A Reduced-Form Stochastic Intensity Model of Timing Prediction Markets

ERIK GÖRANSSON-GASPAR

Master's thesis
2026:E34

LUND UNIVERSITY

Faculty of Science
Centre for Mathematical Sciences
Mathematical Statistics

CENTRUM SCIENTIARUM MATHEMATICARUM

Abstract

In prediction markets actors trade contracts whose payout is linked to the occurrence of real-world events. Markets based on questions of the type "Will event A occur before time T?" are particularly interesting in geopolitical contexts, but are understudied in comparison to standard election-style contracts. I present, to the best of my knowledge, the first pricing model for this type of market, based on a reduced-form framework. Suppose that the underlying event arrives with a doubly stochastic Poisson process, driven by a random intensity. The corresponding prediction market price is then related to the process' first-arrival distribution. This allows for tractable pricing equations for a range of reasonable intensity dynamics. I study, in particular, a Cox-Ingersoll-Ross specification with compound Poisson jumps, and propose a Kalman-filter-based quasi-maximum likelihood estimator of the model parameters. Future prices are forecasted using a Monte Carlo procedure. We can also compute the model-implied probability of the underlying event having occurred by any specified date. Such a model enables sophisticated risk management and market making strategies, which have previously been unfeasible for this type of prediction market.

Popular Scientific Summary

In *prediction markets* participants trade financial contracts which pay out a fixed amount—typically one dollar—to the holder if some real-world event occurs. The most traded contracts have historically been bets on which candidate will win a U.S. presidential election. In these markets the price of a contract can be interpreted as an estimate of the probability of the underlying event occurring. For example, if you think that there is a 30 % chance of occurrence, then the most you should rationally be willing to pay for the corresponding one dollar contract is 30 cents. In this thesis I study a less common form of prediction market based on questions of the type "Will event A occur *before* time T?" The probability of the answer being *yes* goes to zero close to the end date T, and with it the contract price. This is qualitatively different from the behaviour of election-style markets, where there is typically a lot of uncertainty left near the end. I call the former type *timing markets*. They are particularly interesting in geopolitical contexts; a series of markets with various end dates have been run, e.g., on the ousting of Ayatollah Khamenei.

I present the first, to the best of my knowledge, mathematical model of timing market prices and of how they evolve. Consider the probability that an underlying event will occur in the next (infinitesimally) small interval of time; this probability is called the *intensity*. By assuming that it changes randomly over time in specific ways I am able to compute the implied probability that the underlying event will occur before the market's end date T. This corresponds to the market price of the associated timing prediction contract. The model I present is able to make forecasts for future market prices. It can also be used to estimate the probability that the event will have occurred by any specified date, not just the end date T. By studying how good the model's predictions are on real timing markets, I conclude that although the model as presented is flawed, the proposed modeling framework is promising. A good model of this type, towards the development of which I have made a first contribution, could enable traders in timing prediction markets to employ sophisticated risk management strategies that have previously been infeasible. This is necessary if they are to approach the maturity of traditional financial markets.

Contents

1 Introduction

In a typical prediction market participants trade contracts that pay out one dollar if some real-world event occurs and nothing otherwise. The market's equilibrium price reflects their average belief that the event will occur. We interpret a price of 30 cents as a market-implied 30 % probability. This ability to produce aggregate probability estimates, which have been shown to be generally accurate, makes prediction markets interesting objects to study.

Prediction contracts are traded on online platforms which host a large number of markets based on everything from geopolitics and economic indicators to the outcome of football matches. The current leading platforms, Polymarket and Kalshi, rose to mainstream prominence during the 2024 U.S. presidential election, when prediction market prices were frequently cited by the media. Polymarket hosted the largest prediction market in history on the outcome of the election, with $3.7 billion traded; the price history of this market is shown in Figure 1. (Polymarket, 2024) In 2025 Polymarket and Kalshi had a combined trading volume of $49 billion. (HTX Insights, 2026)

Figure 1: Daily price of the contract "Will Trump Win the 2024 U.S. Presidential Election?" on Polymarket, from January to November 2024. The market ended on election day, November 5. Data was retrieved using the Polymarket API.

The first modern prediction market was launched by researchers at the University of Iowa in advance of the 1988 U.S. presidential election. The implied predictions outperformed traditional polls, which stimulated academic research into these markets. In the following decades a large literature emerged, studying their theoretical and empirical properties. Substantial work has also been done on how to design prediction markets to best aggregate information. Less attention has been paid to the mathematical modeling of prediction market prices, which is my concern in this thesis. Section 2 below contains an overview of the existing literature.

Most prediction markets are based on questions of the type "Will event A occur *at* time T?", e.g. will a given candidate win on the day of the election. Such events generally have multiple mutually exclusive outcomes which are all tradeable and whose respective prices add to one. Typically, the true outcome is not certain even right before the market's end time T, resulting in prices removed from zero or one until market resolution. In geopolitical contexts, however, it is often more interesting to consider questions like "Will event A occur *before* time T?" For example, a series of markets were held on the ousting of Ayatollah Khamenei as the supreme leader of Iran, up until his killing on February 28 this year. One of these price histories is shown in Figure 2. Such markets are concerned with the timing of events; I therefore call them *timing markets* to distinguish them from the more common *outcome markets.* This terminology is mine. In timing markets the price goes to zero *ceteris paribus*, as long as the underlying event does not occur, because there is less and less time left for something unexpected to happen. This convergence behaviour is a consequence of the payout structure of timing contracts and is therefore not observed in outcome markets. Consequently, we should not use the same pricing models for both types of markets.

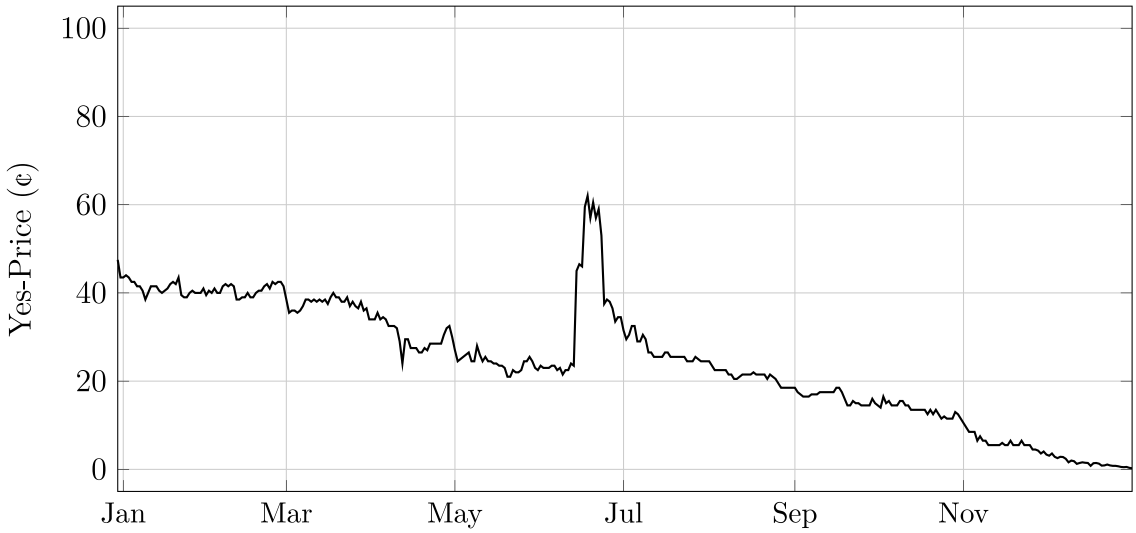


Figure 2: Daily price of the contract “Khamenei Out as Supreme Leader of Iran in 2025?” on Polymarket, from January to the end of December 2025. Data was retrieved using the Polymarket API.

Timing prediction markets have been previously studied — Leigh, Wolfers, and Zitzewitz (2003) and Wörle (2013) look at markets on the ousting of Saddam Hussein and Muammar Gaddafi, respectively — but the distinction between them and outcome markets has not been clearly articulated. I present here the first time series model of timing market prices, to the best of my knowledge. In Section 3 I argue that it is natural to treat such market prices as survival probabilities of a doubly stochastic Poisson process. This is the same problem as the one treated in the literature on reduced-form models of defaultable corporate bonds. Drawing on this body of work, I propose a pricing model based on a latent jump-diffusion Cox-Ingersoll-Ross process. Section 4 describes a Kalman filter-based quasi-maximum likelihood estimation procedure for my proposed model, while Section 5 explains how to use Monte Carlo simulations to produce price forecasts. In Section 6 I evaluate the model's predictive performance on real timing market data. Section 7 concludes.

2 An Overview of Prediction Markets

Over the last half-century, the study of prediction markets has produced a significant body of research. This section contains a brief overview of this literature, in order to set the scene for the modeling work that constitutes the original contribution of this thesis. The interested reader can find a more thorough introduction in Wolfers and Zitzewitz (2006b). Comprehensive literature reviews have been published by Tziralis and Tatsiopoulos (2007), Horn et al. (2014), and Forestal, Zhang, and Pi (2020).

Betting through financial exchanges is an old phenomenon. Rhode and Strumpf (2004) report that prominent prediction markets on U.S. presidential elections existed between 1868 and 1940. At their peak, ca. $300 million (2026 equivalent) were wagered on the election of 1916. These early markets fell out of favor in the 1940s with the advent of modern political polling, but were to make a return decades later. The contemporary interest in prediction markets was spurred by the *Iowa Electronic Market*, which was launched in advance of the 1988 presidential election. It allowed faculty and students at the University of Iowa to trade contracts tied to various election outcomes. Berg et al. (2008) summarize the numerous studies made of its various iterations. They note that these markets matched or exceeded the accuracy of opinion polls, both weeks before the election and on election eve. Market prices quickly incorporated new information, without simply tailing published polls. The authors also cite evidence for both arbitrage-free pricing between related markets and weak-form efficiency. Early experiments were also carried out elsewhere. Beckmann and Werding (1996) observed similar accuracy on the *Passauer Wahlbörse* market platform hosted for the German elections of 1994. A prediction market organized in advance of the Swedish referendum on joining the European Union the same year also outperformed traditional polling. Bohm and Sonnegard (1999) attribute this success in particular to the large number of respondents telling pollsters that they were undecided. These early markets are characterized by relatively low trading volumes and few traders, mostly consisting of students and researchers. In the context of these limitations, their predictive accuracy is remarkable.

Around the same time as the first experimental prediction markets were being held, the economist Robin Hanson published a series of influential speculative articles. (Hanson, 1990, 1992, 1995) He envisioned prediction markets replacing or supplementing a large number of social functions, from the allocation of research funding to informing long-term policy making. Hanson would later publish important work on the design of automated market makers for prediction markets, introducing the so-called *constant function market maker.* (Hanson, 2007, 2009) This work helped found a line of research which still sees a lot of activity. Frongillo, Papireddygari, and Waggoner (2024), Nueve and Waggoner (2026), and Gao et al. (2025) are examples of recent publications in this tradition.

The successes of early prediction market experiments attracted significant interest and, with it, theoretical scrutiny. Manski (2006) posed a famous challenge to the interpretation of market prices as probability estimates by presenting a market model in which mean trader belief diverged significantly from the equilibrium price. In response, Wolfers and Zitzewitz (2006a) and Gjerstad and Hall (2005) conducted detailed studies of such models. They were able to show that the interpretability of market prices is sensitive to assumptions about traders' risk aversion and the distribution of their beliefs; they established mild conditions under which the conventional interpretation holds either exactly or approximately.

