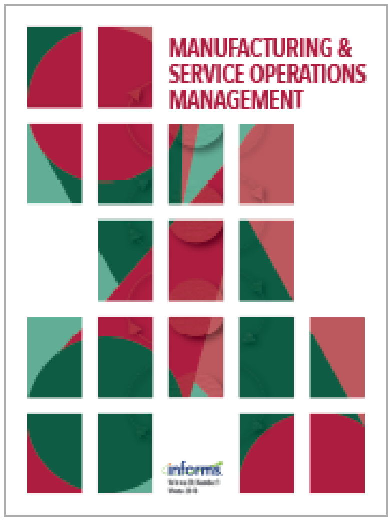


Manufacturing & Service Operations Management

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

The Nonstationary Newsvendor with (and Without) Predictions

Lin An, Andrew A. Li, Benjamin Moseley, R. Ravi

https://pubsonline.informs.org/journal/msom

MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Articles in Advance, pp. 1–16
ISSN 1523-4614 (print), ISSN 1526-5498 (online)

The Nonstationary Newsvendor with (and Without) Predictions

Lin An,[a,*] Andrew A. Li,[a] Benjamin Moseley,[a] R. Ravi[a]

[a] Tepper School of Business, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213
*Corresponding author

Contact: linan@andrew.cmu.edu, https://orcid.org/0009-0005-7840-2194 (LA); aali1@andrew.cmu.edu, https://orcid.org/0000-0002-9552-6421 (AAL); moseleyb@andrew.cmu.edu, https://orcid.org/0000-0001-8162-017X (BM); ravi@andrew.cmu.edu, https://orcid.org/0000-0001-7603-1207 (RR)

Received: July 2, 2024
Revised: November 29, 2024
Accepted: February 3, 2025
Published Online in Articles in Advance: March 4, 2025

https://doi.org/10.1287/msom.2024.1168

Abstract. ***Problem definition*:** The classic newsvendor model yields an optimal decision for a "newsvendor" selecting a quantity of inventory under the assumption that the demand is drawn from a known distribution. Motivated by applications such as cloud provisioning and staffing, we consider a setting in which newsvendor-type decisions must be made sequentially in the face of demand drawn from a stochastic process that is both unknown and nonstationary. All prior work on this problem either (a) assumes that the level of nonstationarity is known or (b) imposes additional statistical assumptions that enable accurate *predictions* of the unknown demand. Our research tackles the Nonstationary Newsvendor without these assumptions both with and without predictions. ***Methodology/results*:** In the setting without predictions, we first design a policy that we prove (via matching upper and lower bounds) achieves order-optimal regret; ours is the first policy to accomplish this without being given the level of nonstationarity of the underlying demand. We then, for the first time, introduce a model for generic (i.e., with no statistical assumptions) predictions with arbitrary accuracy and propose a policy that incorporates these predictions without being given their accuracy. We upper bound the regret of this policy and show that it matches the best achievable regret had the accuracy of the predictions been known. ***Managerial implications*:** Our findings provide valuable insights on inventory management. Managers can make more informed and effective decisions in dynamic environments, reducing costs and enhancing service levels despite uncertain demand patterns. This study advances understanding of sequential decision-making under uncertainty, offering robust methodologies for practical applications with nonstationary demand. We empirically validate our new policy with experiments based on three real-world data sets containing thousands of time-series, showing that it succeeds in closing approximately 74% of the gap between the best approaches based on nonstationarity and predictions alone.

History: This paper was selected for Fast Track in the M&SOM Journal from the 2024 MSOM Service Operations SIG Conference.

Funding: L. An and B. Moseley were supported in part by a Google Research Award, an Infor Research Award, a Carnegie Bosch Junior Faculty Chair, NSF [Grants CCF-2121744 and CCF-1845146] and ONR [Grant N000142212702].

Supplemental Material: The online appendix is available at https://doi.org/10.1287/msom.2024.1168.

Keywords: newsvendor model • decision-making with predictions • regret analysis

1. Introduction

The newsvendor problem is a century-old model (Edgeworth 1888) that remains fundamental to the practice of operations management. In its original instantiation, a "newsvendor" is tasked with selecting a quantity of inventory before observing the demand for that inventory, with the demand itself randomly drawn from a *known* distribution. The newsvendor incurs a per-unit underage cost for unmet demand and a per-unit overage cost for unsold inventory. The objective is to minimize the total expected cost, and the classic result is that the optimal inventory level is a certain problem-specific quantile (depending only on the underage and overage costs) of the demand distribution.

This paper is concerned with a modern instantiation of the same model, consisting of a *sequence* of newsvendor problems over time, each with *unknown* demand distributions that *vary* over time. Although this version of the problem is arguably ubiquitous in practice today, it may be worth highlighting a few motivating examples:

- **Cloud provisioning:** Consider a website that provisions computational resources from a commercial cloud provider to serve its web requests. Such provisioning is typically done *dynamically*, say, on an hourly basis, with the aim of satisfying incoming requests at a sufficiently high service level. Thus, the website faces a single newsvendor problem every hour, with an hourly "demand" that can (and does) vary drastically over time.

• **Staffing:** A more traditional example is staffing, say, for a brick-and-mortar retailer, a call center, or an emergency room. Each day (or even each shift) requires a separate newsvendor problem to be solved, with demand that is highly nonstationary.

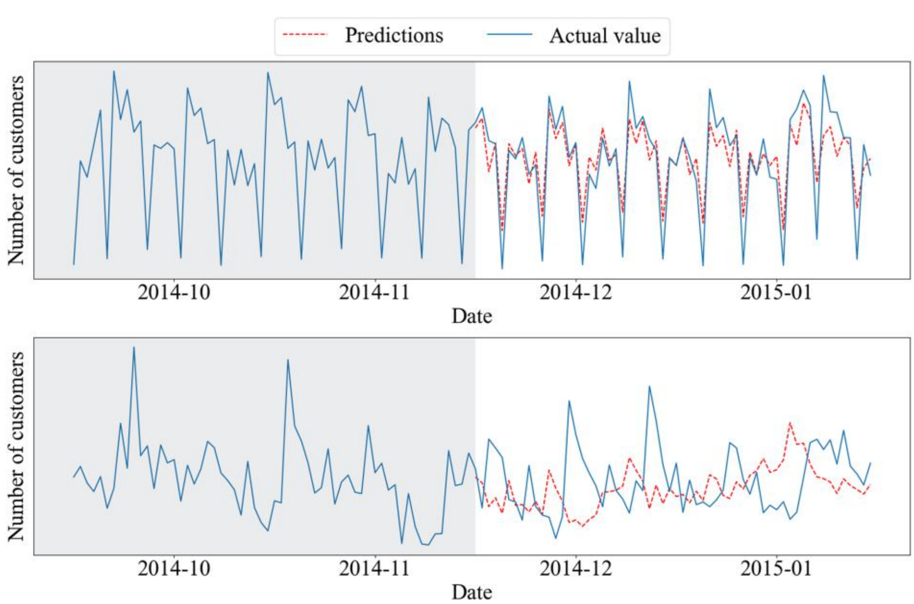
Despite its ubiquity, this problem is far from resolved precisely because the demand (or sequence of demand distributions) is both *nonstationary* and *unknown*; indeed, the repeated newsvendor with stationary but unknown demand was solved by Levi et al. (2015), and the same setting with known but nonstationary demand can be treated simply as a sequence of completely separate newsvendor problems. At present, there are by and large two existing approaches to this problem:

1. **Limited nonstationarity:** One approach is to design policies that "succeed" under limited nonstationarity; that is, the cost incurred by the policy should be parameterized by some carefully chosen measure of nonstationarity (e.g., quadratic variation) and nothing else. This approach has proved fruitful across a diverse set of problems ranging from dynamic pricing (Keskin and Zeevi 2017) to multiarmed bandit problems Besbes et al. (2014) to stochastic optimization (Besbes et al. 2015). Most relevant here, the recent work of Keskin et al. (2023) applies this lens to the newsvendor setting (we will discuss this work in detail momentarily). This approach yields policies with theoretical guarantees that are quite robust; no assumption on the demand (beyond the limited nonstationarity) is required. However, this is far removed from practice, where the next approach is more common.

2. **Predictions:** The second approach is to utilize some sort of *predictions* of the unknown demand. These predictions can be generated from simple forecasting algorithms for univariate time series all the way to state-of-the-art machine-learning models that leverage multiple time series and additional feature information. Therefore, these predictions may contain much more information than past demand data points, such as various features/contexts, or even black-box-type information that is nonidentifiable. In addition to being the de facto approach in practice, the use of predictions in newsvendor-type problems is well studied, and in fact, provable guarantees exist for many specific prediction-based approaches (Ban and Rudin 2019, Huber et al. 2019, Oroojlooyjadid et al. 2020, Zhang et al. 2024). All such guarantees rely on (at the very least) the demand and potential features being generated from a known family of stochastic models so that the framework and tools of statistical learning theory can be applied. Absent these statistical assumptions, it is unclear a priori whether the resulting predictions will be sufficiently accurate to outperform robust policies such as those generated in the previous approach. As a concrete example of this, see Figure 1, which demonstrates on a real set of retail data that prediction accuracy may vary drastically and unexpectedly, even when those predictions are generated according to the same procedure and applied during the same time period.

To summarize, the repeated newsvendor with unknown, nonstationary demand (which from here on we refer to as the *Nonstationary Newsvendor*) admits policies with nontrivial guarantees, which can be made significantly better or worse by following predictions. This suggests the opportunity to design a policy that

Figure 1. (Color online) Daily Number of Customers (in Solid Line) from September 2014 to January 2015, at Two Different Stores in the Rossmann Drug Store Chain



Notes. Predictions (in dashed line), starting November 2014, are generated using Exponential Smoothing with the same fitting process. The store in the upper subfigure has substantially more accurate predictions ($R^2 = 0.88$) than that of the lower subfigure ($R^2 = 0.11$).

uses predictions *optimally* in the sense that the predictions are utilized when accurate and ignored when inaccurate. Ideally, such a policy would run without knowledge of (a) the accuracy of the predictions and (b) the method with which they are generated. This is precisely what we accomplish in this paper.

1.1. The Nonstationary Newsvendor, With and Without Predictions

The primary purpose of this paper is to develop a policy that optimally incorporates *predictions* (defined in the most generic sense possible) into the *Nonstationary Newsvendor* problem. Naturally, a prerequisite to this is a fully solved model of the Nonstationary Newsvendor without predictions. At present, this prerequisite is only partially satisfied (via the work of Keskin et al. 2023), so a nontrivial portion of our contributions will be to fully solve this problem.

Without predictions, the Nonstationary Newsvendor consists of a sequence of newsvendor problems indexed by periods $t \in 1,\dots,T$, each with *unknown* demand distribution D_t. The level of nonstationarity is characterized via a *variation parameter* $v \in [0,1]$, where $v=0$ essentially amounts to stationary demand, and $v=1$ is effectively arbitrary (in a little more detail: a deterministic analog of quadratic variation is applied to the sequence of means $\{\mathbb{E}[D_1],\dots,\mathbb{E}[D_T]\}$, and $v \in [0,1]$ is the exponent such that this quantity equals T^v). Finally, we measure the performance of any policy using *regret*, which is the expected difference in the total cost incurred by the policy versus that of an optimal policy that "knows" the demand distributions. At minimum, we aim to design a policy that achieves *sublinear* (i.e., $o(T)$) regret, because such a policy would incur a per-period cost that is on average no worse than the optimal, as T grows. We will in fact design policies that achieve *order-optimal* regret with respect to the variation parameter v.

