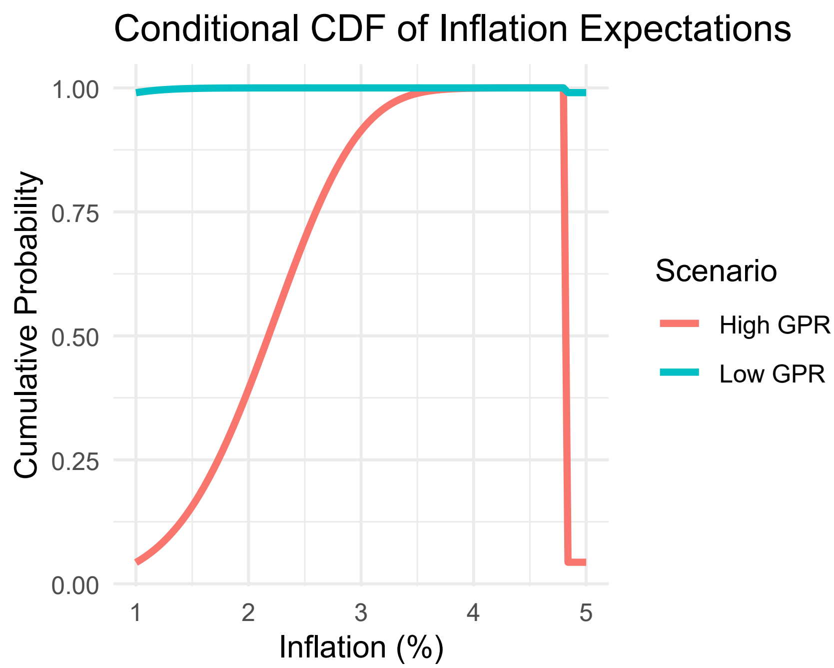


Figure 12: Conditional Distribution Function DCTM 1

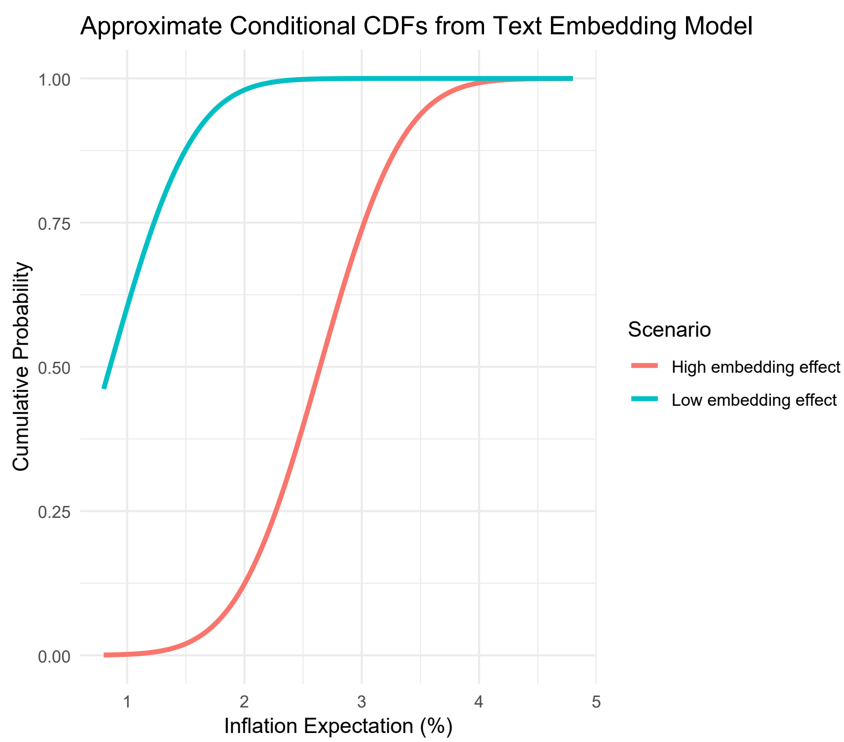


Figure 13: Conditional Distribution Function DCTM 2

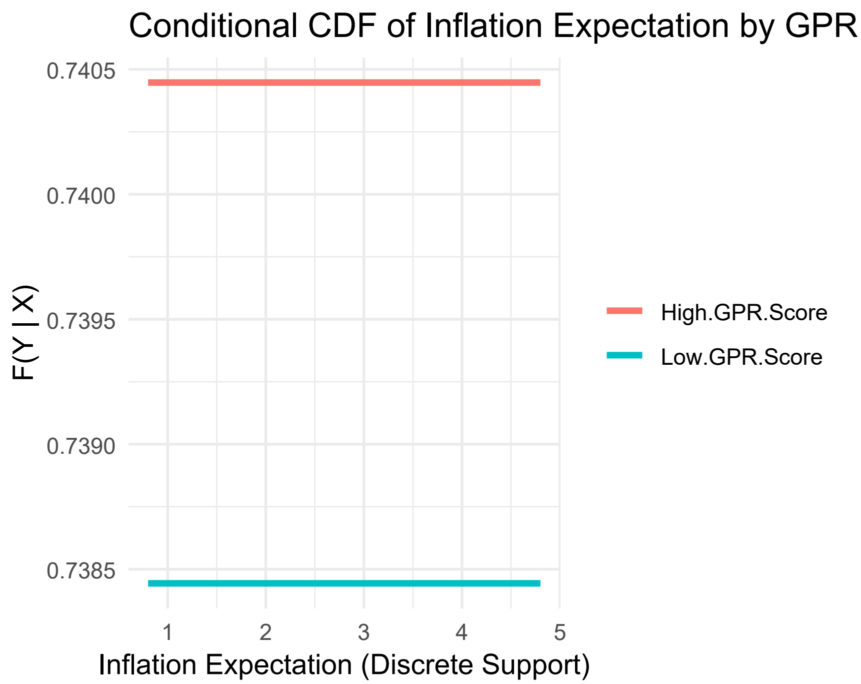


Figure 14: Conditional Distribution Function DCTM 3

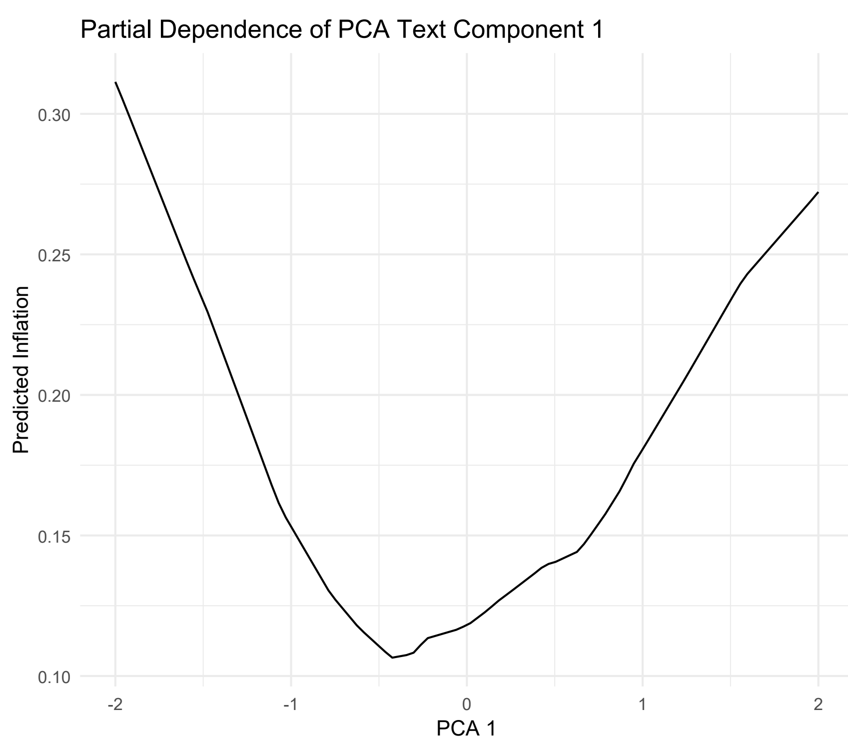


Figure 15: Partial Dependence Plot Text Embeddings DCTM 2

B APPENDIX

To assess the causal relevance of geopolitical sentiment in ECB communications for inflation expectations, a series of counterfactual analyses should be subject of further research. The goal is to simulate how professional forecasters might have formed inflation expectations in the absence of, or under alternative configurations of, ECB sentiment and topic emphasis related to geopolitical risk. While statistical correlation between speech sentiment and SPF inflation expectations is informative, it does not by itself establish whether geopolitical tone has an economically meaningful effect. Counterfactual simulation allows the approximation such effects by isolating the impact of specific features, particularly sentiment polarity and topic prevalence, on the predicted distribution of inflation expectations, holding all other observed features constant. The underlying idea is to estimate a factual prediction based on observed inputs, then simulate hypothetical predictions where selected input variables are modified. The difference between the factual and counterfactual outcomes reveals the marginal impact of the altered features.

The first set of experiments should focus on modifying the geopolitical sentiment score derived from FinBERT or lexicon-based methods. For selected ECB speeches, particularly those delivered during high-risk geopolitical periods such as the 2014 Crimea crisis or the 2022 Ukraine invasion, the sentiment variable needs to be re-coded to represent a