Only two papers have been published, as far as I am aware, which study timing markets in particular, although they do not explicitly distinguish them from outcome markets. These are of special interest to this thesis. Leigh, Wolfers, and Zitzewitz (2003) describe markets held in 2002 and 2003 on the ousting of Saddam Hussein as President of Iraq, which they take as a proxy for the probability of a U.S. invasion. They claim to show that changes to this probability explain a large portion of the variation observed in S&P500 and oil futures prices during the period in question. Wörle (2013) performs a similar analysis on a prediction market related to the removal of Muammar Gaddafi as the leader of Libya.

There has also been comparatively little interest in modeling the evolution of prediction market prices over time. The first such model was presented by Chen, Ingersoll, and Kaplan (2008). They study a set of prediction markets related to the 2004 U.S. presidential election. Their key contribution is to model a latent sentiment process for each state as a random walk, which determines which candidate that state votes for. This is, in effect, treating the prediction market contracts as derivatives of an unobservable stochastic process. Within this framework Chen, Ingersoll, and Kaplan are able to derive pricing formulas for a contract on the national election winner, based on state-level market prices. They show that actual prices agree well with the theoretical relationship; this serves as a test of inter-market pricing efficiency. Archak and Ipeirotis (2009) extend this model to allow for a more general specification of the latent process as a geometric Brownian motion. In the appendix to a working version of their paper they also generalize this model to prediction markets with more than two mutually exclusive outcomes. (Archak & Ipeirotis, 2008) A recent paper by Dalen (2025) takes a more direct approach, by modeling market prices as the log-odds transform of a jump-diffusion geometric Brownian motion.

The model presented in this thesis follows in the tradition of Chen, Ingersoll, and Kaplan by modeling contracts as derivatives of a latent process. In the context of timing markets, a natural choice is to suppose that the underlying event arrives with a doubly stochastic Poisson process and model its random intensity.

3 Modeling a Timing Market

3.1 Reduced-Form Framework

Consider a timing market based on the question "Will event A arrive before time T?". It is natural to model the market price as the survival probability of a Poisson process in the following way. Suppose that the underlying event A arrives with a doubly stochastic Poisson process N_t with intensity λ_t, which is adapted to a filtration $\mathcal{F}_t$. Let π_t^{no} be the *no-price* of the market, i.e. the price of a contract which pays out one dollar if the answer to the question "Will the event A occur before time T?" turns out to be no; we call the price of a contract with the regular payoff structure the *yes-price*. Both yes- and no-contracts are typically tradeable and their prices add to one.

Let R be the first-arrival time of N_t, i.e. $R := \inf\{t : N_t > 0\}$. The price of a no-contract is zero after arrival; before that it is equal to the conditional probability of first-arrival occurring after the end-date T. Neglecting discounting and other price distortions, the no-price compatible with interpreting market prices as probabilities is then

$$\pi_t^{\text{no}} = \mathbb{P}\left[R > T | \mathcal{F}_t\right].$$

This is the survival probability of N_t by the end date T, conditional on the market not having resolved by time t. The unconditional price process is trivially a martingale, because for any $s > t$

$$\mathbb{E}\left[\pi_s^{\text{no}} | \mathcal{F}_t\right] = \mathbb{E}\left[\mathbb{E}\left[R > T | \mathcal{F}_s\right] | \mathcal{F}_t\right] = \mathbb{E}\left[R > T | \mathcal{F}_t\right] = \pi_t^{\text{no}} \quad \text{a.s.}$$

The standard properties of doubly stochastic Poisson processes give a pricing equation

$$\pi_t^{\text{no}} = \mathbb{E}\left[e^{-\int_t^T \lambda_s \mathrm{d}s} \,\middle|\, \mathcal{F}_t\right]. \tag{$\star$}$$

Proof. Since the market has not resolved by time t, we have that $N_t = 0$ and thus $N_T = N_T - N_t$. Denote the cumulative intensity $\Lambda(t,T) := \int_t^T \lambda_s \,\mathrm{d}s$. The fair no-price is the conditional probability that the market has not resolved by the end time T, i.e. that $N_T = 0$. By the tower property

$$\pi_t^{\text{no}} = \mathbb{E}\left[\mathbb{P}[N_T - N_t = 0 \mid \Lambda(t,T)] \,\middle|\, \mathcal{F}_t\right].$$

Since N_t is a doubly stochastic Poisson process, the conditional increment $N_T - N_t | \Lambda(t,T)$ is Poisson-distributed with mean $\Lambda(t,T)$. Therefore

$$\pi_t^{\text{no}} = \mathbb{E}\left[e^{-\int_t^T \lambda_s \mathrm{d}s} \,\middle|\, \mathcal{F}_t\right],$$

which completes the proof. □

A reader familiar with the theory of bond pricing will recognize the above equation. A no-contract has the same payoff structure as a defaultable bond of unit principal with zero yield or recovery. The so called *reduced-form* family of models of such bonds assume that default arrives with a doubly stochastic Poisson process, giving rise to Eq. ($\star$). Bakshi, Gao, and Zhong (2022) give an introduction to this literature, which inspired the present work. It is of even greater significance that Eq. ($\star$) gives the price of a zero-coupon bond if λ_t is taken to be a stochastic short rate. This analogy allows us to exploit the theoretical developments made in the vast literature on classical bond pricing. In Section 3.3 we draw on this work to get an analytically tractable model of timing market prices with a jump-diffusion intensity process.

3.2 Constant Intensity

The no-price of a timing market at time t is entirely determined by the current and future values of the intensity process. In the simplest possible model we assume that the future intensity is constant, i.e. $\lambda_s \equiv \lambda > 0$ for $s \geq t$. Eq. ($\star$) then reduces to

$$\pi_t^{\text{no}} = e^{-(T-t)\lambda},$$

which allows us to compute future prices π_s^{no} for $s \geq t$.

There are multiple ways one could estimate the constant intensity λ given a price history $\{\pi_k^{\text{no}}\}_{k=1}^n$ observed at times $t_1, t_2, \ldots, t_n$. If we only consider the last observed price π_n^{no}, then we can invert the pricing equation to find the corresponding constant intensity. We get

$$\hat{\lambda} = -\frac{\log \pi_n^{\text{no}}}{T - t_n}.$$

Let us call this estimator the *instantaneous intensity.* Another natural choice is to use the entire observed price history, assuming that the intensity is constant for all times, past and future. The least squares estimator of what we can call the *homogeneous intensity* is then

$$\hat{\lambda} = \arg\min_{\lambda>0} \sum_{k=1}^{n} \left(\pi_k^{\text{no}} - e^{-(T-t_k)\lambda}\right)^2.$$

Both of these estimators can be derived as special cases of a weighted least squares estimator $\hat{\lambda}_\gamma$ with exponential decay factor $\gamma \in [0,1]$, where

$$\hat{\lambda}_\gamma = \arg\min_{\lambda>0} \sum_{k=1}^{n} \gamma^{n-k} \left(\pi_k^{\text{no}} - e^{-(T-t_k)\lambda}\right)^2.$$

It may be more intuitive to reparameterize this estimator in terms of an effective rectangular window size $W = 1/(1-\gamma)$. The instantaneous estimator is recovered by choosing $\gamma = 0, W = 1$ (with the convention that $0^0 = 1$) and the homogeneous by $\gamma = 1, W = \infty$.

In order to compute the weighted least-squares estimator, we derive a standard Newton-Raphson algorithm. Since we are minimizing a smooth objective, differentiating yields a score equation which has the estimator as a zero:

$$S(\lambda) := \sum_{k=1}^{n} \gamma^{n-k} (T - t_k) e^{-\lambda(T-t_k)} \left(\pi_{t_k}^{\text{no}} - e^{-\lambda(T-t_k)}\right).$$

By taking the derivative of the score

$$S'(\lambda) = \sum_{k=1}^{n} \gamma^{n-k} (T - t_k)^2 e^{-\lambda(T-t_k)} \left(2e^{-\lambda(T-t_k)} - \pi_{t_k}^{\text{no}}\right)$$

we can construct a Newton step

$$\lambda_{i+1} = \lambda_i - \frac{S(\lambda_i)}{S'(\lambda_i)},$$

starting at any $\lambda_0 > 0$. Iterating until convergence yields the exponentially weighted least-squares estimator $\hat{\lambda}_\gamma$. In Section 6 we will use the predictions of these constant-intensity models as baselines to compare against a more sophisticated jump-diffusion model.

3.3 Jump-Diffusion Intensity

Within the reduced-form pricing framework, the key modeling decision is the choice of dynamics of the intensity process. They should guarantee that the intensity is non-negative and be consistent with observed prices. Drawing on the defaultable bond pricing literature, it is natural to consider an intensity process λ_t with Cox-Ingersoll-Ross (CIR) dynamics

$$\mathrm{d}\lambda_t = \kappa(\mu - \lambda_t)\mathrm{d}t + \sigma\sqrt{\lambda_t}\mathrm{d}W_t, \quad (\mu, \kappa, \sigma > 0)$$

where W_t is a standard Wiener process. Such a process is mean-reverting around a level μ with strength determined by κ; σ is a volatility parameter. An analytic solution to Eq. ($\star$) with CIR intensity was presented by Cox, Ingersoll, and Ross (1985) in order to price zero-coupon bonds. The CIR process is non-negative and strictly positive if the so called Feller condition $2\kappa\mu > \sigma^2$ is satisfied; this was first proven by Feller (1951). The process' transition probability density is a scaled non-central χ^2-distribution such that for $s > t$

$$2C\lambda_s|\lambda_t \sim \text{non-central}\,\chi^2(2q+2, 2u),$$

where

$$C := \frac{2\kappa}{\sigma^2(1 - e^{-\kappa(s-t)})}, \quad u := C\lambda_t e^{-\kappa(s-t)}, \quad q := \frac{2\kappa\mu}{\sigma^2} - 1.$$

Prediction market prices often exhibit jumps, corresponding to the revelation of new information. A clear example of this can be seen in Figure 2, which shows the Polymarket price of a timing contract based on the question "Khamenei Out as Supreme Leader of Iran in 2025?" The spike in yes-prices coincides with the U.S.-Israeli bombings carried out against Iran in June 2025. To model such behaviour, we take the intensity to be a jump-diffusion process. We will consider a jump-CIR-process (JCIR) with compound Poisson jumps. It has dynamics

$$\mathrm{d}\lambda_t = \kappa(\mu - \lambda_t)\mathrm{d}t + \sigma\sqrt{\lambda_t}\mathrm{d}W_t + \mathrm{d}J_t,$$

where J_t is a compound Poisson process. It is defined by

$$J_t := \sum_{i=1}^{M(t)} Y_i,$$

where the number of jumps is determined by a Poisson process $M(t)$ with constant intensity $\alpha > 0$. The jump sizes Y_i are i.i.d. exponentially distributed with mean $\gamma > 0$. Under this model, jumps are strictly positive. This is clearly an unrealistic assumption for prediction market prices, but is made in order to preserve tractability. Downward jumps are difficult to model, because in order to maintain the non-negativity of the intensity process, the jump sizes would have to depend on the current intensity level.