To this base problem, we introduce the notion of predictions. In each period, we receive a prediction a_t of the mean demand $\mu_t = \mathbb{E}[D_t]$ before selecting the order quantity. Our predictions are generic; no assumption is made on how they are generated. We measure the accuracy of the predictions through an *accuracy parameter* $a \in [0,1]$, defined such that $\sum_{t=1}^{T} |a_t - \mu_t| = T^a$. Notice that when $a=0$ the predictions are almost perfect, and when $a=1$ the predictions are effectively useless. We will characterize a precise threshold on a (which depends on v) that determines when the predictions should be utilized. Our primary challenge will be to design a policy that makes use of the predictions only when they are sufficiently accurate and without having access to a. As to the variation parameter v, we will separately consider policies that do and do not have access to v; this distinction will turn out to be the critical factor in classifying what is and is not achievable.

1.2. Our Contributions

Our primary contributions can be summarized as follows.

1. ***Nonstationary Newsvendor (without predictions):*** We completely solve the Nonstationary Newsvendor problem. This consists of first constructing a policy and proving an upper bound on its regret:

Theorem 1 (Informal). *There exists a policy that achieves $\tilde{O}(T^{(3+v)/4})$ regret*[1] *without knowing v.*

We then show that this regret is minimax optimal up to logarithmic factors.

Proposition 1 (Informal). *No policy can achieve regret better than $O(T^{(3+v)/4})$, even if v is known.*

As alluded to earlier, Keskin et al. (2023) previously initiated the study of the Nonstationary Newsvendor. Our results are distinct in terms of both modeling and theoretical contributions. We will expound these distinctions more carefully later on.

• **Modeling:** The most crucial difference in our model is that we allow both the demand and the set of possible ordering quantities to be *discrete*. This is certainly of practical concern (e.g., physical inventory, employees, and virtual machines are all indivisible units of demand), but moreover, we will show that the results of Keskin et al. (2023) require both the demand and set of feasible ordering quantities to be continuous. Thus, there is no overlap in our theoretical results.

• **Results:** Keskin et al. (2023) succeeded in designing a policy that achieved order-optimal regret, but crucially, their policy required that the variation parameter v be known. In addition to being concerning from a practical standpoint, this leaves open the theoretical question of what exactly is achievable in settings for which v is unknown. Our results show that the same regret can be achieved without knowing v.

2. ***Nonstationary Newsvendor with Predictions:*** We construct a policy that optimally leverages predictions; that is, it is robust to unknown prediction accuracy. To be precise, the previous contribution offers a policy that achieves $\tilde{O}(T^{(3+v)/4})$ regret, and predictions yield a simple policy that achieves $O(T^a)$ regret, so we would expect that the best possible regret is the minimum of these two quantities. We show this formally.

Proposition 2 (Informal). *No policy can achieve regret better than $O(T^{\min\{(3+v)/4, a\}})$, even if v and a are known.*

Our main algorithmic contribution is a policy that achieves this lower bound (up to log factors) without knowing the prediction accuracy:

Theorem 2 (Informal). *There exists a policy that achieves regret $\tilde{O}(T^{\min\{(3+v)/4, a\}})$, knowing v, and without knowing a.*

Finally, because our policy relies on knowledge of the variation parameter v, the remaining question is whether the same regret is achievable if both v and a are unknown. We show that in fact predictions cannot be incorporated in any meaningful way in this case.

Proposition 3 (Informal). *If v and a are unknown, then no policy can achieve regret better than $O(T^{\max\{(3+v)/4,a\}})$ for all $v, a \in [0,1]$.*

Our theoretical results are summarized in the Table 1. Each entry has a corresponding policy that achieves the stated regret, along with a matching lower bound.

3. ***Empirical Results:*** Finally, we demonstrate the practical value of our model (namely, the Nonstationary Newsvendor with Predictions) and our policy via empirical results on three real-world data sets that span our motivating applications above: daily web traffic for Wikipedia.com (of various languages), daily foot traffic across the Rossmann store chain, and daily visitors at a certain Japanese restaurant. These data sets together contain more than 1,000 individual time series on which we generate predictions of varying quality, using four different popular forecasting and machine-learning algorithms. We apply our policy and compare its performance against the two most-natural baseline policies: our optimal policy without predictions and the simple policy that always utilizes the predictions (these correspond to the two "existing approaches" described previously). A snapshot of our results, for the Rossmann stores depicted in Figure 1, is given in Table 2.

More generally, on any given experimental instance (i.e., a time series and a set of predictions), the minimum (maximum) of the costs incurred by these two baselines can be viewed as the best (worst) we can hope for. Thus we measure performance in terms of the proportion of the gap between these two costs incurred by our policy, so if this "optimality gap" is close to 0, then our policy performs almost as good as the better one of the two baselines. Note that *randomly* selecting between the two baseline policies yields an (expected) optimality gap of 0.5. We find that in the Rossmann data set, the average optimality gap is 0.26 when the predictions are accurate and 0.28 when the predictions are inaccurate. In the Wikipedia data set, the average optimality gap is 0.40 when the predictions are accurate and 0.07 when the predictions are inaccurate. In the Restaurant data set, the average optimality gap is 0.10 when the predictions are accurate and 0.39 when the predictions are inaccurate. This demonstrates that our policy performs well, irrespective of the quality of the predictions.

Table 1. Summary of Main Theoretical Results (Each Entry Has a Corresponding Policy That Achieves the Stated Regret, Along with a Matching Lower Bound)

	Without predictions	With predictions of unknown accuracy
Known variation	$\tilde{O}(T^{(3+v)/4})$	$\tilde{O}(T^{\min\{(3+v)/4,a\}})$
Unknown variation	$\tilde{O}(T^{(3+v)/4})$	$\tilde{O}(T^{\max\{(3+v)/4,a\}})$

Table 2. Continuation of Figure 1: Costs Incurred by an Optimal Policy That Makes No Use of Predictions, a Policy That Relies Entirely on Predictions, and Our Policy

	No prediction	Prediction	Our policy
Upper store	\$28,303	\$14,454	\$14,454
Lower store	\$23,460	\$35,600	\$23,899

1.3. Literature Review

The earliest works on the newsvendor model assume that the demand distribution is known (Arrow et al. 1958, Scarf et al. 1960). This has since been relaxed, with the resulting approaches being divided into parametric and nonparametric ones. Among parametric approaches, much work is Bayesian, where a prior distribution is assumed over the parameters. Scarf (1959) applied the Bayesian approach to inventory models, and later this was studied in many works (Karlin 1960, Iglehart 1964, Azoury 1985, Lovejoy 1990). Liyanage and Shanthikumar (2005) introduced another parametric approach called operational statistics, which, unlike the Bayesian approach, does not assume any prior knowledge on the parameter values, instead using past demand observations to directly estimate the optimal ordering quantity.

Nonparametric approaches have been developed in recent years. The first example is the sample average approximation (SAA) method, first proposed by Kleywegt et al. (2002) and Shapiro (2003). Levi et al. (2007) applied SAA to the newsvendor problem, and Levi et al. (2015) improved significantly upon their bounds. Other nonparametric approaches have included stochastic gradient descent algorithms (Burnetas and Smith 2000, Kunnumkal and Topaloglu 2008, Huh and Rusmevichientong 2009) and the concave adaptive value estimation (CAVE) method (Godfrey and Powell 2001, Powell et al. 2004). With the development of machine learning, Ban and Rudin (2019) and Oroojlooyjadid et al. (2020) proposed machine-learning/deep-learning algorithms using demand features and historical data.

All of the above studies treated the newsvendor in a static environment. There are two common approaches to nonstationarity. The first is to model (stochastically) the nonstationarity and utilize past demand observations according to the model. One common way is to model the nonstationarity as a Markov chain. For example, Treharne and Sox (2002) applied this idea to inventory management, and Aviv and Pazgal (2005) and Chen et al. (2019a) applied this idea to revenue management. Another approach is to bound the nonstationarity via a variation budget, which has been applied to stochastic optimization (Besbes et al. 2015), dynamic pricing (Keskin and Zeevi 2017), multiarmed bandit (Besbes et al. 2014), and the newsvendor problem (Keskin et al. 2023), among others. Some of these works are applicable in the sense that our problem can be mapped to their settings (e.g., multiarmed bandit, such as in Besbes et al. 2015, Karnin and Anava 2016, Luo et al. 2018, and Cheung et al. 2022), but these connections do not appear to be fruitful. In particular, the multiarmed bandit papers cited above typically considered a *limited-feedback* setting rather than the *full-feedback* setting explored in this work. Related to feedback, whereas our study provides a complete characterization of the regret behavior for the nonstationary newsvendor problem with *uncensored* demand, practical applications often involve *censored* demand. The nonstationary newsvendor problem under censored demand is an interesting direction for future research.

Beyond the bandit literature, it is worth mentioning recent work on online convex optimization (OCO) with limited nonstationarity. When the level of nonstationarity is *known*, the standard first-order OCO algorithms can be modified with carefully chosen restarts and updating rules (Besbes et al. 2015, Yang et al. 2016, Chen et al. 2019b). There are also recent works that concern *unknown* nonstationarity, such as Zhang et al. (2018), Baby and Wang (2019), Bai et al. (2022), and Huang and Wang (2023). Finally, as mentioned before, Keskin et al. (2023) is particularly relevant, so we delay a careful comparison with Sections 2 and 3.

The second common practice is to use predictions. A recent line of work has looked to help decision-making by incorporating predictions into online optimization problems such as revenue optimization (Munoz and Vassilvitskii 2017, Balseiro et al. 2022, An et al. 2024), caching (Rohatgi 2020, Lykouris and Vassilvitskii 2021), online scheduling (Lattanzi et al. 2020), and the secretary problem (Dütting et al. 2021). In this paper, we combine the nonstationarity framework and the prediction framework on the newsvendor problem.

Finally, most previous works involving algorithms with predictions have analyzed algorithms' performances using competitive analysis (see, e.g., Mahdian et al. 2012, Antoniadis et al. 2020, Balseiro et al. 2022, and Jin and Ma 2022) and obtained optimal *consistency-robustness* trade-offs, where *consistency* is an algorithm's competitive ratio when the prediction is accurate and *robustness* is the competitive ratio regardless of the prediction's accuracy. However, competitive ratio transfers to a regret bound that is linear in T. In contrast, we do regret analysis under this framework and design an algorithm that has near-optimal worst-case regret without knowing the prediction quality. Other papers with regret analyses under the prediction model have included Munoz and Vassilvitskii (2017) (revenue optimization in auctions), Hao et al. (2023) (Thompson sampling), Hu et al. (2024) (constrained online two-stage stochastic optimization), and An et al. (2024) (online resource allocation).