neutral tone while keeping all other variables constant. The DCTM is then used to generate new conditional distributions of SPF inflation expectations under the altered input. The difference in median forecasts and tail probabilities between the actual and counterfactual distributions provides an interpretable measure of how much a more neutral or less alarmist tone would have shifted inflation expectations. This procedure repeated in the opposite direction, would simulate a hypothetically more negative tone to assess whether increased alarmism would have elevated expected inflation. These dual counterfactuals allow for symmetry testing: i.e., whether sentiment effects are linear or asymmetric across tone polarity.

In a second series of tests, the prevalence of specific geopolitical topics, which are extracted via Bayesian topic modeling are systematically altered. For example, the model is re-run assuming reduced emphasis on topics relating to "energy security" or "military escalation," while holding sentiment and other covariates fixed. Such manipulations test the marginal influence of issue salience versus tone. In cases where sentiment remains constant but expectations still shift, this would suggest that topic framing alone, independent of emotional polarity, plays a role in shaping forecaster beliefs. In each scenario, Shapley value decompositions or accumulated local effect (ALE) plots can be used to summarize which input features are most responsible for changes in the predicted expectations distribution. These techniques enhance the interpretability of counterfactual simulations in high-dimensional NLP-based models.

Beyond single-speech scenarios, counterfactual simulations could be extended to an aggregated temporal analysis. For each forecast round, an alternative neutral sentiment baseline should be constructed in which the geopolitical sentiment variable is set to zero across all ECB speeches in that period. The DCTM is then used to simulate the hypothetical trajectory of SPF inflation expectations under this neutral communication regime. Comparing the actual and counterfactual time paths allows estimations of the cumulative influence of geopolitical tone over time. These results can be interpreted as an "expectations impulse response" to shifts in central bank sentiment, as a form of quasi-structural analysis that bridges textual economics with forward-looking belief formation.

The counterfactual results provide a practical sense of whether, and to what extent, ECB communication about geopolitical developments shapes inflation expectations. If sentiment and topic modifications induce large changes in the predicted SPF distributions, this suggests that central bank communication is an active channel through which geopolitical risk is priced into expectations. Nonetheless, limitations remain. The simulations rely on the assumption that all non-manipulated variables remain unaffected by the hypothetical change. In reality, a less alarmist speech might differ not only in sentiment but also in length, context, or audience reaction. Hence, the counterfactuals should be interpreted as partial equilibrium responses, not full behavioral general equilibrium predictions.

Counterfactual analysis complements the statistical and calibration-based validations by directly simulating the economic relevance of ECB communication. The results show that variations in geopolitical tone and topic framing can materially shift the predicted distribution of inflation expectations, underscoring the communicative power of the ECB during periods of heightened geopolitical uncertainty. This highlights the importance of central bank rhetoric as a forward-guidance mechanism, even in domains extending beyond monetary policy in the narrow sense.