Since the JCIR process is affine, we can derive an explicit solution to the pricing equation using the standard methods of affine options pricing. Following Brigo and Mercurio (2006) we have

$$\pi_t^{\text{no}} = A(t)e^{-B(t)\lambda_t},$$

where

$$\begin{aligned}
h &:= \sqrt{\kappa^2 + 2\sigma^2}, \\
A(t) &:= \left[\frac{2he^{(\kappa+h)(T-t)/2}}{(\kappa+h)\left(e^{h(T-t)}-1\right)+2h}\right]^{2\kappa\mu/\sigma^2} \left[\frac{2he^{(h+\kappa+2\gamma)(T-t)/2}}{2h+(\kappa+h+2\gamma)(e^{h(T-t)}-1)}\right]^{\frac{2\alpha\gamma}{\sigma^2-2\kappa\gamma-2\gamma^2}}, \\
B(t) &:= \frac{2\left(e^{h(T-t)}-1\right)}{(\kappa+h)\left(e^{h(T-t)}-1\right)+2h}.
\end{aligned}$$

Note that the denominator in the second exponent of $A(t)$

$$\sigma^2 - 2\kappa\gamma - 2\gamma^2 = \frac{h^2 - (\kappa+2\gamma)^2}{2}$$

is equal to zero when $h^2 = (\kappa + 2\gamma)^2$. In this limit we get the alternate expression

$$A(t) := \left[\frac{2he^{(\kappa+h)(T-t)/2}}{(\kappa+h)\left(e^{h(T-t)}-1\right)+2h}\right]^{2\kappa\mu/\sigma^2} \exp\left(-2\alpha\gamma\left[\frac{T-t}{\kappa+h+2\gamma} + \frac{e^{-h(T-t)}-1}{h(\kappa+h+2\gamma)}\right]\right).$$

Under this model, if we know the parameters $(\mu, \kappa, \sigma, \alpha, \gamma)$ and the current intensity λ_t, then we can compute the implied no-price π_t^{no}. This also allows us to simulate the price dynamics following a method proposed by Mikulevičius and Platen (1988). Assuming that the Brownian motion and Poisson process driving the JCIR process are independent, they suggest simulating the diffusion and jumps separately. We first draw the number of jumps on the time interval $[t, T]$ as a $\text{Po}(\alpha[T-t])$-distributed random variable. The jump times are then uniformly distributed on $[t, T]$, while the jump sizes are independently drawn from an exponential distribution with mean γ. Once the jumps have been obtained, a CIR-process can be simulated on a discrete grid using the exact non-central χ^2 transition density, adding jumps as appropriate. This will give us an intensity series simulated conditionally on the market not resolving, which we can use to compute the corresponding prices. To account for the possibility of the market resolving before its end date, we must check for resolution at each step of the simulation. At time t_k with current intensity λ_k the probability of the market resolving before the next time step t_{k+1} is

$$1 - \mathbb{E}\left[e^{-\int_{t_k}^{t_{k+1}} \lambda_s \,\mathrm{d}s} | \mathcal{F}_t\right] \approx 1 - e^{-\Delta\lambda_k},$$

for small grid sizes $\Delta := t_{k+1} - t_k$. The remainder of this thesis is dedicated to implementing and evaluating the doubly stochastic model of timing prediction market prices with JCIR intensity. In order to make predictions we first need to be able to fit the model's parameters to an observed price history. In the next section I present a Kalman-filter-based quasi-maximum likelihood procedure for parameter estimation.

4 Estimation Procedure

4.1 Kalman Quasi-Maximum Likelihood Estimator

I have above presented a pricing model for timing market prices driven by a JCIR intensity process. In this section I develop a method for estimating its parameters from historical price series. A standard CIR process has a closed form non-central χ^2 transition density, which allows us to also derive a transition density for the price process through a standard change of variables. This would allow us to use ordinary maximum likelihood estimation to find its parameters. The JCIR process, however, lacks such closed form expressions and we must resort to a more elaborate estimation scheme. Keeping the notation of the previous section, we have a model consisting of one stochastic differential equation and one deterministic pricing equation:

$$\begin{aligned} \mathrm{d}\lambda_t &= \kappa(\mu - \lambda_t)\mathrm{d}t + \sigma\sqrt{\lambda_t}\mathrm{d}W_t + \mathrm{d}J_t, \\ \pi_t^{\text{no}} &= A(t)e^{-B(t)\lambda_t}. \end{aligned}$$

We will approximate this system by a quasi-linear Gaussian state space model. The quasi-likelihood of Kalman filter innovation residuals will then give us a suitable objective to maximize. Let us first discretize the state equation on a time grid $\{t_k\}_{k=1}^n$. Denote the corresponding price and intensity values by $\{\pi_k^{\text{no}}\}_{k=1}^n$ and $\{\lambda_k\}_{k=1}^n$, respectively. An Euler-Maruyama scheme with constant time step Δ yields

$$\lambda_{k+1} = \lambda_k + \kappa(\mu - \lambda_k)\Delta + \sigma\sqrt{\Delta\lambda_k}Z_k + J_k',$$

where Z_k are i.i.d. standard normal random variables and J_k' are increments of a compound Poisson process $M(t)$ such that

$$J_k' = \sum_{i=M(t_k)+1}^{M(t_{k+1})} Y_i.$$

Let $m_k := M(t_{k+1}) - M(t_k)$. We can then compute the first two moments of J_k' by the laws of total expectation and variance:

$$\begin{aligned} \mathbb{E}\left[J_k'\right] &= \mathbb{E}\left[\mathbb{E}\left[J_k'|m_k\right]\right] = \mathbb{E}\left[m_k\gamma\right] = \Delta\alpha\gamma, \\ \mathrm{Var}[J_k'] &= \mathbb{E}\left[\mathrm{Var}[J_k'|m_k]\right] + \mathrm{Var}\left[\mathbb{E}\left[J_k'|m_k\right]\right] = \mathbb{E}\left[m_k\gamma^2\right] + \mathrm{Var}[m_k\gamma] = 2\Delta\alpha\gamma^2. \end{aligned}$$

If we approximate J_k' by a Gaussian matching these moments we get a total state equation

$$\lambda_{k+1} = \lambda_k(1 - \kappa\Delta) + \Delta\kappa\mu + \Delta\alpha\gamma + \eta_k,$$

where $\eta_k|\lambda_k \sim \mathcal{N}\left(0, \sigma^2\Delta\lambda_k + 2\Delta\alpha\gamma^2\right)$. We also wish to transform the observation equation into a linear Gaussian form. Consider its logarithm

$$\log\pi_k^{\text{no}} = \log A(t_k) - B(t_k)\lambda_k \iff y_k := \log A(t_k) - \log\pi_k^{\text{no}} = B(t_k)\lambda_k.$$

Suppose that the shifted log-price y_k is observed with Gaussian error having zero mean and variance s^2. Although we make this assumption as a concession to the estimation methodology, this observation error could help account for small discrepancies between modeled and observed prices arising due to a bid-ask spread or other market inefficiencies. By a slight redefinition of terms we then have

$$y_k = B(t_k)\lambda_k + \varepsilon_k, \quad \varepsilon_k \sim \mathcal{N}(0, s^2) \text{ i.i.d.}$$

This is equivalent to log-normal multiplicative noise in the untransformed pricing equation. In summary, we have transformed our original model into state space form

$$\begin{aligned}
\lambda_{k+1} &= \lambda_k(1-\kappa\Delta) + \Delta\kappa\mu + \Delta\alpha\gamma + \eta_k, \\
y_k &= B(t_k)\lambda_k + \varepsilon_k.
\end{aligned}$$

Note that this is not a standard linear Gaussian state-space system, since the variance of η_k depends on the current level λ_k. When implementing the Kalman filter, we can circumvent this by simply using the filtered estimate $\mathbb{E}\left[\lambda_k|Y_k\right]$ as a plug-in estimator when computing the state variance. Here Y_k denotes the observations $y_k, y_{k-1}, \ldots$ available up to and including time t_k. The filter's prediction equations then become

$$\begin{aligned}
\mathbb{E}\left[\lambda_{k+1}|Y_k\right] &= \mathbb{E}\left[\lambda_k|Y_k\right](1-\kappa\Delta) + \Delta\kappa\mu + \Delta\alpha\gamma, \\
\operatorname{Var}\left[\lambda_{k+1}|Y_k\right] &= (1-\kappa\Delta)^2 \operatorname{Var}\left[\lambda_k|Y_k\right] + \sigma^2\Delta\,\mathbb{E}\left[\lambda_k|Y_k\right] + 2\Delta\alpha\gamma^2.
\end{aligned}$$

This filter then produces innovation residuals

$$e_k := y_k - B(t_k)\mathbb{E}\left[\lambda_k|Y_{k-1}\right] \sim \mathcal{N}\left(0, B(t_k)^2\operatorname{Var}[\lambda_k|Y_{k-1}] + s^2\right).$$

Let ℓ_{KF} be the joint quasi-log-likelihood of the residuals

$$\ell_{\mathrm{KF}} = -\frac{1}{2}\sum_{k=1}^{n-1}\left[\log(2\pi) + \log\left(B(t_k)^2\operatorname{Var}[\lambda_k|Y_{k-1}] + s^2\right) + \frac{e_k^2}{B(t_k)^2\operatorname{Var}[\lambda_k|Y_{k-1}] + s^2}\right].$$

The above analysis is done under the assumption that the market has not resolved. Recall that the probability of the market not resolving in the interval $[t_k, t_{k+1}]$ is approximately $e^{-\lambda_k\Delta}$. To account for this, we add upp the log-probability

$$\ell_{\mathrm{Res}} := \sum_{k=1}^{n-1} -\Delta\,\mathbb{E}\left[\lambda_k|Y_k\right],$$

accounting for the fact that we are only able to observe the price series $\{\pi_k^{\mathrm{no}}\}_{k=1}^n$ because the market did not resolve before time t_k. This has the effect of penalizing high intensity levels which would make it unlikely that the market would remain unresolved long enough for us to observe the entire price history. Note that this is an approximate correction. Parameter estimates can finally be found by maximizing the total quasi-log likelihood

$$\ell_{\mathrm{QML}} := \ell_{\mathrm{KF}} + \ell_{\mathrm{Res}}.$$

The observation variance s^2 can either be estimated jointly with the rest of the parameters, or fixed to some reasonable value. A sensible starting point would be to let the standard deviation be equal to half of a typical bid-ask spread of 0.02 dollars, implying that $s^2 = 10^{-4}$. A benefit of the Kalman filter-based estimation procedure is that it is easy to jointly fit a model to data from multiple markets on the same underlying event. All that would need to be done is to add more observation equations to the Kalman filter, with different end dates but the same model parameters.

4.2 Numerical Optimization

We find parameter estimates by maximizing a quasi-likelihood derived from the innovation residuals of a Kalman filter. I do this using the unconstrained Broyden-Fletcher-Goldfarb-Shanno (BFGS) optimization algorithm as implemented in the SciPy Python package. (Virtanen et al., 2020) In order to benefit from the improved performance and stability characteristics of unconstrained optimization relative to algorithms capable of accommodating constraints, we need to reparameterize our model. The parameters $\mu, \kappa, \alpha, \gamma$, and s^2 are optimized on the logarithmic scale in order to ensure positivity. I have found that numerical issues arise when evaluating the quasi-likelihood with parameters violating the Feller condition $2\mu\kappa > \sigma^2$. The following transformation of σ ensures that it condition is always satisfied: suppose we optimize over an unconstrained variable $x \in \mathbb{R}$. The sigmoid of x

$$\operatorname{sigmoid}(x) := \frac{1}{1 + e^{-x}}$$

then takes values in the interval $(0, 1)$. Finally, parameterize σ by

$$\sigma^2 := 2\mu\kappa \operatorname{sigmoid}(x) \in (0, 2\mu\kappa).$$

This ensures that both the Feller condition and the non-negativity of σ are satisfied. Note that the Feller constraint is imposed here only for numerical reasons. BFGS optimization also returns an estimate of the inverse Hessian of the parameters. As an approximation of the inverse Fisher information this can be used to construct corresponding Wald confidence intervals. We will use these in the following section to account for the effect of uncertainty in parameter estimates on predictive distributions. To ensure that the confidence intervals do not cover impossible negative values we will construct normal approximate confidence intervals of the optimized log-parameters and transform them back to the real domain. Because the volatility σ is a function of multiple optimized parameters, we find its confidence interval using the delta method. Let $\theta := (\log \mu, \log \kappa, x, \log \alpha, \log \gamma, \log s^2)$ be the transformed parameter vector returned by the optimizer and H^{-1} its inverse Hessian. We then have that

$$\operatorname{Var}[\sigma] = \nabla_\theta \sigma^{\mathrm{T}} H^{-1} \nabla_\theta \sigma.$$

Note that

$$\frac{\partial \sigma^2}{\partial x} = 2\mu\kappa \operatorname{sigmoid}'(x) = 2\mu\kappa \operatorname{sigmoid}(x) \left[1 - \operatorname{sigmoid}(x)\right].$$

Therefore the full gradient of σ with respect to θ is

$$\nabla_\theta \sigma = \frac{\sigma}{2}(1, 1, 1 - \operatorname{sigmoid}(x), 0, 0, 0)^{\mathrm{T}}.$$

The resulting Wald intervals should be interpreted as rough measures of parameter uncertainty, not as reliable confidence intervals.