2. Model: The Nonstationary Newsvendor (Without Predictions)

We begin this section with a formal description of the Nonstationary Newsvendor along with a comparison with the problem of the same name from Keskin et al. (2023). Consider a sequence of newsvendor problems over T time periods labeled $t = 1, \ldots, T$. At the beginning of each time period t, the decision-maker selects a quantity $q_t \in Q$, where Q is a fixed subset of $\mathbb{R}^+$ bounded above by a quantity we denote as $Q_{\max}$.[2] Then the period's demand d_t is drawn from an (unknown) demand distribution D_t, which depends on the time period t. These demand distributions are independent over time. Finally a cost is incurred; specifically, there is a (known) per-unit underage cost $b_t \in [0, b_{\max}]$ and a (known) per-unit overage cost $h_t \in [0, h_{\max}]$ so that the total cost is equal to

$$b_t(d_t - q_t)^+ + h_t(q_t - d_t)^+,$$

where $x^+ = \max\{0, x\}$. The decision-maker observes the realized demand d_t[3] and thus the cost. Note that requiring $q_t \in Q$ does not impose any restriction on *modeling*, because Q could simply be selected to be $\mathbb{R}$ (as in much of the literature). In fact, introducing Q allows for modeling important practical concerns such as batched inventory or even simply the integrality of physical items. As we will discuss momentarily, this is a nontrivial concern insofar as theoretical guarantees are concerned.

To complete our description of the Nonstationary Newsvendor, we will need to (a) impose a few assumptions on the demand distributions and then (b) describe how "nonstationarity" is quantified. These are, respectively, the subjects of the following two subsections.

2.1. Demand Distributions

We will assume that the demand distributions come from a known parameterized family of distributions $\mathcal{D}$.

Assumption 1. *Every demand distribution D_t comes from a family of distributions $\mathcal{D}$ satisfying the following:*

(a) $\mathcal{D} = \{\mathcal{D}_\mu : \mu \in [\mu_{\min}, \mu_{\max}]\}$; *that is, $\mathcal{D}$ is parameterized by a scalar μ taking values in some bounded interval.*

(b) *Each distribution $\mathcal{D}_\mu \in \mathcal{D}$ is sub-Gaussian.*[4]

Assumption 1 is fairly minimal. Parsing it in reverse, the sub-Gaussianity in part (b) allows for many commonly-used variables, such as the Gaussian distribution and any bounded random variable, while letting us eventually apply Hoeffding-type concentration bounds. Part (a) is particularly minimal at the moment, because μ represents an arbitrary parameterization of $\mathcal{D}$, but will become meaningful when combined with Assumption 2. The choice of the symbol "μ" might suggest that μ represents the mean of $\mathcal{D}_\mu$, and indeed this is what we will assume from here on. But it should be emphasized that our taking $\mu = \mathbb{E}[\mathcal{D}_\mu]$ is strictly for notational convenience (because we will frequently need to refer to the means of these distributions); if μ were any other parameterization of $\mathcal{D}$, we could simply define a mapping from μ to the mean values.

Now define $C(\mu, b, h, q)$ to be the expected newsvendor cost when selecting quantity $q \in Q$, given underage/overage costs b and h, and demand distribution $\mathcal{D}_\mu$:

$$C(\mu, b, h, q) = \mathbb{E}_{d\sim\mathcal{D}_\mu}[b(d-q)^+ + h(q-d)^+].$$

The critical assumption, with respect to the parameterization in Assumption 1(a), is that the expected cost is well behaved as a function of μ.

Assumption 2. *For every $b \in [0, b_{\max}]$, $h \in [0, h_{\max}]$, and $q \in Q$, the function $C(\cdot, b, h, q)$ is Lipschitz on its domain $[\mu_{\min}, \mu_{\max}]$; that is, there exists $\ell \in \mathbb{R}^+$ such that for every $\mu_1, \mu_2 \in [\mu_{\min}, \mu_{\max}]$, we have*

$$|C(\mu_1, b, h, q) - C(\mu_2, b, h, q)| \le \ell|\mu_1 - \mu_2|.$$

Note that in the above description, the Lipschitz constant ℓ may depend on b, h, and q, but by continuity, there exists a single ℓ so that the above holds for all b, h, and q simultaneously.

Some useful examples of families $\mathcal{D}$ satisfying Assumptions 1 and 2 are the following:[5]

1. $\mathcal{D}_\mu \sim \mathcal{N}(\mu, \sigma^2)$, the family of normal distributions with fixed variance σ^2. In this case, $\ell = O(\sigma(b_{\max} + h_{\max}))$. A relaxation is that the variances may vary (continuously) with μ.
2. $\mathcal{D}_\mu = \mu + \epsilon$, where ϵ is any mean-zero, sub-Gaussian variable.
3. The Poisson distribution is frequently used to model demand (because arrivals are often modeled as a Poisson process). Although the Poisson distribution is *not* sub-Gaussian, any reasonable truncation satisfies our assumptions. For example, $\mathcal{D}_\mu \sim \min\{\text{Poisson}(\mu), K\mu\}$ for some constant K. Here, K can be taken to be large enough so that the truncation happens with small probability (in fact, this probability is $O(e^{-K\mu})$).

To understand the reasoning behind Assumption 2, consider the problem faced at some time t. The optimal choice for the decision-maker here is

$$q_t^* \in \arg\min_{q\in Q} \; C_t(\mu_t, q), \tag{1}$$

where μ_t is the mean of D_t (i.e., $D_t \sim \mathcal{D}_{\mu_t}$), and $C_t(\mu, q) = C(\mu, b_t, h_t, q)$ to simplify the notation.[6] Because D_t is unknown, it is likely that some $q_t \ne q_t^*$ will ultimately be selected, and we could measure the suboptimality of this decision (i.e., regret, to be defined soon): $C_t(\mu_t, q_t) - C_t(\mu_t, q_t^*)$. It would be natural then to try to characterize this suboptimality as a function of $|q_t - q_t^*|$, but in fact all of the algorithms we will consider "work" by making an estimate $\hat{\mu}_t$ of μ_t and then selecting $\hat{q}_t \in \arg\min_{q\in Q} C_t(\hat{\mu}_t, q)$. So motivated, the purpose of Assumption 2 is to allow us to "translate" error in our estimate of μ_t to (excess) costs. The following structural lemma makes this precise and will be used throughout the paper.

Lemma 1. *Fix any b and h (we will suppress them from the notation). For any $D_{\mu_1}, D_{\mu_2} \in \mathcal{D}$, let $q_1^* \in \arg\min_{q\in Q} C(\mu_1, q)$ and $q_2^* \in \arg\min_{q\in Q} C(\mu_2, q)$. Then, we have*

$$C(\mu_1, q_2^*) - C(\mu_1, q_1^*) \le 2\ell|\mu_1 - \mu_2|.$$

Lemma 1 states that estimation error of the mean μ_t translates *linearly* to excess cost. The proof of Lemma 1 appears in Online Appendix A.

Aside: Comparison with Keskin et al. (2023). The final component in describing the Nonstationary Newsvendor is defining a proper quantification of nonstationarity. Before doing so, we delineate the *modeling* differences between our Nonstationary Newsvendor and that of Keskin et al. (2023). There are two primary differences.

1. The demand distributions D_t in Keskin et al. (2023) are assumed to be of the form $D_t = \mu_t + \epsilon_t$, where μ_t is the mean of D_t that drifts across time and ϵ_t is the noise distribution that is i.i.d., continuous, and bounded. Effectively, the demand distributions fall into a *nonparametric* family of distributions with the same "shape." In contrast, our demand distributions fall into a *parametric* family of distributions, albeit not necessarily of the same "shape."
2. Our set of allowed order quantities Q is bounded but otherwise arbitrary. In particular, it need not contain the optimal unconstrained order quantity $\arg\min_{q\in\mathbb{R}} C(\mu, q)$ for each μ (or any μ, for that matter). Keskin et al. (2023) assumed that $Q = \mathbb{R}^+$.[7]

Besides the practical reasons why discrete quantities arise in practice (nondivisible items, batched inventory, etc.), the primary consequence of either of the two differences above is that they preclude a critical lemma used in Keskin et al. (2023) (and in fact by Levi et al. 2007) that states that $C(\mu_1, q_2^*) - C(\mu_1, q_1^*)$ (as defined in our Lemma 1) scales as $(q_1^* - q_2^*)^2$. This scaling does not necessarily hold when either the demand distribution or Q is discrete. These relaxations in assumptions yield different lower bounds in the worst-case regret from Keskin et al. (2023), which we will discuss in detail later.

2.2. Demand Variation

Just as in Keskin et al. (2023) (and Keskin and Zeevi 2017 before that), we measure the level of nonstationarity via a deterministic analog of quadratic variation for the sequence of means $\mu_1, \ldots, \mu_T$. Specifically, define a *partition* of the time horizon $\{1, \ldots, T\}$ to be any subset of time periods $\{t_0, \ldots, t_K\}$, where $1 \le t_0 < \cdots < t_K \le T$. Here, the subset can have any size between 1 and T, that is, $0 \le K \le T-1$. Then, for any sequence of means $\boldsymbol{\mu} = \{\mu_1, \ldots, \mu_T\}$, its demand variation is

$$V_{\boldsymbol{\mu}} = \max_{0 \le K \le T-1} \max_{\{t_0, \ldots, t_K\} \in \mathcal{P}} \left\{ \sum_{k=1}^{K} |\mu_{t_k} - \mu_{t_{k-1}}|^2 \right\}, \tag{2}$$

where $\mathcal{P}$ is the set of all partitions.

To motivate the use of partitions in the definition of $V_{\boldsymbol{\mu}}$, it is worth contrasting with a measure that may feel more natural, namely, the sum of squared differences (*SSD*) between consecutive terms, $\sum_{t=2}^{T} (\mu_t - \mu_{t-1})^2$, which corresponds to taking the densest possible partition $\{1, 2, \ldots, T\}$. The maximum in the definition of $V_{\boldsymbol{\mu}}$ is not necessarily achieved by selecting the densest possible partition but rather by setting $t_0, \ldots, t_K$ to be the periods when the sequence $\mu_1, \ldots, \mu_T$ changes direction. Thus, the demand variation penalizes *trends*, or consecutive increases/decreases, more so than the SSD. For example, the mean sequences $\boldsymbol{\mu_1} = \{1, 2, 3, 4, 5\}$ and $\boldsymbol{\mu_2} = \{1, 0, 1, 0, 1\}$, respective variations $V_{\boldsymbol{\mu_1}} = (5-1)^2 = 16$ and $V_{\boldsymbol{\mu_2}} = 1^2 + 1^2 + 1^2 + 1^2 + 1^2 = 5$, despite having identical SSDs.

All of our theoretical guarantees (upper and lower bounds) will be parameterized by $V_{\boldsymbol{\mu}}$. This quantity of course depends on T, and so it is natural to allow $V_{\boldsymbol{\mu}}$ to grow T. It will turn out that the most natural parameterization of this growth is via what we will simply call the variation parameter $v \in [0, 1]$, such that $V_{\boldsymbol{\mu}} = BT^v$, where B is some constant (which we take to be equal to 1 from here on). We denote the set of demand distribution sequences $\{D_1, \ldots D_T\}$ whose means $\boldsymbol{\mu} = \{\mu_1, \ldots, \mu_T\}$ satisfy $V_{\boldsymbol{\mu}} \le T^v$ as

$$\mathcal{D}(v) = \{\{D_1, \ldots D_T\} : D_t \in \mathcal{D} \text{ for all } t \text{ and } V_{\boldsymbol{\mu}} \le T^v\}.$$

In the next section, we will show via a minimax lower bound that nontrivial guarantees are achievable only when $v < 1$ and provide an algorithm that achieves the same bound.