In order to ensure the stability and performance of the numerical optimization initial values must also be chosen with some care. I have had success with the following selection procedure: the instantaneous intensity defined in Section 3.2 gives us a rough estimate $\{\hat{\lambda}_t\}_{t=1}^n$ of the latent intensity process, which turns out to agree reasonably well with the Kalman filter estimate. We take the mean of this series as an initial guess for the value of μ and the standard deviation around this mean as the initial value of σ. The mean-reversion strength κ can then be chosen such that the Feller condition is satisfied with some margin; I have used $2\mu\kappa = 1.5\sigma^2$. For lack of a more sophisticated method, I have taken the initial values of α and γ to simply be equal to 2 and 1, respectively. An initial value of $s = 0.001$ has been used. This procedure turns out to produce an acceptably stable and performant optimization routine, in particular given the identifiability issues inherent to the model.

In a pure CIR process, the volatility parameter σ pushes the process away from its mean μ, while the mean-reversion strength κ counteracts this effect. This manifests in the process' stationary distribution, the variance of which depends on the ratio σ^2/κ. The result is poor identifiability, where a large value of κ can be compensated for by increasing σ, in an inexact manner of speaking. Adding the jump component complicates parameter identification further. A large volatility manifests itself similarly to many small jumps, making it difficult to independently identify σ, α and γ. This all results in the joint optimization of $\mu, \sigma, \kappa, \alpha, \gamma$ and s^2 being fairly unstable. As we will see in Section 6, I have only been able to achieve good results on real data after experimenting with which parameters to keep fixed and which to optimize, in a trial-and-error fashion.

5 Forecasting

A fitted price series model can be used to forecast future market prices. Since the JCIR stochastic differential equation has no closed form solution, it is not possible to derive an exact predictive distribution. It is, however, easy to simulate the price process. This situation lends itself to a Monte Carlo procedure: use the available data to estimate parameters and current intensity, as given by the Kalman filter, and use these to simulate a large number of future price paths. The pointwise empirical distributions of these simulations can then be used to compute the mean and quantiles of the model-implied predictive distribution. Apart from the inherent unpredictability of the intensity process, these simulations have two key sources of uncertainty: about the correct parameter values, and about the current intensity level. The former is captured by the normal approximate confidence intervals derived in Section 4.2. The best available estimate of the latter uncertainty is the variance of the intensity prediction given by the Kalman filter. To account for these two effects, the Monte Carlo simulations are initialized with random parameters and initial intensity drawn from the corresponding distributions. Figure 3 shows a 75-day price forecast and 95 % prediction intervals for the market "Khamenei Out as Supreme Leader of Iran in 2025?" starting on 31 July, using 1,000 Monte Carlo samples. The parameters only were estimated only on prices up to that point, with the same estimation configuration as the one described below in Section 6. Note that the asymmetric prediction intervals arise due to the JCIR process, and thus also the yes-price, only admitting positive jumps.

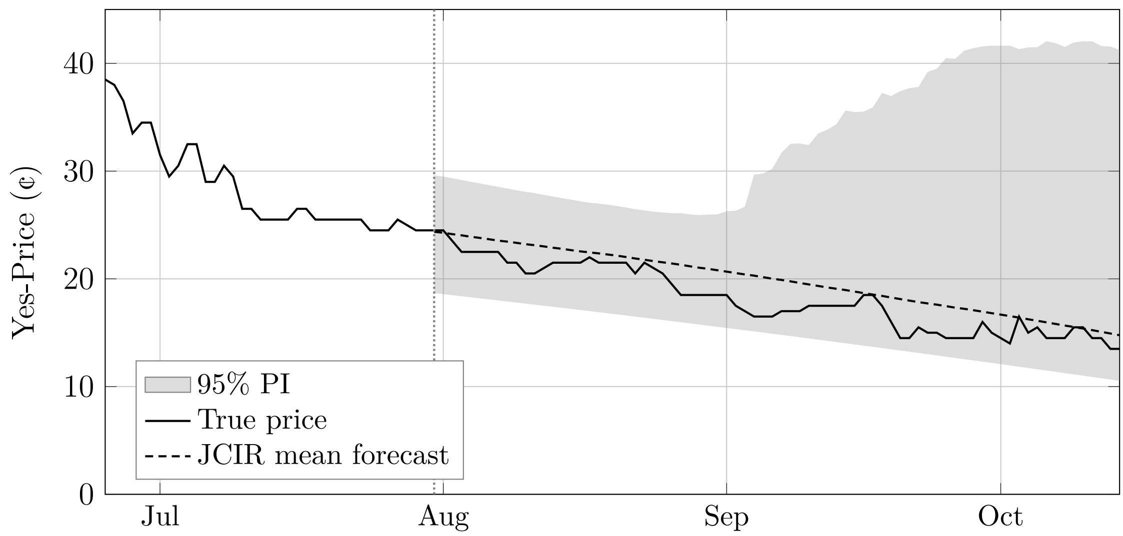


Figure 3: 75-day price forecast for the market “Khamenei Out as Supreme Leader of Iran?", starting at 31 July, using 1,000 Monte Carlo simulations, with randomized parameters and initial intensities. The solid line shows the true price and the dashed line the mean forecasted price. The shaded band marks 95 % prediction intervals. The dotted vertical line marks the forecast origin.

A fitted price model also allows us to study when the market will resolve. Recall that we interpret the yes-price as the probability of the market resolving before its end date T. The corresponding probability of resolution before some other time s is simply the price of another yes-contract with s as its end date. Given parameter estimates and an initial intensity at some time t we can compute this probability for any $s > t$ using Eq. ($\star$).

To account for uncertainty in the intensity and parameter estimates we can use a similar Monte Carlo procedure as for the price forecasts. We draw parameters and an initial intensity a large number of times from the distributions previously described and use these to compute resolution probabilities. The mean and empirical quantiles of this sample then give us point estimates with confidence intervals. This is done in Figure 4 for the market “Khamenei Out as Supreme Leader of Iran in 2025?” using the same parameter estimate as the forecast above. Confidence intervals were calculated from 2,000 Monte Carlo samples. Note that for $s = T$ the mean probability coincides with the market price, as expected. There is, of course, nothing preventing us from computing the probability of market resolution by some time after the end date T. Currently, market platforms like Polymarket host multiple contracts with different end dates on the same event. The pricing model developed here can not only be used to detect arbitrage opportunities between such markets, but also calls into question the need to run multiple markets in the first place. With a well-calibrated model, only one market would be needed to compute the probability of resolution by any given date. It would be valuable to study the accuracy of these computed probabilities and how they vary with distance from traded end dates.

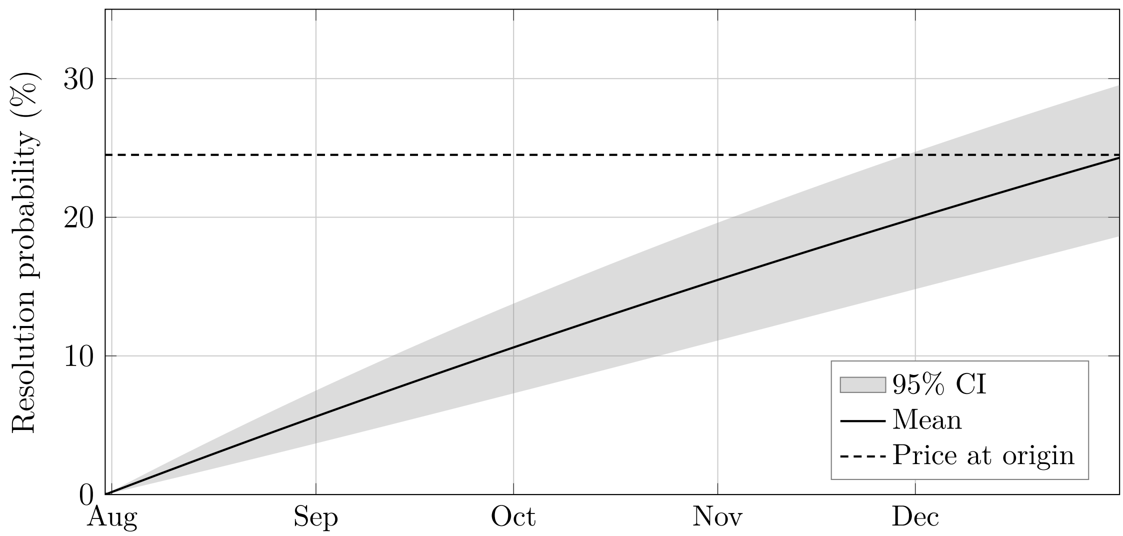


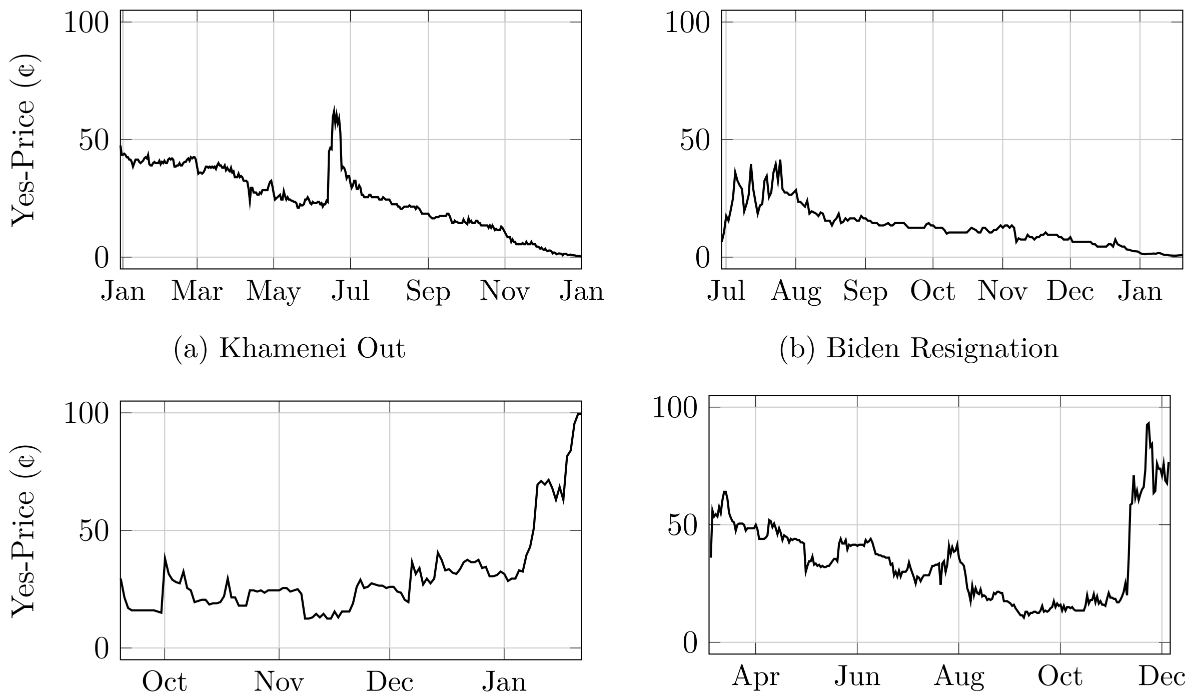
Figure 4: Model-implied probability of resolution by a given date for the market "Khamenei Out as Supreme Leader of Iran in 2025?", using parameters estimated on data up to 31 July 2025. The solid line is the point estimate. The shaded 95% confidence band was computed from 2,000 Monte Carlo samples, with randomized parameters and initial intensity. The dashed horizontal line shows the observed yes-price at the forecast origin.

6 Model Evaluation

In order to evaluate the quality of the proposed model, we will study its ability to predict the prices of four historical prediction markets hosted on Polymarket:

1. Khamenei Out as Supreme Leader of Iran in 2025?
2. Will Biden Resign Before the End of his Term?
3. TikTok Banned in the U.S. Before May 2025?
4. Will Bitcoin hit $100k in 2024?

The daily price histories of these markets are shown in Figure 5. They were retrieved using the Polymarket API. Before market resolution, no prices reached zero or one. The market on the resignation of President Biden was originally listed under the question "Will Biden Complete his Term as President?". Since our model assumes that the market resolving to *yes* corresponds with a change from the current *status quo*, I have inverted the phrasing, which swaps the yes- and no-prices.



(a) Khamenei Out

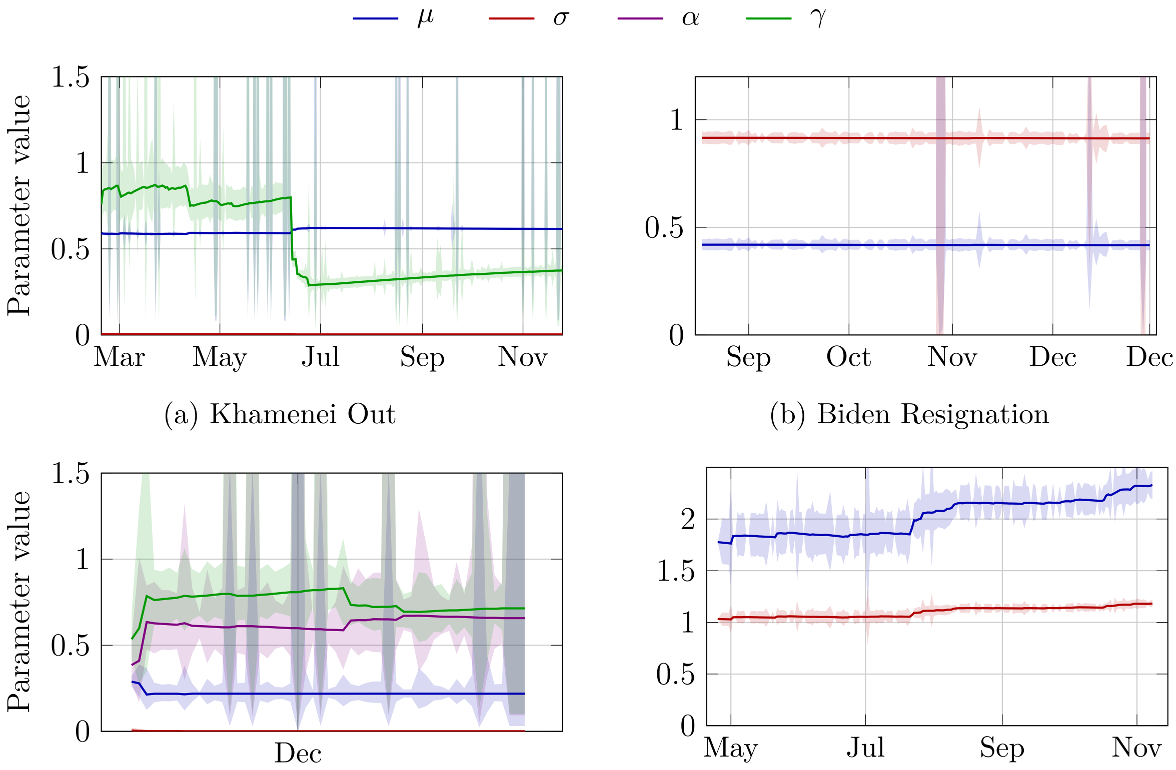
(b) Biden Resignation

(c) TikTok Ban

(d) Bitcoin $100k

Figure 5: Daily yes-prices of four timing prediction markets hosted on Polymarket.

Fitting the model to real data is not straightforward, due to the identifiability issues discussed in Section 4.2. Appropriate optimizer settings are most easily found through trial and error. I have found the best model fit and optimization convergence using the settings described in Table 1. This is done to ensure stability, but does not adress the potential risk of over-fitting. The empirical evaluation done here should therefore be considered exploratory in nature. The model was re-estimated daily on expanding windows for each market, with a minimum size of 50 days; for the sake of computational performance, each parameter optimization was initialized at the previous estimate. The first estimate used the initial value selection procedure described in Section 4.2. The last 10 % of the price



(a) Khamenei Out

(b) Biden Resignation

(c) TikTok Ban

(d) Bitcoin $100k

Figure 6: Daily parameter estimates on expanding-window estimation with a minimum of 50 observations, excluding the last 10% of each price history. Shaded bands indicate 95% Wald confidence intervals.

series was discarded, in order to avoid numerical issues with terms involving the inverse of the remaining time until the market's end. The estimates of the non-fixed parameters as shown in Figure 6. The estimates of the observation standard deviation s for the Khamenei-market are not shown, because they were insignificantly small, ranging from ca. 10^{-6} to 10^{-3}; this parameter was held fixed for the other markets. Note that the 95 % Wald confidence intervals vary greatly, with some briefly being implausibly wide. This reveals that the BFGS-optimizer's estimate of the inverse Hessian is unstable. We will see this propagated to the model-implied prediction intervals. A more robust measure of parameter uncertainty could be attained by a parametric bootstrap procedure, at significant computational expense.

Table 1: Optimizer configurations used for parameter estimation. They are necessary to ensure stability due to weak identification.

Market	Configuration
Khamenei Out	Fixed $\kappa \equiv 2$, $\alpha \equiv 0.3$
Biden Resignation	Fixed $\kappa \equiv 1$, $\alpha \equiv 0$, $s \equiv 0.01$
TikTok Ban	Fixed $\kappa \equiv 1$, $s \equiv 0.001$
Bitcoin $100k	Fixed $\kappa \equiv 0.3$, $\alpha \equiv 0$, $s \equiv 0.001$

A simple way to evaluate the pricing model's predictive performance is to consider the accuracy of point predictions at various horizons. Table 2 and Figure 7 contain root mean squared errors of predictions at 1, 3, and 10 day horizons on the four markets considered. They were made using 2,000 Monte Carlo samples. As with the parameter estimates, no predictions were made on the last 10 % of the price series, in order to avoid numerical problems. The table also includes the root mean squared errors of a naive persistence estimator as a reference point. The performance of the instantaneous, homogeneous, and exponentially weighted estimators developed in Section 3.2 and assuming constant future intensity is also listed. The exponential weighting was chosen to imply an effective rectangular window size of 10. The predictions over time are also shown in Figures 8–11. The prediction intervals are shown at a level of 67 %, because ordinary 95 % intervals are very noisy. We clearly see how instability in the parameter confidence intervals are propagated to the price prediction intervals, through the random parameter draw for each Monte Carlo sample.

Across all four markets, the instantaneous and naive estimators produce nearly identical errors, and the JCIR model rarely improves on naive persistence. The exponentially weighted and homogeneous constant-intensity estimators occasionally beat it at the ten-day horizon, but no model consistently dominates across markets. This difficulty in outperforming a simple persistence forecast would be explained by daily returns being uncorrelated. Conditional on the market not resolving, the modeled price process is not a martingale; the yes-price goes to zero as the market's end date approaches. However, looking at the autocorrelation functions of the considered markets daily log returns in Figure 12, we see little systematic or significant autocorrelation. Log-returns were computed as the logarithm of daily price differences. The dashed gray lines mark the 95 % confidence bands under a white-noise null. This suggests that at the daily level, this systematic trend to zero is weak in comparison to the effect of the random diffusion of the intensity process. It is therefore not surprising that future prices are difficult to predict. This is consistent with the reports of Majumder et al. (2009), who observe uncorrelated returns in a range of political prediction markets.

In many applications we are not only interested in making point predictions, but also in modeling the whole transition distribution. One way to check whether the model produces well-calibrated predictive distributions is to consider the coverage rates of prediction intervals constructed at various levels. At each time step and for a given horizon, we expect a well-calibrated prediction interval at, say, a level of 50 % to cover the true future price

Table 2: Root mean squared prediction errors in cents at horizons of 1, 3, and 10 days. All models were re-estimated daily on expanding windows with a minimum window size of 50 observations. Predictions were made using 2,000 Monte Carlo samples.

	Khamenei			Biden			TikTok			Bitcoin		
Model	1d	3d	10d	1d	3d	10d	1d	3d	10d	1d	3d	10d
JCIR	3.2	4.7	7.9	1.3	1.5	1.7	5.3	6.3	8.9	2.8	3.8	6.0
Naive	2.2	3.8	8.1	1.0	1.4	2.2	3.7	5.6	8.2	2.3	3.3	5.9
Instantaneous	2.2	3.8	7.8	1.0	1.4	1.9	4.0	5.9	8.9	2.4	3.4	5.9
Exp. W=10	4.9	5.8	7.3	1.3	1.4	1.5	6.8	7.7	10.3	4.2	4.8	6.3
Homogeneous	6.7	6.7	6.8	2.4	2.4	2.2	11.0	11.9	14.1	5.5	5.7	6.0

around half of the time. Computing such coverage rates for a range of levels gives us information about model calibration across the entire predictive distribution. Nominal coverage rates at low levels indicate that the model captures the central part of the distribution well, while at high levels this suggests that tail behaviour is adequately represented. Note that this analysis is incomplete, because it does not also consider the width of the prediction intervals. Figure 13 contains plots of actual vs. nominal coverage rates for the four markets at 1, 3, and 10 day horizons, computed using 10,000 Monte Carlo samples. The results show reasonable calibration for several market-horizon combinations, but the pattern is not uniform. Shorter horizons are not consistently better calibrated than longer ones, despite short-horizon prediction being an easier problem. This suggests that different sources of uncertainty dominate at different time scales. At short horizons, uncertainty about the current intensity level dominates, and the filtered estimate appears poorly calibrated. Perhaps the Gaussian approximation made by the Kalman filter is inadequate here. At longer horizons, parameter estimation uncertainty and accumulated stochastic intensity variation become the primary drivers of the predictive distribution.

Looking at the predictions over time, the longer horizons suffer especially when large jumps in price occur. This contributes to the under-coverage at the 3 and 10 day horizons for the Khamenei- and TikTok-markets. The estimated jump-parameters are seemingly unable to capture this behavior, suggesting the need for separating common small jumps and rare large jumps in the dynamics of the price process. Despite the poor point predictions, we observe reasonably good calibration for some markets and horizons. Even this limited model might therefore be useful in applications where modeling the entire predictive distribution, not just its mean, is important.