Aside: Time-Series Modeling. At this point, we have fully described our model for the Nonstationary Newsvendor. All that remains is to define our performance metric, which we will do in the next subsection. We conclude this subsection with an important practical consideration with respect to time series models and our variation parameter.

Consider, as an example, the following class of time series models:

$$d_t = R(t) + S(t) + \epsilon_t. \tag{3}$$

Here, $R(t)$ represents a deterministic (and usually simple, e.g., linear) function representing some notion of "trend," and $S(t)$ represents a deterministic, periodic function representing some notion of "seasonality." Finally, all stochastic behavior is captured by the random variables ϵ_t, which are assumed to be independent and mean-zero. This time series model is classic and yet drives forecasting algorithms (e.g., exponential smoothing) that are still competitive in modern forecasting competitions (Makridakis and Hibon 2000).

The above model raises an important practical issue; if there exists any (nontrivial) trend $R(\cdot)$ or seasonality $S(\cdot)$, then the demand variation of the sequence of means that $\mu_t = R(t) + S(t)$ would scale at least as T, meaning that $v = 1$, and no meaningful guarantee will be achievable. Our main observation is that time series effects like trend and seasonality are easily detected and estimated so that in any practical setting, estimates $\hat{R}(\cdot)$ and $\hat{S}(\cdot)$ should be available and used to "detrend" and "de-seasonalize" the data. Concretely, the Nonstationary Newsvendor would take place on the sequence

$$\tilde{d}_t = d(t) - \hat{R}(t) - \hat{S}(t) = (R(t) - \hat{R}(t)) + (S(t) - \hat{S}(t)) + \epsilon_t.$$

The resulting sequence of means $\mu_t = (R(t) - \hat{R}(t)) + (S(t) - \hat{S}(t))$ does not stem from the trend and seasonality but rather the error in estimating the trend and seasonality. It is this error that is assumed to be nonstationary but with reasonable variation parameter.

2.3. Performance Metric: Regret

We conclude this section by formally defining our performance metric for any policy. A policy is simply a sequence of mappings $\pi = \{\pi_1, \ldots, \pi_T\}$, where each π_t is a mapping from $d_1, \ldots, d_{t-1}$ to an order quantity $q_t \in Q$ at time t (by convention, π_1 is a constant function).[8] We measure the performance of a policy by its regret. Fix a sequence of demand distributions $\boldsymbol{D} = \{D_1, \ldots, D_T\}$. Following the earlier notation from (1), the regret incurred by a policy that selects order quantities $q_1, \ldots, q_T$ is

$$\mathbb{E}_{\boldsymbol{D}}^{\pi}\left[\sum_{t=1}^{T} (C_t(\mu_t, q_t) - C_t(\mu_t, q_t^*))\right],$$

where the expectation is with respect to the randomness of the realized demands. Recall that the demand distributions are independent, so q_t^* as defined in (1) depends only on D_t. In words, the regret measures the difference between the (expected) total cost incurred by the policy and that of a clairvoyant that knows the underlying demand distributions $\boldsymbol{D} = \{D_1, \ldots, D_T\}$.[9]

We will be concerned with the worst-case regret of a policy across families of instances (i.e., sequences of demand distributions) controlled by the variation parameter v:

$$\mathcal{R}^{\pi}(T) = \sup_{\boldsymbol{D} \in \mathcal{D}(v)} \mathbb{E}_{\boldsymbol{D}}^{\pi}\left[\sum_{t=1}^{T} (C_t(\mu_t, q_t) - C_t(\mu_t, q_t^*))\right].$$

Note that if the worst-case regret $\mathcal{R}^\pi(T)$ of some policy is sublinear in T, then that policy is essentially cost-optimal on average as T goes to infinity. In the next section, we will prove a lower bound on the achievable across all policies and describe an algorithm that achieves this lower bound.

3. Solution to the Nonstationary Newsvendor (Without Predictions)

This section contains a complete solution (i.e., matching lower and upper bounds on regret) to the Nonstationary Newsvendor. We begin with the lower bound.

Proposition 1 (Lower Bound: Nonstationary Newsvendor). *For any variation parameter $v \in [0,1]$ and any policy π (which may depend on the knowledge of v), we have*

$$\mathcal{R}^\pi(T) \geq cT^{(3+v)/4},$$

where $c > 0$ is a universal constant.

Proposition 1 is a corollary of a more general lower bound (Proposition 2 in the next section); it will turn out that the Nonstationary Newsvendor is a special case of the Nonstationary Newsvendor with predictions, so the proof is omitted. Proposition 1 states that the regret of any policy is at least $\Omega(T^{(3+v)/4})$. It is useful to contrast this with two existing results.

1. **Stationary Newsvendor:** In the special case of i.i.d. demand, it is known that the optimal achievable regret is $\Theta(T^{1/2})$. Example 1 in Besbes and Muharremoglu (2013) demonstrates the lower bound, and the SAA method of Levi et al. (2007, 2015) achieves the upper bound. This point might appear to be incompatible with our result, which states a lower bound of $\Omega(T^{3/4})$ when $v = 0$, but in fact the case of $v = 0$ is more general than i.i.d. demand because it allows $O(T^0) = O(1)$ demand variation, whereas i.i.d. demand amounts to zero demand variation. Indeed, our proof of Proposition 1, for the special case of $v = 0$, utilizes instances for which the demand distribution is allowed to change $T^{1/2}$ times (by an amount of $T^{-1/4}$, resulting in $O(1)$ variation).

As an aside, this discussion raises a natural question. Is the disconnect here between stationary (i.i.d.) demand and variation parameter $v = 0$ a consequence of our use *quadratic* variation, and would the same disconnect arise for other measures of demand variation? In Online Appendix E, we answer both questions in the affirmative by showing that if the exponent 2 in the demand variation (Equation (2)) is instead some $\theta \geq 0$, then Proposition 1 generalizes to a lower bound of $\Omega(T^{(1+\theta+v)/(2+\theta)})$. Thus, for any $\theta > 0$, the case of variation parameter $v = 0$ is meaningfully more general than stationary demand.

2. **Continuous Newsvendor:** A similar "story" plays out in the setting of Keskin et al. (2023), which recall (among other key differences with our model, as described in Section 2.1) requires the additional assumption that both the demand distributions and the possible order quantities be continuous. Keskin et al. (2023) showed an optimal achievable regret of $\Theta(T^{(1+v)/2})$, which could be contrasted with the stationary (i.i.d.) setting for which an $\Omega(\log T)$ lower bound exists (Besbes and Muharremoglu 2013). Table 3 summarizes these lower bounds.

In the next two subsections, we will first analyze a simple algorithm that achieves the lower bound of Proposition 1 when the variation parameter v is known and then use this as a building block for an algorithm that achieves the same bound when v is unknown.

3.1. Upper Bound with Known Variation Parameter *v*

If we assume that v is known, then designing a policy that achieves regret matching Proposition 1 is fairly straightforward. In fact, a simple policy based on averaging a fixed number of past demand observations does the job (Keskin et al. 2023 used the same policy). That policy, which we call the Fixed-Time-Window Policy, is defined in Algorithm 1.

Algorithm 1 (Fixed-Time-Window Policy)

Inputs: variation parameter $v \in [0,1]$ and scaling constant $\kappa > 0$
Initialization: $n \leftarrow \lceil \kappa T^{(1-v)/2} \rceil$
for $t = 1, \ldots, n$ **do**
  select $q_t \in Q$ arbitrarily;
for $t = n+1, \ldots, T$ **do**
  $\hat{\mu}_t \leftarrow \frac{1}{n}\sum_{s=t-n}^{t-1} d_s$ (if $\hat{\mu}_t \notin [\mu_{\min}, \mu_{\max}]$, round $\hat{\mu}_t$ to the nearest value in $[\mu_{\min}, \mu_{\max}]$);
  $q_t \leftarrow \arg\min_{q \in Q} C_t(\hat{\mu}_t, q)$.

The Fixed-Time-Window Policy uses a carefully selected "window" size n that is on the order of $T^{(1-v)/2}$. At each time period t, it constructs an estimate $\hat{\mu}_t$ of the mean by averaging the observed demands from the previous n periods and then selects the optimal order quantity corresponding to $\hat{\mu}_t$. Note that Algorithm 1 also includes a "scaling constant" κ; this should be thought of as a practical tuning parameter, but for the coming theoretical result, it can be chosen arbitrarily (e.g., $\kappa = 1$ suffices).

The following result bounds the worst-case regret of the Fixed-Time-Window Policy.

Lemma 2 (Upper Bound: Nonstationary Newsvendor with Known v). *Fix any variation parameter $v \in [0,1]$. The Fixed-*

Table 3. Summary of Newsvendor Lower Bounds Under Different Settings

	Continuous	General
Stationary (i.i.d.)	$\log T$	$T^{1/2}$
Nonstationary	$T^{(1+v)/2}$	$T^{(3+v)/4}$

Time-Window Policy π^{fixed} *achieves worst-case regret*

$$\mathcal{R}^{\pi^{\text{fixed}}}(T) \le CT^{(3+v)/4},$$

where $C \le 3\max\{b_{\max}, h_{\max}\}(\delta + Q_{\max}) + \ell\left(2\sqrt{\kappa} + \sqrt{\frac{\pi}{36e}}\frac{\delta}{\sqrt{\kappa}}\right)$, *and* $\delta = \sup_{\mathcal{D}_\mu \in \mathcal{D}} \|\mathcal{D}_\mu\|_{\psi_2}$.

As promised, Lemma 2 shows that the Fixed-Time-Window Policy achieves regret that matches the lower bound in Proposition 1. Its proof can be found in Online Appendix B and amounts to bounding the estimation error incurred by demand noise (which is worse for smaller time windows) and demand mean variation (which is worse for larger time windows). The exact time window used in the policy comes from balancing these two sources of error.

3.2. Upper Bound with Unknown Variation Parameter *v*

The lower bound in Proposition 1 holds for policies that "know" v. Naturally, it also holds for policies that do not know v, but an unanswered question at the moment is whether (a) the lower bound should be even larger when v is unknown or (b) there exists a policy that matches Proposition 1 without knowing v. We show here that case (b) holds by constructing such a policy.[10] Our policy, which we call the Shrinking-Time-Window Policy (Algorithm 2), at a high level uses the Fixed-Time-Window Policy with the smallest variation parameter that is consistent with the demand observed so far. In more detail:

1. It begins with a discrete set of candidate variation parameters $\mathcal{V} = \{v_1, \ldots, v_k\}$:

$$v_j = \left(1 + \frac{1}{\log T}\right)^{j-1} \frac{1}{\log T}, \quad j = 1, \ldots, k, \tag{4}$$

where k is chosen so that $v_{k-1} < 1 \le v_k$. $\mathcal{V}$ is defined specifically so that the variation parameters are increasing ($v_{j-1} < v_j$) and so that it discretizes the interval $[0,1]$ at a sufficiently fine granularity.

2. At any time period t, there is a "current" candidate parameter v_i (initialized to be v_1 at $t = 1$) that is assumed to be the true variation v, and so the corresponding Fixed-Time-Window Policy is applied; a time window of

$$n_i = \lceil \kappa T^{(1-v_i)/2} \rceil \tag{5}$$

is used, and an estimate of μ_t is made:

$$\hat{\mu}_t^i = \frac{1}{n_i} \sum_{s=t-n_i}^{t-1} d_s, \quad \text{rounded to the nearest value in } [\mu_{\min}, \mu_{\max}]. \tag{6}$$

3. The index i of the "current" candidate parameter v_i is incremented at any period in which the policy gathers sufficient evidence that $v_i < v$. This is possible because of the following observation; if $v_i \approx v$, then by Lemma 2 we have that for any $v_j > v_i$, the regret incurred by the Fixed-Time-Window Policy corresponding to v_j is $O(T^{(3+v_j)/4})$, and thus the *cumulative difference* between the estimated mean demands ($|\hat{\mu}_t^i - \hat{\mu}_t^j|$) cannot exceed $O(T^{(3+v_j)/4})$. Thus, if this is observed for some $v_j > v_i$, then we can conclude that $v_i < v$, and i is incremented.

Algorithm 2 (Shrinking-Time-Window Policy)

Inputs: scaling constants $\kappa > 0$, and γ sufficiently large (Equation (8) in Online Appendix C);
Initialization: Set $\mathcal{V} = \{v_1, \ldots, v_k\}$ and $\{n_1, \ldots, n_k\}$ according to Equations (4) and (5);
for $t = 1, \ldots, T^{3/4}$ **do**
  select $q_t \in Q$ arbitrarily;
Initialize $i \leftarrow 1$ and $t_{\text{if}} \leftarrow T^{3/4} + 1$;
for $t = T^{3/4} + 1, \ldots, T$ **do**
  if $\sum_{s=t_{\text{if}}}^{t} |\hat{\mu}_s^i - \hat{\mu}_s^j| \ge 2(\gamma\sqrt{\log T} + \sqrt{\kappa})T^{(3+v_j)/4}$ *for some* $j > i$ **then**
    $i \leftarrow i + 1$;
    $t_{\text{if}} \leftarrow t$;
  $q_t \leftarrow \arg\min_{q \in Q} C_t(\hat{\mu}_t^i, q)$.

This policy's regret matches (up to log factors) the lower bound in Proposition 1:

Theorem 1 (Upper Bound: Nonstationary Newsvendor with Unknown ***v***). *For any variation parameter* $v \in [0,1]$, *the Shrinking-Time-Window Policy* $\pi^{\text{shrinking}}$ *achieves worst-case regret*

$$\mathcal{R}^{\pi^{\text{shrinking}}}(T) \le CT^{(3+v)/4}\log^{5/2} T,$$

where $C \le 3\max\{b_{\max}, h_{\max}\}(\delta + Q_{\max}) + 12e^{1/4}\ell(\gamma + \sqrt{\kappa}) + e^{1/4}C_{\text{Lemma 2}}$, *and* $C_{\text{Lemma 2}}$ *is the constant in Lemma 2.*

The proof of Theorem 1 can be found in Online Appendix C.

This concludes our discussion of the Nonstationary Newsvendor. In the next section, we turn to the second subject of this paper, which is the same problem with predictions.

4. The Nonstationary Newsvendor with Predictions

As described in the Introduction, it is likely that when the Nonstationary Newsvendor is faced with practice, some notion of a "prediction" of future demand will be made. Such predictions can come from a diverse set of sources ranging from simple human judgment to forecasting algorithms built on previous demand data to more sophisticated machine-learning algorithms trained on feature information. The process of sourcing or constructing such predictions is orthogonal to our work. Instead, we treat these predictions as given to us endogenously (and in particular, we make no assumption on the accuracy of these predictions) and attempt to use these predictions optimally.

4.1. Model

The **Nonstationary Newsvendor with Predictions** problem assumes all of the setup, assumptions, and notation of the previous Nonstationary Newsvendor problem. In addition, at each time period t, we assume that the decision-maker receives a prediction a_t before selecting an order quantity $q_t \in Q$.[11] This prediction is meant to be an estimate of μ_t, and so we measure the prediction error of a sequence $\boldsymbol{a} = \{a_1, \ldots, a_T\}$ with respect to a sequence of means simply as

$$\sum_{t=1}^{T} |a_t - \mu_t|.$$

Note that, unlike demand variation, we have not used partitions here (and in fact, introducing partitions would not have any effect because we are measuring *absolute* rather than squared differences). Intuitively, we do not want to require the sequence of errors to be meaningful time series; the predictions are generic, and their accuracy is allowed to change rapidly. Just as for the demand variation, the prediction error is expected to grow with the time horizon T, and the proper parameterization of this growth is via an exponent; we call the accuracy parameter the smallest $a \in [0,1]$ such that the prediction error is at most T^a. We will always assume that a is unknown to the decision-maker.

Algorithm 3 (Prediction Policy)
for $t = 1, \ldots, T$ **do**
  $\hat{\mu}_t \leftarrow a_t$ (if $\hat{\mu}_t \notin [\mu_{\min}, \mu_{\max}]$, round $\hat{\mu}_t$ to the nearest value in $[\mu_{\min}, \mu_{\max}]$);
  $q_t \leftarrow \arg\min_{q \in Q} C_t(\hat{\mu}_t, q)$.

Naturally, the notion of a policy π expands to include the predictions $\pi = \{\pi_1, \ldots, \pi_T\}$, where each π_t is a mapping from $d_1, \ldots, d_{t-1}$ and $a_1, \ldots, a_t$ to an order quantity $q_t \in Q$. The simplest policy, which "should" be used if the prediction error is known to be sufficiently small, is to simply behave as if the predictions were perfect. We call this the Prediction Policy (Algorithm 3). The following observation collects a few (likely unsurprising) facts about the performance of this policy with respect to worst-case regret (generalized in the "obvious" manner to incorporate prediction accuracy via the accuracy parameter a):

Observation 1 (Upper and Lower Bounds: Prediction Policy). Fix any variation parameter $v \in [0,1]$ and any accuracy parameter $a \in [0,1]$.

a) The Prediction Policy $\pi^{\text{prediction}}$ achieves worst-case regret

$$\mathcal{R}^{\pi^{\text{prediction}}}(T) \leq CT^a,$$

where $C \leq 2\ell$.

b) For any policy π (which may depend on the knowledge of a) that is solely a function of the predictions (i.e., does not depend on the observed demands), we have

$$\mathcal{R}^{\pi}(T) \geq cT^a,$$

where $c > 0$ is a universal constant.

Observation 1a states that the Prediction Policy translates prediction error directly to regret (incidentally, it does this without "knowing" a). There are, of course, other ways in which the predictions could be used, but Observation 1b essentially states that there is nothing to be gained by doing so (even if a is known). The proof of Observation 1a appears in Online Appendix A. Observation 1b is a direct corollary of Proposition 2, which is given in the next subsection.

4.2. Extreme Cases

What exactly is achievable for the Nonstationary Newsvendor with Predictions depends heavily on whether v and a are known to the policy. To see this, it is worth first considering the two extremes.

Case 1: Known v and a.

A simple policy is available when v and a are both known. Compare the quantities $(3+v)/4$ and a. If the former quantity is smaller, apply the Fixed-Time-Window Policy. If the latter is smaller, apply the Prediction Policy. Lemma 2 and Observation 1 together imply that this achieves a worst-case regret of $O(T^{\min\{(3+v)/4, a\}})$. This is optimal, as demonstrated by the following result.

Proposition 2 (Lower Bound: Known v and a). *Fix any variation parameter $v \in [0,1]$ and any accuracy parameter $a \in [0,1]$. For any policy π (which may depend on the knowledge of v and a), we have*

$$\mathcal{R}^{\pi}(T) \geq cT^{\min\{(3+v)/4, a\}},$$

where $c > 0$ is a universal constant.

The proof of this result can be found in Online Appendix D and relies on an explicit construction of a family of problem instances. Our construction breaks the total time horizon into cycles wherein the demand distribution is i.i.d. We tune the length of each cycle to be small enough so that it is (provably) hard to detect the change in demand distributions, and the predictions are essentially useless for most time periods in the cycle and large enough so that the demand variation is within T^v and the prediction error is within T^a.

Case 2: Unknown v and a.

At the opposite extreme, if v and a are both unknown, is it still possible to achieve $O(T^{\min\{(3+v)/4, a\}})$ worst-case regret? The answer is no.

Proposition 3 (Lower Bound: Unknown v and a). *For any policy that does not depend on the knowledge of v or a, there exists a problem instance such that $a \neq (3+v)/4$, and the policy incurs regret at least $cT^{\max\{(3+v)/4, a\}}$ on the instance, where $c > 0$ is a universal constant.*

Proposition 3 states that the best we can hope for, when v and a are unknown, is a worst-case regret of at least $\Omega(T^{\max\{(3+v)/4, a\}})$.[12] Note that Proposition 3 shows that there exists a pair of v and a, and a corresponding problem instance, such that this lower bound holds.

This is in contrast to a result showing that for any pair of v and a, there exists a problem instance such that the lower bound holds, as is common in the literature (see, e.g., theorem 1 of Keskin and Zeevi 2017). This lower bound is easily achieved, for example, by applying the Shrinking-Time-Window Policy or the Prediction Policy (or any blind randomization of the two). The proof of Proposition 3 is in Online Appendix F. In contrast to Case 1, the lower bound construction here relies heavily on the fact that we do not know which one of $(3+v)/4$ and a is smaller.

4.3. Final Solution

We have finally reached the problem that motivates this entire paper: designing an optimal policy for the Nonstationary Newsvendor with Predictions when the prediction error a is unknown. We will assume that v is known, because when v is unknown, Proposition 3 rules out the possibility of using the predictions to improve on what is already achievable without predictions. On the other hand, by Proposition 2, the absolute best we could hope for is a policy that achieves a worst-case regret of $O(T^{\min\{(3+v)/4,a\}})$. In words, we would like a policy that, without knowing a, achieves the same regret had a been known. Our main result is the design of such a policy.

Our policy is called the Prediction-Error-Robust Policy (PERP) and is given in Algorithm 4. PERP utilizes the Fixed-Time-Window policy π^{fixed} in Section 3 as an estimate of the true mean to track the quality of the predictions over time.

Algorithm 4 (Prediction-Error-Robust Policy (PERP))

Inputs: variation parameter $v \in [0,1]$ and scaling constants $\kappa > 0$, γ sufficiently large (Equation (9) in Online Appendix C).