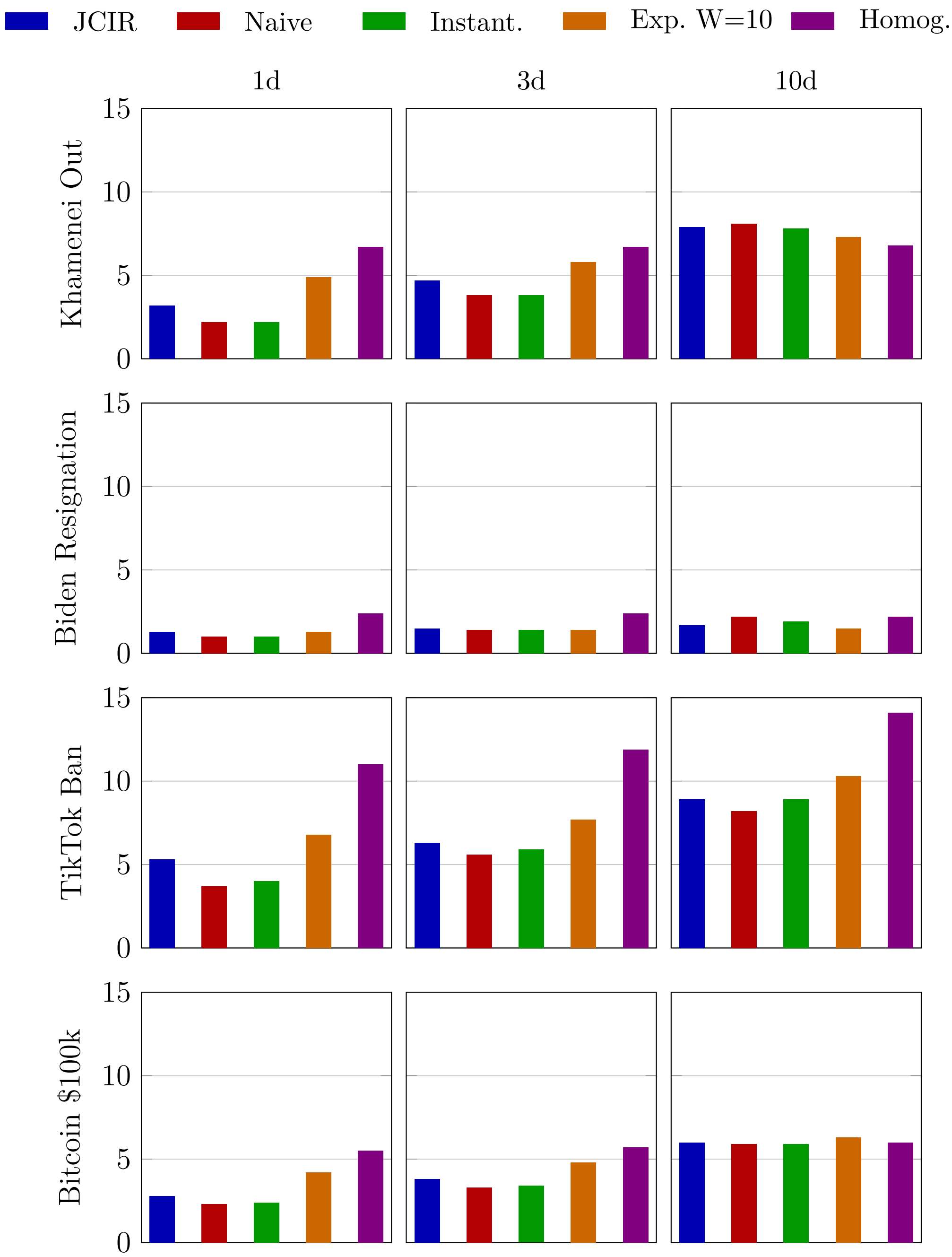


Figure 7: Root mean squared prediction errors at 1, 3, and 10 day horizons. Models were re-estimated daily on expanding windows with a minimum of 50 observations. Predictions were made using 2,000 Monte Carlo samples.

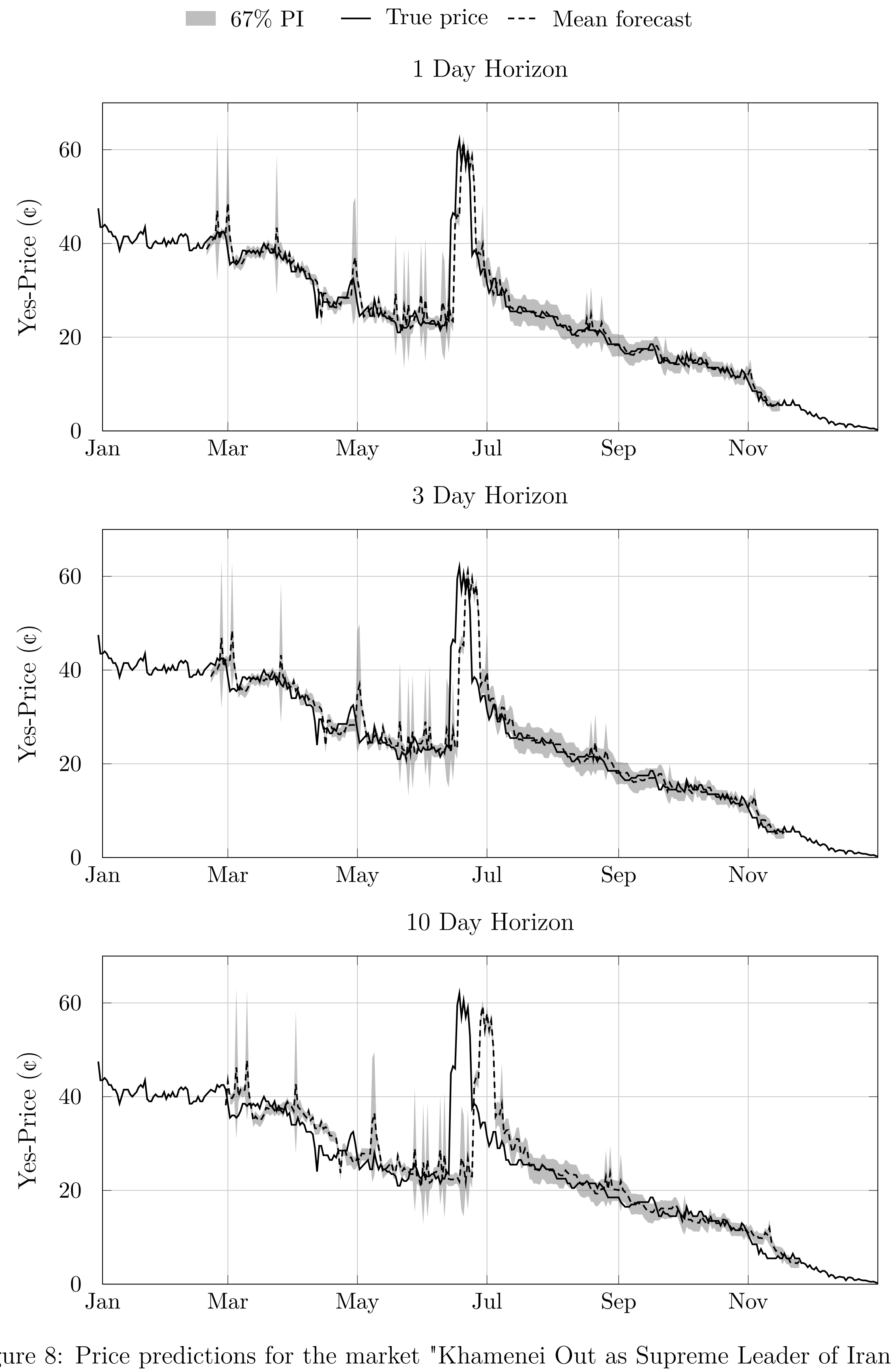


Figure 8: Price predictions for the market "Khamenei Out as Supreme Leader of Iran in 2025?" at 1, 3, and 10 day horizons. The solid line marks the true price history and the dashed the mean prediction made using 2,000 Monte Carlo samples. The shaded bands mark 67 % prediction intervals.

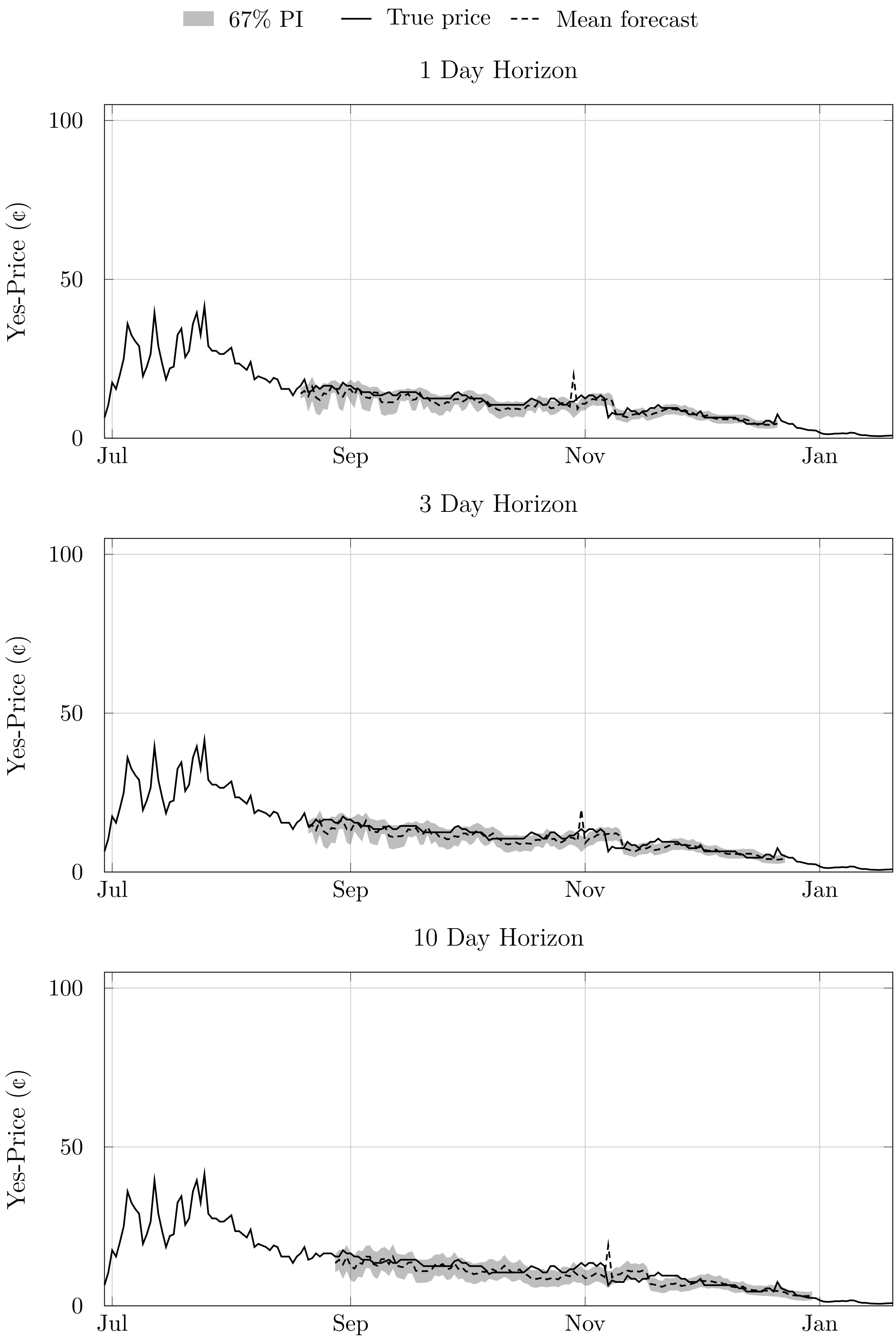


Figure 9: Price predictions for the market "Will Biden Resign Before the End of his Term?" at 1, 3, and 10 day horizons. The solid line marks the true price history and the dashed the mean prediction made using 2,000 Monte Carlo samples. The shaded bands mark 67 % prediction intervals.