Initialization: $n \leftarrow \lceil \kappa T^{(1-v)/2} \rceil$;

for $t = 1, \dots, n$ **do**
  $\pi_t \leftarrow \pi_t^{\text{prediction}}$;

for $t = n+1, \dots, T$ **do**
  $\hat{\mu}_t^{\text{fixed}} \leftarrow \frac{1}{n}\sum_{s=t-n}^{t-1} d_s$;
  $\hat{\mu}_t^{a} \leftarrow a_t$;
  if $\sum_{s=n+1}^{t} |\hat{\mu}_s^{a} - \hat{\mu}_s^{\text{fixed}}| \geq (\gamma\sqrt{\log T} + \sqrt{\kappa} + 1) \cdot T^{(3+v)/4}$, **then**
    $\pi_t \leftarrow \pi_t^{\text{fixed}}$;
    break
  else
    $\pi_t \leftarrow \pi_t^{\text{prediction}}$.

Theorem 2 (Upper Bound: Known v and Unknown a). *For any variation parameter $v \in [0,1]$ and any accuracy parameter $a \in [0,1]$, the Prediction-Error-Robust Policy π^{PERP} achieves worst-case regret,*

$$\mathcal{R}^{\pi^{\text{PERP}}}(T) \leq \min\{3C_{\text{Lemma 2}}\sqrt{\log T} \cdot T^{(3+v)/4}, 2\ell \cdot T^{a}\},$$

where C_{Lemma2} is the constant in Lemma 2 (and 2ℓ matches the constant in Observation 1a).

The intuition behind PERP is to follow the predictions until a time that is late enough to have evidence that the prediction quality is bad (compared with the Fixed-Time-Window Policy) but early enough to not incur much regret caused by the poor quality of the predictions. Because we do not observe the true past mean μ_t after time period t, we naturally use $\hat{\mu}_t^{\text{fixed}}$ from the Fixed-Time-Window policy π^{fixed} as an estimation of μ_t and in turn keep tracking the cumulative difference in the prediction quality $|a_t - \mu_t|$. We carefully choose the parameters in π^{PERP} so that this estimation is not accurate only with a small probability, and we can identify that the prediction quality is bad if this cumulative difference is too large. By Proposition 2, any policy can achieve only worst-case regret on the order of $T^{\min\{(3+v)/4,a\}}$, so PERP is order optimal.

Aside: Unknown *v* and Known *a*. There are four possible scenarios depending on the knowledge of v and a: known/unknown v and known/unknown a. So far, we have discussed three of them: known v and a (Proposition 2), unknown v and a (Proposition 3), and known v and unknown a (Theorem 2). For the sake of completeness, we discuss the remaining case of unknown v and known a in Online Appendix H, where we give a policy that achieves worst-case regret $\tilde{O}(T^{\min\{(3+v)/4,a\}})$. This is order optimal by Theorem 2.

5. Experiments

Finally, we describe a set of experiments we performed to evaluate our policy (`PERP`) for the Nonstationary Newsvendor with Predictions. In all of our experiments, we compared `PERP` against the Shrinking-Time-Window Policy (`NO-PRED`) and the Prediction Policy (`PURE-PRED`). The main takeaways are as follows:

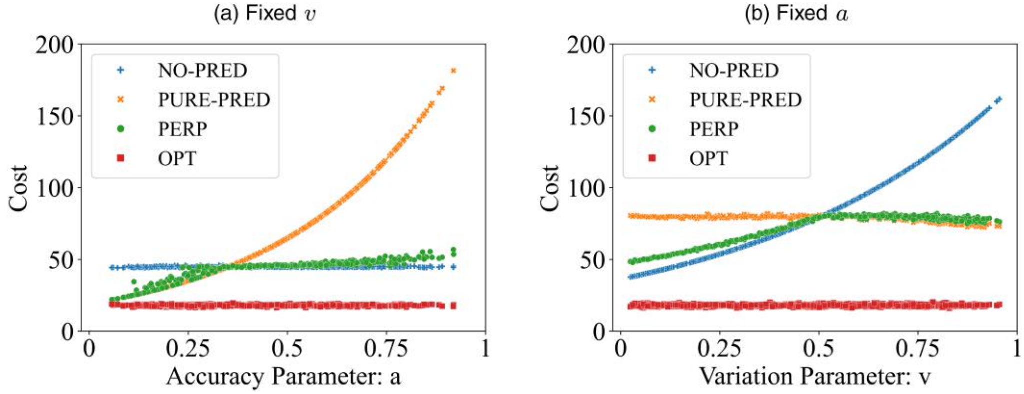
1. `PERP`'s performance is robust with respect to the quality of the predictions without knowing the prediction quality beforehand. Specifically, the (newsvendor) cost it incurs is consistently "close" to the lower of the costs incurred by `NO-PRED` and `PURE-PRED`.

2. `PERP` performs especially well when the absolute difference between the costs of `NO-PRED` and `PURE-PRED` is large, that is, when the "stakes" are highest.

5.1. Experiments on Synthetic Data

The objective of our first batch of experiments was to fix one of the two theoretical parameters (v or a) and test `PERP`'s performance as the other parameters change. To generate demand sequences, we used the parametric time series model that corresponds to triple exponential smoothing (Holt Winters), a classic model for time series data in the family of Equation (3). We give the exact formulas and our choices of parameters in Online Appendix I; for more on triple exponential

Figure 2. (Color online) The Costs of NO-PRED, PURE-PRED, PERP, and OPT When (a) the Variation Parameter v is Fixed and (b) the Accuracy Parameter a is Fixed



Note. Each dot represents the cost of the corresponding policy on a given instance.

smoothing, see Winters (1960). In our experiment, each demand sequence consisted of the demands for the next 365 time periods, with the realized demands generated as Poisson variables with the corresponding means. We ran two sets of experiments:

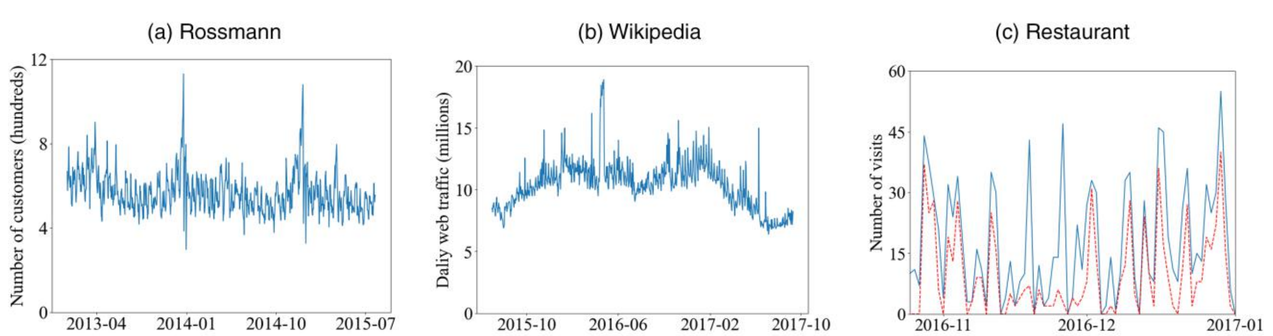
- **Fixed v:** We fixed a single set of parameters for the demand sequence and generated 1,000 different predictions of this demand sequence, each from a set of "predicted" parameters with different accuracy. Thus the variation parameter v was fixed, and the accuracy parameter a varied across instances.
- **Fixed a:** We generated 1,000 demand sequences by selecting the parameters randomly. We then generated predictions by changing each parameter 10% and using the corresponding sequence. Thus the variation parameter v varied across instances, but the accuracy parameter a was (roughly) fixed.

For each demand sequence and corresponding prediction, we ran NO-PRED, PURE-PRED, and PERP with equal overage and underage costs and scaling constants $\kappa = \gamma = 1$. Because the experiment was synthetic, the true underlying demand distribution was known at each time period. Therefore we also ran OPT as a benchmark, which simply ordered the optimal quantile at each time period. The variation parameter v in PERP was calculated using the past demands of the prefixed 30 time periods by the definition in Section 2.2. We calculated the parameters v and a by their definitions (given in Sections 2.2 and 4.1, respectively), scaled appropriately to make them lie in $[0,1]$. The resulted scatter plots are shown in Figure 2. In (a), v is fixed, so the cost of NO-PRED (blue dots) is approximately the same for all instances. The cost of PURE-PRED (orange dots) is approximately exponential in a, which follows by Observation 1a. In (b), a is fixed, so the cost of PURE-PRED is approximately the same for all instances. The cost of NO-PRED is approximately exponential in v, which follows by Theorem 1. Note that in both (a) and (b), the cost of PERP (green dots) is close to the minimal cost of NO-PRED and PURE-PRED across all instances, showing that PERP's performance is robust in both v and a.

5.2. Experiments on Real Data

We used real-world data sets to represent the "demand" sequences in our experiments. Figure 3 depicts example

Figure 3. (Color online) An Example of a Single Time Series from Each Data Set



Note. In (c), the dashed line represents an additional feature: daily online reservations.

Table 4. Description of Experimental Instances

	Rossmann	Wikipedia	Restaurant
No. of time series	1,115	9	185
Critical quantiles (%)	30,40,50,60,70	95,98,99,99.9	50
Experimental period (days)	300,400,500,600	300,400,500,600,700	100
Prediction update frequency (days)	2,4,10,20	2,4,10,20	5, 10
Total no. of instances	1,000 (sampled)	2,880 (exhaustive)	740 (exhaustive)

time series from each of these data sets. All data sets include multiple daily time series and are publicly available:

- **Rossmann:**[13] Daily number of customers that visited each of 1,115 stores in the *Rossmann* drug store chain during a 781-day period from 2013 to 2015.
- **Wikipedia:**[14] Daily web traffic across Wikipedia.com pages of nine different languages for an 803-day period from 2015 to 2017.
- **Restaurant:**[15] Daily number of visitors and online reservations across 185 restaurants in Japan, during a 478-day period from 2016 to 2017. We treated the number of visitors as the "demand" and the reservations as a predictive feature.

Each instance of our experiment represented a single Nonstationary Newsvendor with Predictions problem, with the realized demands taken from a single time series in our data (a single Rossmann store, a single language on Wikipedia, or a single restaurant). The overage and underage costs were constant within each instance, and without loss of generality the two costs for an instance can be characterized by the corresponding critical quantile (specifically, the ratio of the underage cost to the sum). The time horizon for each instance was a set number of days taken from the end of the time series, with the preceding days used to train one of four prediction methods. These predictions were also updated over the course of the instance at a set frequency. For the Wikipedia data set, this yielded a total of 2,880 possible instances, all of which were tested. The Rossmann data set has multiple orders of magnitude more instances, so we randomly sampled 1,000 from this set. For the Restaurant data set, we used a single prediction method to generate two sets of predictions for each restaurant; one only utilized the number of past visitors, and the other incorporated the number of reservations as a feature, which gave 740 instances. Table 4 describes all of the instances used.