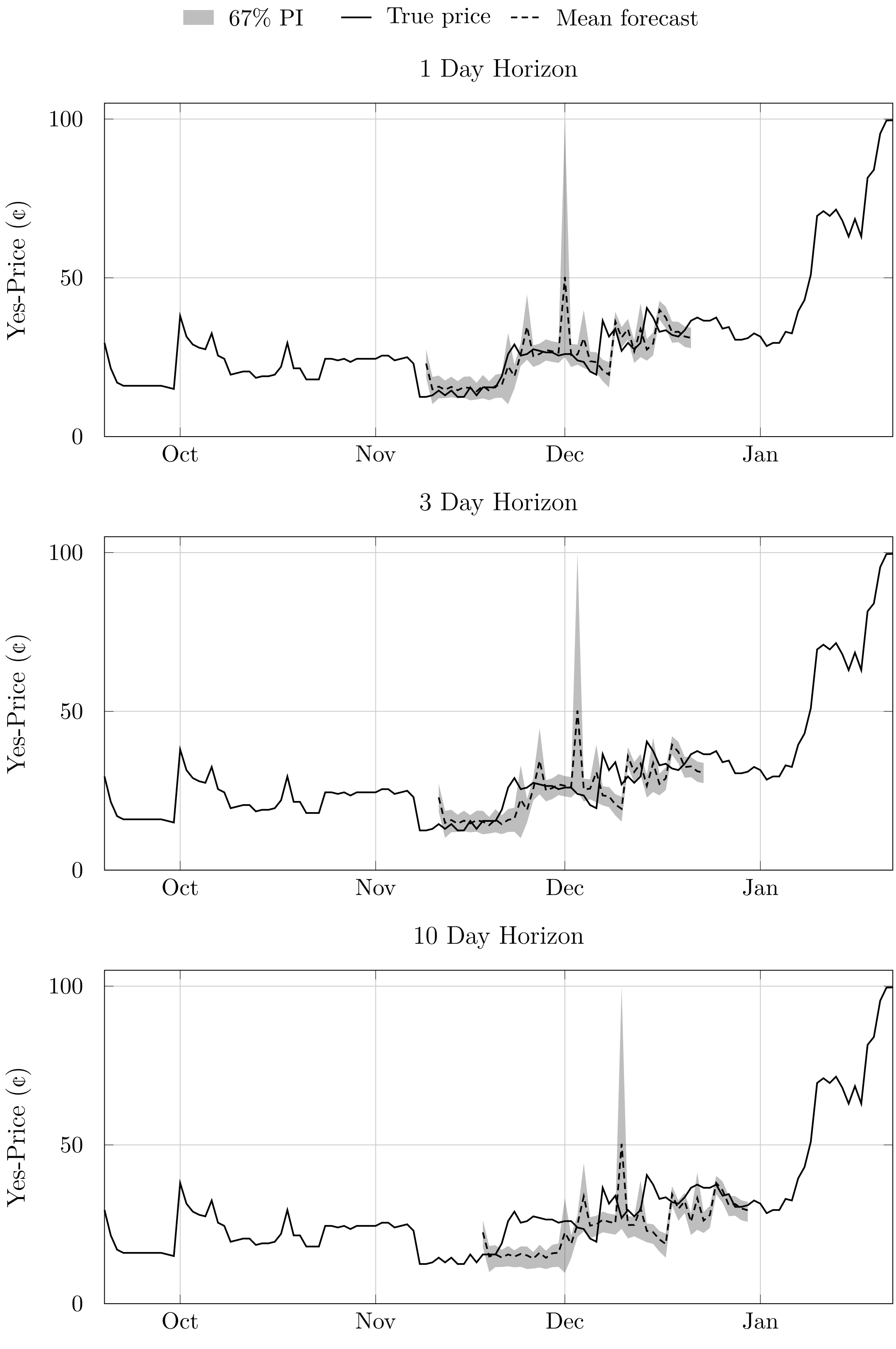


Figure 10: Price predictions for the market "TikTok Banned in the U.S. Before May 2025?" at 1, 3, and 10 day horizons. The solid line marks the true price history and the dashed the mean prediction made using 2,000 Monte Carlo samples. The shaded bands mark 67 % prediction intervals.

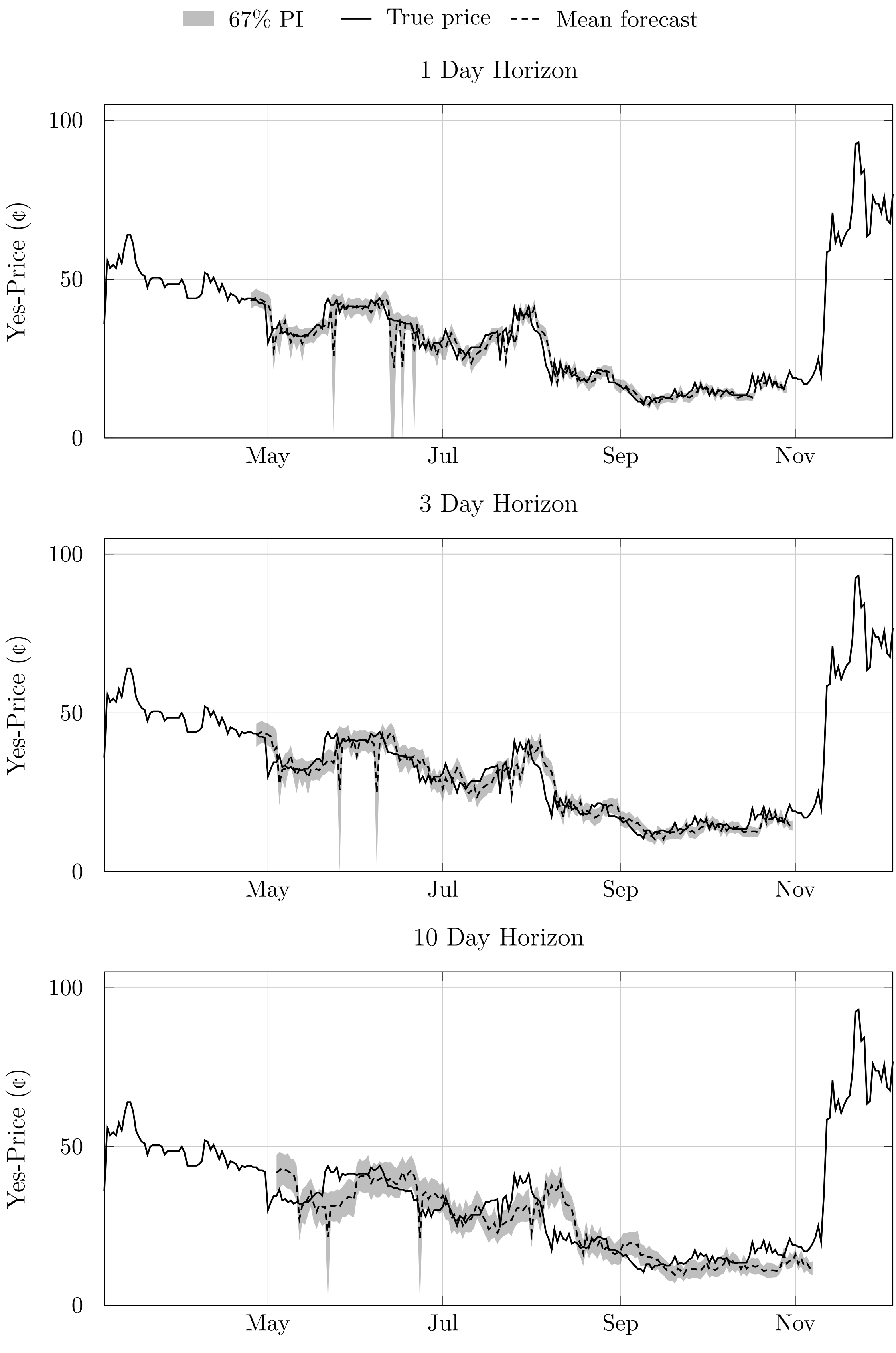


Figure 11: Price predictions for the market "Will Bitcoin hit $100k in 2024?" at 1, 3, and 10 day horizons. The solid line marks the true price history and the dashed the mean prediction made using 2,000 Monte Carlo samples. The shaded bands mark 67 % prediction intervals.

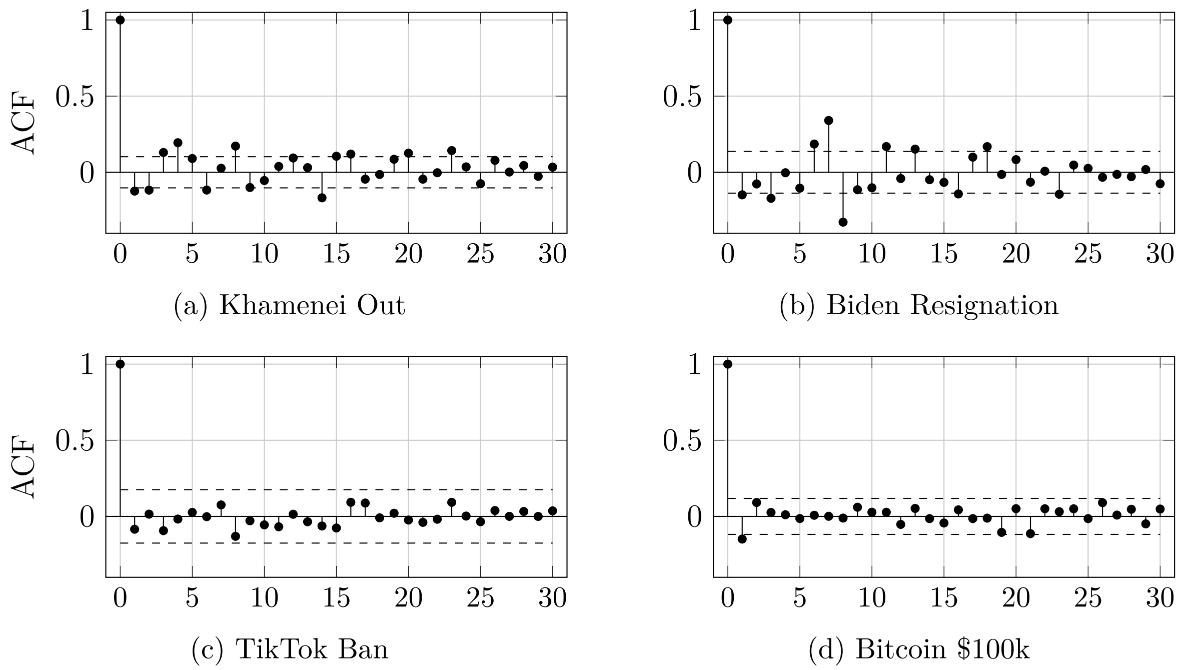


Figure 12: Autocorrelation functions of daily log-returns for the markets in Figure 5. Dashed lines mark 95% confidence bands under a white-noise null.

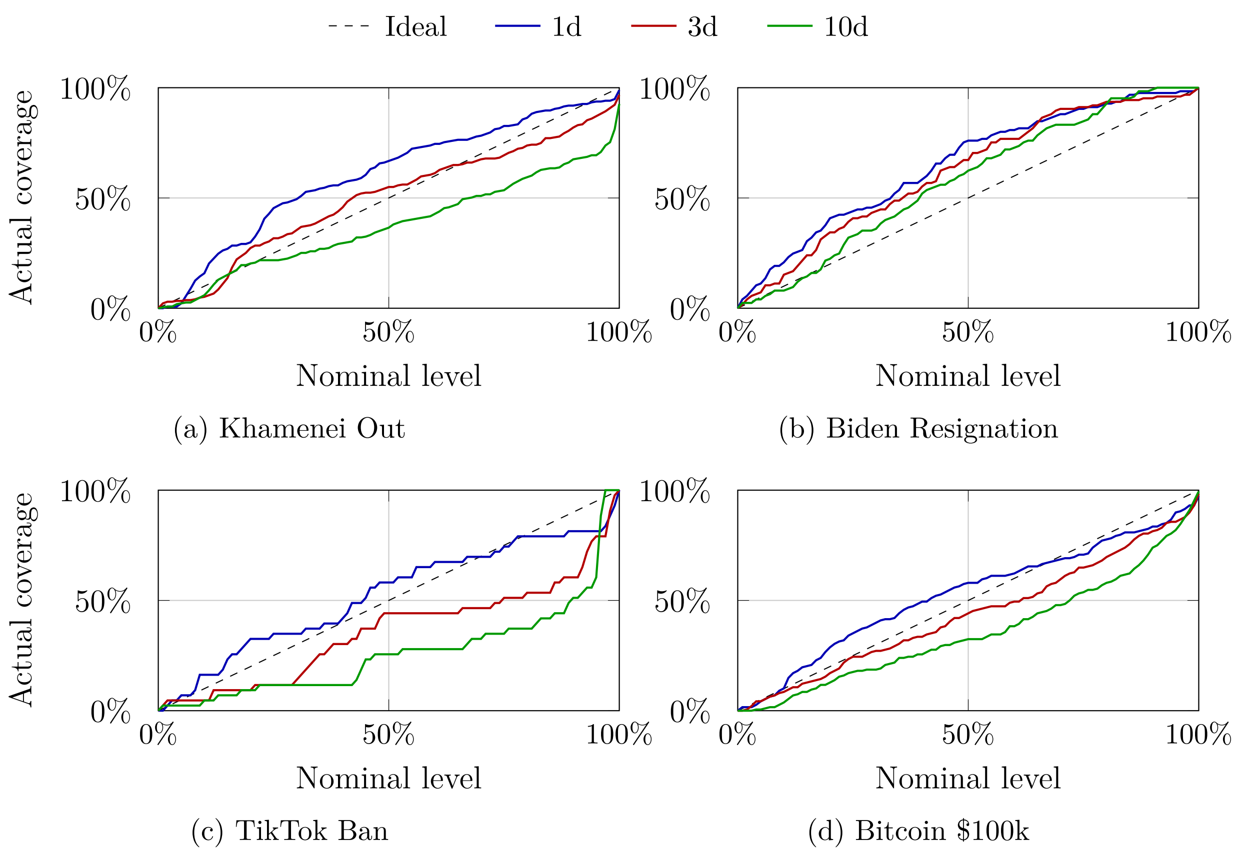


Figure 13: Actual vs. nominal coverage rates of prediction intervals at 1, 3, and 10 day horizons. The dashed diagonal marks ideal calibration. Based on 10,000 Monte Carlo samples.

7 Conclusion

The last few years have seen an explosion of interest in prediction markets. For the first time, platforms exist on which a wide variety of liquid markets are open for trading. Timing markets are an important part of this development, especially in geopolitical contexts, but they have received comparatively little attention in the literature. In this thesis I present a first pricing framework for these markets. By assuming that the underlying event arrives with a doubly stochastic Poisson process, the no-price of a timing contract can be interpreted as a conditional survival probability of the process' first arrival time. This makes timing-market pricing mathematically equivalent to both reduced-form models of defaultable bonds and zero coupon bond-pricing under a stochastic interest rate. This analogy turns the problem of pricing timing contracts into the more familiar one of specifying and estimating a stochastic intensity process.

The particular specification studied in this thesis, a Cox-Ingersoll-Ross intensity with compound Poisson jumps, has important limitations. Its positive-jump structure is not well suited to prediction-market data, where news can move prices sharply in either direction. The model also suffers from substantial parameter-identification problems, especially when volatility and jump components are estimated jointly. These weaknesses are visible in the empirical results, with the JCIR model not consistently outperforming simple persistence or constant-intensity baselines. At the same time, the model sometimes produces reasonably calibrated predictive intervals, suggesting that the reduced-form approach may be more useful for modeling the full predictive distribution than for improving short-horizon mean forecasts.

The main contribution of the thesis is therefore not the final JCIR specification itself, but the reduced-form framework in which it sits. Once timing contracts are viewed as first-arrival claims, a wide range of alternative intensity dynamics and estimation methods become available. A natural next step would be to allow the intensity to mean-revert around a piecewise constant level, with jumps or regime changes representing discrete news shocks. We have also seen indications that it could be useful to separate the dynamics of small common jumps and large rare ones, representing different types of news events. There are also various ways to improve the estimation procedure. Parametric bootstrapping could easily provide more robust estimates of parameter uncertainty, although it is computationally expensive. More sophisticated filtering techniques could also be used to avoid the simplifications made in order to adapt the model to a standard Kalman filter. These include the unscented Kalman filter and various types of particle filters. Another important extension would be to estimate the model jointly across several timing markets based on the same underlying event but with different end dates. Such an approach would make better use of available market information and would allow the study of inter-market pricing consistency.

This thesis shows that timing prediction markets are naturally modeled by traditional techniques in mathematical finance. Even though the empirical performance of the specific model considered here is mixed, the approach opens a path toward richer models of timing-market dynamics. If prediction markets continue to grow in liquidity and institutional relevance, models of this kind may become useful for market making, risk management, and the comparison of probabilities across contracts with different maturities.

References

Archak, N., & Ipeirotis, P. G. (2008). *Modeling volatility in prediction markets* (Working Paper No. CeDER-08-07). Leonard Stern School of Business, New York University.

Archak, N., & Ipeirotis, P. G. (2009). Modeling volatility in prediction markets. *Proceedings of the 10th ACM conference on Electronic commerce*, 275–284.

Bakshi, G., Gao, X., & Zhong, Z. (2022). Decoding default risk: A review of modeling approaches, findings, and estimation methods. *Annual Review of Financial Economics*, *14*(1), 391–413.

Beckmann, K., & Werding, M. (1996). Passauer Wahlbörse: Information processing in a political market equipment. *Kyklos*, *49*(2).

Berg, J., Forsythe, R., Nelson, F., & Rietz, T. (2008). Results from a dozen years of election futures markets research. *Handbook of experimental economics results*, 742–751.

Bohm, P., & Sonnegard, J. (1999). Political stock markets and unreliable polls. *Scandinavian Journal of Economics*, *101*(2), 205–222.

Brigo, D., & Mercurio, F. (2006). *Interest rate models—theory and practice: With smile, inflation and credit* (2nd). Springer.

Chen, M. K., Ingersoll, J. E., & Kaplan, E. H. (2008). Modeling a presidential prediction market. *Management Science*, *54*(8), 1381–1394.

Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica*, *53*(2), 385–407.

Dalen, S. (2025). Toward Black Scholes for prediction markets: A unified kernel and market maker's handbook. *arXiv:2510.15205*.

Feller, W. (1951). Two singular diffusion problems. *Annals of mathematics*, *54*(1), 173–182.

Forestal, R. L., Zhang, P. C., & Pi, S.-M. (2020). Prediction markets: A systematic review and meta-analysis [Paper 51]. *Proceedings of the 20th International Conference on Electronic Business (ICEB 2020)*.

Frongillo, R., Papireddygari, M., & Waggoner, B. (2024). An Axiomatic Characterization of CFMMs and Equivalence to Prediction Markets. In V. Guruswami (Ed.), *15th innovations in theoretical computer science conference (ITCS 2024)* (51:1–51:21, Vol. 287). Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

Gao, J., Wang, Z., Wu, W., & Yu, D. (2025). Price interpretability of prediction markets: A convergence analysis. *Operations Research*, *73*(1), 157–177.

Gjerstad, S., & Hall, M. (2005). Risk aversion, beliefs, and prediction market equilibrium. *Economic Science Laboratory, University of Arizona*.

Hanson, R. (1990). Market-based foresight-a proposal. *Foresight Update*, *10*(1), 3.

Hanson, R. (1992). Idea futures: Encouraging an honest consensus. *Extropy*, *3*(2), 7–17.

Hanson, R. (1995). Could gambling save science? Encouraging an honest consensus. *Social Epistemology*, *9*, 3–33.

Hanson, R. (2007). Logarithmic market scoring rules for modular combinatorial information aggregation. *The Journal of Prediction Markets*, *1*(1), 3–15.

Hanson, R. (2009). On market maker functions. *The Journal of Prediction Markets*, *3*(1), 61–63.

Horn, C. F., Ivens, B. S., Ohneberg, M., & Brem, A. (2014). Prediction markets: A literature review 2014. *The Journal of Prediction Markets*, *8*(2), 89–126.

HTX Insights. (2026). *2025 prediction market review: Total trading volume exceeds $50 billion, top two players hold over 97.5% market share*. Retrieved May 8, 2026, from https://www.htx.com/news/2025-prediction-market-review-total-trading-volume-exceeds-5-Q4y3TtCj/

Leigh, A., Wolfers, J., & Zitzewitz, E. (2003). *What do financial markets think of war in Iraq?* (Working Paper No. 9587). National Bureau of Economic Research. Cambridge, MA.

Majumder, S. R., Diermeier, D., Rietz, T. A., & Amaral, L. A. N. (2009). Price dynamics in political prediction markets. *Proceedings of the national academy of sciences*, *106*(3), 679–684.

Manski, C. F. (2006). Interpreting the predictions of prediction markets. *Economics Letters*, *91*(3), 425–429.

Mikulevičius, R., & Platen, E. (1988). Time discrete Taylor approximations for Itǒ processes with jump component. *Mathematische Nachrichten*, *138*(1), 93–104.

Nueve, E., & Waggoner, B. (2026). Smooth quadratic prediction markets. *Advances in Neural Information Processing Systems*, *38*, 61464–61487.

Polymarket. (2024). *Presidential election winner 2024* [Total trading volume: $3.68B.]. Retrieved May 8, 2026, from https://polymarket.com/event/presidential-election-winner-2024

Rhode, P. W., & Strumpf, K. S. (2004). Historical presidential betting markets. *Journal of Economic Perspectives*, *18*(2), 127–142.

Tziralis, G., & Tatsiopoulos, I. (2007). Prediction markets: An extended literature review. *The journal of prediction markets*, *1*(1), 75–91.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., . . . SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, *17*, 261–272.

Wolfers, J., & Zitzewitz, E. (2006a). *Interpreting prediction market prices as probabilities* (Working Paper No. 12200). National Bureau of Economic Research. Cambridge, MA. https://www.nber.org/papers/w12200

Wolfers, J., & Zitzewitz, E. (2006b). *Prediction markets in theory and practice* (Working Paper No. 12083). National Bureau of Economic Research. Cambridge, MA. https://www.nber.org/papers/w12083

Wörle, S. M. (2013). Wanna bet there will be war? A time-series analysis of prediction markets during the Libya conflict 2011. *The Journal of Prediction Markets*, *7*(3), 1–28.

Master’s Theses in Mathematical Sciences 2026:E34
ISSN 1404-6342

LUNFMS-3141-2026

Mathematical Statistics
Centre for Mathematical Sciences
Lund University
Box 118, SE-221 00 Lund, Sweden

http://www.maths.lu.se/