For each instance, we applied `NO-PRED`, `PURE-PRED`, and `PERP` with scaling constants $\kappa = \gamma = 1$, and the variation parameter v in `PERP` was calculated using the past demands of the training data by the definition in Section 2.2. To generate predictions, we used four popular forecasting methods ranging from classic to the state-of-the art:

- **Exponential Smoothing (Holt Winters):** A classic algorithm based on a (linear) trend and seasonality decomposition as in Equation (3), known for its simplicity and robust performance. It is frequently used as a benchmark in forecasting competitions (Makridakis and Hibon 2000). Tuning parameter: seasonality of length 50.
- **ARIMA:** Another classic algorithm that is rich enough to model a wide class of nonstationary time series. Tuning parameters: $(p, q, r) = (3, 2, 5)$.
- **Prophet:** A recent algorithm developed by Facebook (Taylor and Letham 2018) based on a (piecewise-linear) trend and seasonality decomposition, as in Equation (3), known to work well in practice with minimal tuning. Tuning parameters: software default.
- **LightGBM:** A recent algorithm developed by Microsoft (Ke et al. 2017) based on tree algorithms.

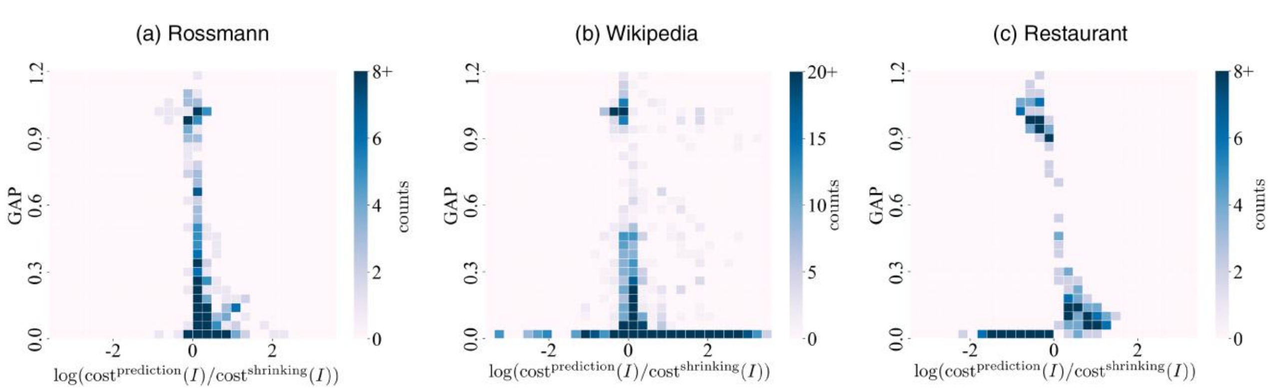


Figure 4. (Color online) Histograms of GAPs Across (a) 1,000 Randomly Sampled Instances on the Rossmann Data Set, (b) 2,880 Instances on the Wikipedia Data Set, and (c) 740 Instances on the Restaurant Data Set

Table 5. Summary of Experimental Results

	Rossmann	Wikipedia	Restaurant
Average GAP with good predictions	0.26	0.40	0.10
Average GAP with bad predictions	0.28	0.07	0.39

LightGBM formed the core of most of the top entries in the recent \$100,000 M5 Forecasting Challenge (Makridakis et al. 2022). Tuning parameters: software default.

For the Restaurant data set, we used Prophet as the forecasting method, with and without the reservations as an additive linear feature. We treated the outputs of these methods as predictions of the *mean* demand. To estimate the demand distribution around this mean, we used the empirical distribution of the residuals of the same predictions on the training period.[16] In practice, even if the prediction quality is good, the predictions of the first few days might incur large costs because of noise/instability of the predictions, which may cause PERP to misidentify the prediction quality. Therefore, we restricted PERP to following the predictions for the first 20 days, allowing switches only afterward.

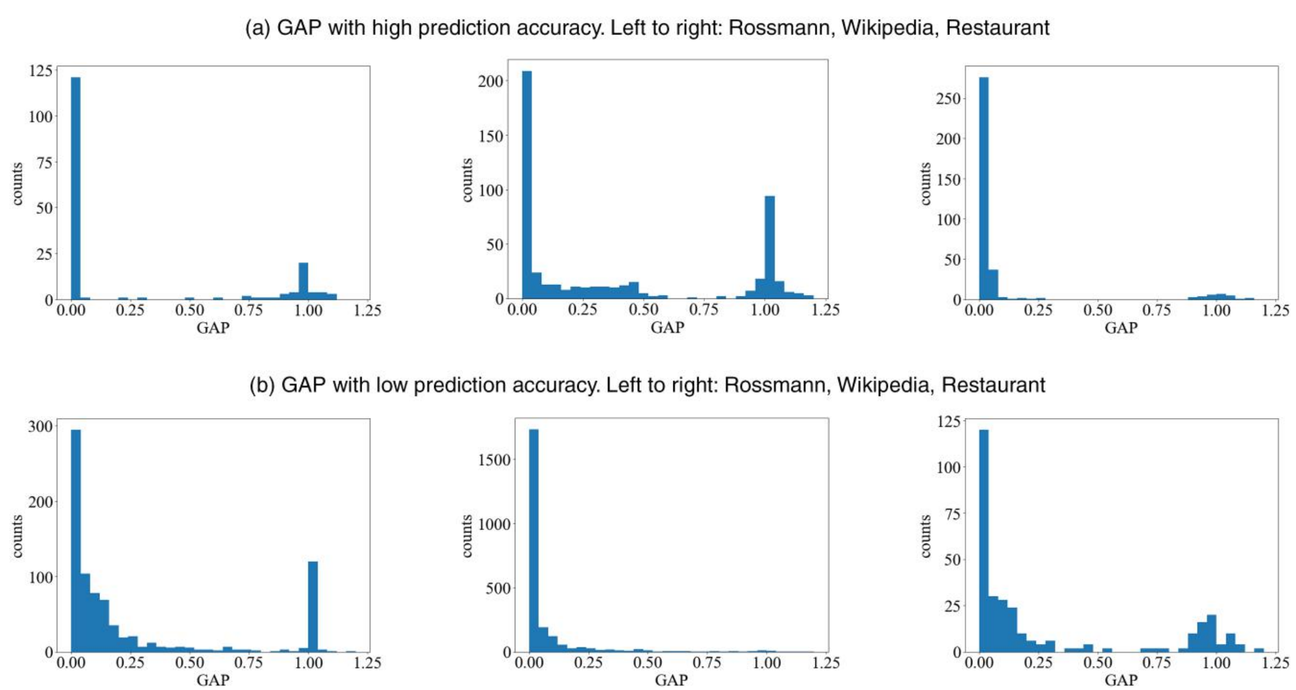
Results. Each instance yields three total costs: one incurred by PERP and two incurred by the benchmark algorithms (NO-PRED and PURE-PRED). The primary performance metric that we report is a form of optimality gap. For an instance I, let $\text{cost}^{\text{PURE-PRED}}(I)$ be the cost of PURE-PRED, and similarly define $\text{cost}^{\text{NO-PRED}}(I)$, $\text{cost}^{\text{PERP}}(I)$. Then, the optimality gap (GAP) of PERP is defined as

$$\text{GAP}(I) = \frac{\text{cost}^{\text{PERP}}(I) - \min\{\text{cost}^{\text{PURE-PRED}}(I), \text{cost}^{\text{NO-PRED}}(I)\}}{|\text{cost}^{\text{PURE-PRED}}(I) - \text{cost}^{\text{NO-PRED}}(I)|}.$$

If we think of PERP as trying to achieve the minimum of the costs incurred by the two benchmark policies, then GAP measures the excess cost that PERP incurs on top of this minimum, which is normalized so that $\text{GAP} = 0$ implies that the minimum has been achieved, and $\text{GAP} = 1$ implies that the maximum of the two costs was incurred.[17]

Experiments on the data sets yielded the histograms in Figure 4. For each instance I, the value on the horizonal axis is $\log(\text{cost}^{\text{PURE-PRED}}(I)/\text{cost}^{\text{NO-PRED}}(I))$, which is greater than 0 if NO-PRED has a lower cost and less than 0 if PURE-PRED has a lower cost. In the 1,000 Rossman instances NO-PRED had a lower cost 82.7% of the time, in the 2,880 Wikipedia instances NO-PRED had a lower cost 81.9% of the time, and in the 740 Restaurant instances NO-PRED had a lower cost 64.3% of the time. The values on the vertical axis are the GAPs. Note that most GAPs are small when the

Figure 5. (Color online) Histograms of the GAPs for (a) 173 Rossmann Instances (left), 522 Wikipedia Instances (Middle), and 476 Restaurant Instances (Right) for Which PURE-PRED Has Lower Cost and (b) 827 Rossmann Instances (Left), 2,358 Wikipedia Instances (Middle), and 264 Restaurant Instances (Right) for Which NO-PRED Has Lower Cost



absolute values of the log difference are large. This shows that PERP performs very well when the difference of costs between NO-PRED and PURE-PRED is large. On the other hand, there are instances where PERP has large GAPs, and in particular there are instances with GAPs equal to 1 when the log difference of costs is close to 0. This happens because when the log difference of costs is close to 0, the cost of NO-PRED and the cost of PURE-PRED are close, so PERP may misidentify the prediction quality. Still, because the maximum cost and the minimum cost of the other two policies are close, even the GAPs are large in these instances, and PERP does not perform badly.

We further divide the instances according to which of NO-PRED and PURE-PRED had lower cost in Table 5 and Figure 5. For comparison, if we did not know the prediction quality beforehand, uniformly random choosing between NO-PRED and PURE-PRED had an expected GAP of 0.5. Therefore, PERP outperforms this natural benchmark in all cases of all data sets.

6. Conclusion

We proposed a new model incorporating predictions into the nonstationary newsvendor problem. We first gave a complete analysis of the Nonstationary Newsvendor (without predictions) by proving a lower regret bound and developing the Shrinking-Time-Window Policy, which was the first policy that achieved the lower bound up to log factors without knowing the variation parameter. We then considered the Nonstationary Newsvendor with Predictions and proposed the Prediction-Error-Robust Policy, which does not need to know the prediction quality beforehand and achieves nearly optimal minimax worst-cast regret.

Endnotes

[1] The $\tilde{O}(\cdot)$ notation hides logarithmic factors.

[2] All of our results carry through if Q is allowed to depend on t.

[3] The demand is not censored here, as is the case in all of the motivating examples in the Introduction. The censored version of our problem is an interesting but separate subject.

[4] A random variable X is *sub-Gaussian* with *sub-Gaussian norm* $\|X\|_{\psi_2}$ if $\mathbb{P}(|X|>x)\le 2\exp(-x^2/\|X\|_{\psi_2}^2)$ for all $x\ge 0$. For sub-Gaussian variables, we have $\mathbb{E}[|X|]\le 3\|X\|_{\psi_2}$.

[5] Unfortunately, Assumption 2 is not guaranteed to hold. For example, for the family of distributions

$$\mathcal{D}_\mu \sim \begin{cases} \mu, & \mu\in[0,1] \\ \mu+\text{Bernoulli}(0.5)-0.5, & \mu\in(1,2], \end{cases}$$

the function $C(\mu,1,1,1)$ is discontinuous (and thus not Lipschitz) at $\mu=1$.

[6] As a sanity check, the classical result for the newsvendor problem (Arrow et al. 1958, Scarf et al. 1960) states that if $Q=\mathbb{R}$, then q_t^* is the $b_t/(b_t+h_t)$-th quantile of D_t.

[7] Although not stated explicitly, the results in Keskin et al. (2023) require Q only to contain points arbitrarily close to every optimal unconstrained order quantity.

[8] Note that we are not considering randomized policies here, but all of our theoretical results (the lower bounds, in particular) hold even when randomization is allowed.

[9] Note that this is different from a clairvoyant that knows the realized demands $d_1,\dots,d_T$. Such a clairvoyant would incur zero cost.

[10] As a final comparison to Keskin et al. (2023), they do not consider the unknown v setting.

[11] We are taking the predictions to be entirely deterministic, so for example, a_t is not allowed to depend on the previously observed demands $d_1,\dots,d_{t-1}$. Our results hold if we extend to the setting in which the predictions are stochastic (and adapted to the demand filtration).

[12] Indeed, it implies that no algorithm can achieve regret $O(T^{f(v,a)})$ for a function $f:[0,1]\times[0,1]\to[0,1]$ satisfying $f(v,a)\le\max\{(3+v)/4,a\}$ for all $v,a\in[0,1]$ and $f(v,a)<\max\{(3+v)/4,a\}$ for some $v,a\in[0,1]$.

[13] Available at https://www.kaggle.com/competitions/rossmann-store-sales/data.

[14] Available at https://www.kaggle.com/competitions/web-traffic-time-series-forecasting/data.

[15] Available at https://www.kaggle.com/competitions/recruit-restaurant-visitor-forecasting/.

[16] That is, if the training data consist of T_{train} periods, which without loss we index as $\{t=-T_{\text{train}}+1,T_{\text{train}}+2,\dots,-1,0\}$, then the demand distribution at any time t was estimated to be

$$\hat{\mu}_t+\text{Uniform}(\{d_s-\hat{\mu}_s : s=-T_{\text{train}}+1,T_{\text{train}}+2,\dots,-1,0\}).$$

[17] GAP may technically be outside of $[0,1]$.

References

An L, Li AA, Moseley B, Visotsky G (2024) Best of many in both worlds: Online resource allocation with predictions under unknown arrival model. Preprint, submitted February 21, https://arxiv.org/abs/2402.13530.

Antoniadis A, Gouleakis T, Kleer P, Kolev P (2020) Secretary and online matching problems with machine learned advice. *Advances in Neural Information Processing Systems*, vol. 33 (Curran Associates, Inc., Red Hook, NY), 7933–7944.

Arrow KJ, Karlin S, Scarf H (1958) *Studies in the Mathematical Theory of Inventory and Production* (Stanford University Press, Stanford, CA).

Aviv Y, Pazgal A (2005) A partially observed Markov decision process for dynamic pricing. *Management Sci.* 51(9):1400–1416.

Azoury KS (1985) Bayes solution to dynamic inventory models under unknown demand distribution. *Management Sci.* 31(9):1150–1160.

Baby D, Wang YX (2019) Online forecasting of total-variation-bounded sequences. *Advances in Neural Information Processing Systems*, vol. 32 (Curran Associates, Inc., Red Hook, NY).

Bai Y, Zhang YJ, Zhao P, Sugiyama M, Zhou ZH (2022) Adapting to online label shift with provable guarantees. *Advances in Neural Information Processing Systems*, vol. 35 (Curran Associates Inc., Red Hook, NY), 29960–29974.

Balseiro S, Kroer C, Kumar R (2022) Single-leg revenue management with advice. Preprint, submitted February 18, https://arxiv.org/abs/2202.10939.

Ban GY, Rudin C (2019) The big data newsvendor: Practical insights from machine learning. *Oper. Res.* 67(1):90–108.

Besbes O, Muharremoglu A (2013) On implications of demand censoring in the newsvendor problem. *Management Sci.* 59(6):1407–1424.

Besbes O, Gur Y, Zeevi A (2014) Stochastic multi-armed-bandit problem with non-stationary rewards. *Advances in Neural Information Processing Systems*, vol. 27 (Curran Associates Inc., Red Hook, NY).

Besbes O, Gur Y, Zeevi A (2015) Non-stationary stochastic optimization. *Oper. Res.* 63(5):1227–1244.

Burnetas AN, Smith CE (2000) Adaptive ordering and pricing for perishable products. *Oper. Res.* 48(3):436–443.

Chen B, Chao X, Ahn HS (2019a) Coordinating pricing and inventory replenishment with nonparametric demand learning. *Oper. Res.* 67(4):1035–1052.

Chen X, Wang Y, Wang YX (2019b) Nonstationary stochastic optimization under l p, q-variation measures. *Oper. Res.* 67(6):1752–1765.

Cheung WC, Simchi-Levi D, Zhu R (2022) Hedging the drift: Learning to optimize under nonstationarity. *Management Sci.* 68(3):1696–1713.

Dütting P, Lattanzi S, Paes Leme R, Vassilvitskii S (2021) Secretaries with advice. *Proc. 22nd ACM Conf. Econom. Comput.* (Association for Computing Machinery, New York), 409–429.

Edgeworth FY (1888) The mathematical theory of banking. *J. Roy. Statis. Soc.* 51(1):113–127.

Godfrey GA, Powell WB (2001) An adaptive, distribution-free algorithm for the newsvendor problem with censored demands, with applications to inventory and distribution. *Management Sci.* 47(8):1101–1112.

Hao B, Jain R, Lattimore T, Van Roy B, Wen Z (2023) Leveraging demonstrations to improve online learning: Quality matters. *Proc. 40th Internat. Conf. Machine Learn*. (JMLR.org).

Hu P, Jiang J, Lyu G, Su H (2024) Constrained online two-stage stochastic optimization: Algorithm with (and without) predictions. Preprint, submitted January 2, https://arxiv.org/abs/2401.01077.

Huang C, Wang K (2023) A stability principle for learning under non-stationarity. Preprint, submitted October 27, https://arxiv.org/abs/2310.18304.

Huber J, Müller S, Fleischmann M, Stuckenschmidt H (2019) A data-driven newsvendor problem: From data to decision. *Eur. J. Oper. Res.* 278(3):904–915.

Huh WT, Rusmevichientong P (2009) A nonparametric asymptotic analysis of inventory planning with censored demand. *Math. Oper. Res.* 34(1):103–123.

Iglehart DL (1964) The dynamic inventory problem with unknown demand distribution. *Management Sci.* 10(3):429–440.

Jin B, Ma W (2022) Online bipartite matching with advice: Tight robustness-consistency tradeoffs for the two-stage model. *Advances in Neural Information Processing Systems*, vol. 35 (Curran Associates Inc., Red Hook, NY), 14555–14567.

Karlin S (1960) Dynamic inventory policy with varying stochastic demands. *Management Sci.* 6(3):231–258.

Karnin ZS, Anava O (2016) Multi-armed bandits: Competing with optimal sequences. *Advances in Neural Information Processing Systems*, vol. 29 (Curran Associates Inc., Red Hook, NY).

Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, Ye Q, Liu T-Y (2017) Lightgbm: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, vol. 30 (Curran Associates Inc., Red Hook, NY).

Keskin NB, Zeevi A (2017) Chasing demand: Learning and earning in a changing environment. *Math. Oper. Res.* 42(2):277–307.

Keskin NB, Min X, Song JSJ (2023) The nonstationary newsvendor: Data-driven nonparametric learning. Preprint, submitted June 7, https://dx.doi.org/10.2139/ssrn.3866171.

Kleywegt AJ, Shapiro A, Homem-de Mello T (2002) The sample average approximation method for stochastic discrete optimization. *SIAM J. Optim.* 12(2):479–502.

Kunnumkal S, Topaloglu H (2008) Using stochastic approximation methods to compute optimal base-stock levels in inventory control problems. *Oper. Res.* 56(3):646–664.

Lattanzi S, Lavastida T, Moseley B, Vassilvitskii S (2020) Online scheduling via learned weights. *SODA* (SIAM, Philadelphia), 1859–1877.

Levi R, Perakis G, Uichanco J (2015) The data-driven newsvendor problem: New bounds and insights. *Oper. Res.* 63(6):1294–1306.

Levi R, Roundy RO, Shmoys DB (2007) Provably near-optimal sampling-based policies for stochastic inventory control models. *Math. Oper. Res.* 32(4):821–839.

Liyanage LH, Shanthikumar JG (2005) A practical inventory control policy using operational statistics. *Oper. Res. Lett.* 33(4): 341–348.

Lovejoy WS (1990) Myopic policies for some inventory models with uncertain demand distributions. *Management Sci.* 36(6):724–738.

Luo H, Wei CY, Agarwal A, Langford J (2018) Efficient contextual bandits in non-stationary worlds. *Conf. Learn. Theory* (PMLR), 1739–1776.

Lykouris T, Vassilvitskii S (2021) Competitive caching with machine learned advice. *J. ACM* 68(4):1–25.

Mahdian M, Nazerzadeh H, Saberi A (2012) Online optimization with uncertain information. *ACM Trans. Algorithms* 8(1):1–29.

Makridakis S, Hibon M (2000) The m3-competition: Results, conclusions and implications. *Internat. J. Forecast.* 16(4):451–476.

Makridakis S, Spiliotis E, Assimakopoulos V (2022) M5 accuracy competition: Results, findings, and conclusions. *Internat. J. Forecast.* 38(4):1346–1364.

Munoz A, Vassilvitskii S (2017) Revenue optimization with approximate bid predictions. *Advances in Neural Information Processing Systems*, vol. 30 (Curran Associates Inc., Red Hook, NY).

Oroojlooyjadid A, Snyder LV, Takáč M (2020) Applying deep learning to the newsvendor problem. *IISE Trans.* 52(4):444–463.

Powell W, Ruszczyński A, Topaloglu H (2004) Learning algorithms for separable approximations of discrete stochastic optimization problems. *Math. Oper. Res.* 29(4):814–836.

Rohatgi D (2020) Near-optimal bounds for online caching with machine learned advice. *Proc. Fourteenth Annual ACM-SIAM Sympos. Discrete Algorithms* (Society for Industrial and Applied Mathematics, Philadelphia), 1834–1845.

Scarf H (1959) Bayes solutions of the statistical inventory problem. *Ann. Math. Statist.* 30(2):490–508.

Scarf H, Arrow K, Karlin S, Suppes P (1960) The optimality of (s, s) policies in the dynamic inventory problem. *Optimal Pricing, Inflation, and the Cost of Price Adjustment*, 49–56.

Shapiro A (2003) Monte Carlo sampling methods. *Handbooks Oper. Res. Management Sci.* 10:353–425.

Taylor SJ, Letham B (2018) Forecasting at scale. *Amer. Statist.* 72(1):37–45.

Treharne JT, Sox CR (2002) Adaptive inventory control for nonstationary demand and partial information. *Management Sci.* 48(5):607–624.

Winters PR (1960) Forecasting sales by exponentially weighted moving averages. *Management Sci.* 6(3):324–342.

Yang T, Zhang L, Jin R, Yi J (2016) Tracking slowly moving clairvoyant: Optimal dynamic regret of online learning with true and noisy gradient. *Internat. Conf. Machine Learn.* (PMLR), 449–457.

Zhang L, Lu S, Zhou ZH (2018) Adaptive online learning in dynamic environments. *NIPS* 31.

Zhang L, Yang J, Gao R (2024) Optimal robust policy for feature-based newsvendor. *Management Sci.* 70(4):2315–2329.