Warsaw University of Life Sciences – SGGW
Institute of Economics and Finance
Department of Econometrics and Statistics

QUANTITATIVE METHODS IN ECONOMICS

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH

Volume XXV, No. 3

Warsaw 2024

ISSN 2082-792X
e-ISSN 2543-8565

Warsaw 2024, Volume XXV, No. 3
Journal homepage: https://qme.sggw.edu.pl
Published by Warsaw University of Life Sciences Press

QUANTITATIVE METHODS IN ECONOMICS
Volume XXV, No. 3, 2024

CONTENTS

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
TOM XXV/3, 2024, s. 99 – 106

DĄŻENIE DO PRAWDY JAKO ŹRÓDŁO ETYKI STATYSTYKA[1]

Czesław Domański https://orcid.org/0000-0001-6144-6231
Uniwersytet Łódzki
Instytut Statystyki i Demografii
e-mail: czedoman@uni.lodz.pl

Streszczenie: Postawa etyczna statystyka to dążenie do prawdy oraz unikanie i eliminowanie wpływów i nacisków wewnętrznych, jak i zewnętrznych na rzetelny końcowy efekt badań. Celem artykułu jest określenie źródeł gwarantujących postawę etyczną, w której uwzględniamy: dążenie do prawdy, profesjonalizm i niezależność badawczą. Najważniejszym przesłaniem „zasad etycznych" jest dążenie do prawdy. Omawiając problem związany z etyką statystyka należy uwzględniać niebezpieczeństwa, gdyż informacja statystyczna stała się „towarem" z jednej strony ze względu na szybki rozwój technologii elektronicznej, które stwarzają możliwości manipulacji tą informacją.

Keywords: dążenie do prawdy, niezależność statystyka

JEL classification: C18, C69, H69

„(…) nie trzeba kłaniać się okolicznościom.
Prawdom kazać by za drzwiami stały"

Cyprian Kamil Norwid (1821-1883)

UWAGI WSTĘPNE

Zachowanie idei autonomii i obiektywizmu w statystyce publicznej oznacza tworzenie i rozwijanie metodologii statystyki i jej systematyczne rozpowszechnianie z pełnym poszanowaniem zasad etycznych.

Trafnie określa Sedláček [2012 s. 26] „cała historia etyki to ciągłe dążenie do stworzenia etycznego zachowania".

[1] Artykuł wygłoszony na XXV Jubileuszowej Konferencji Metody Ilościowe w Badaniach Ekonomicznych, Warszawa, 20 czerwca 2024 r.

https://doi.org/10.22630/MIBE.2024.25.3.9

Zarówno Hebrajczycy, jak i chrześcijanie dużą wagę przykładali do moralności, zwłaszcza odkrywanej dzisiaj na nowo etyki cnót, a dobrego życia nie można sobie wyobrazić bez badania dobra i zła.

„Przez etykę cnót rozumiemy opartą na cnotach: nie na odpowiedzialności, nie na korzyści, nie na użyteczności, nie na kalkulacji wyników" (por. Sedláćek, [2012 s. 32]).

Prekursorem etyki cnót był Platon (427-347 p.n.e.), ale w rzeczywistości rozwinął ją dopiero Arystoteles (384-322 p.n.e.). Stanowiła ona dominującą szkołę etyki w naszej cywilizacji aż do oświecenia, kiedy to częściowo zastąpił utylitaryzm i kantowska deontologia (moralność oparta na odpowiedzialności, dobrych intencjach i przestrzeganiu zasad).

Etyka zniknęła z głównego nurtu ekonomii. Dla ekonomistów etyka stała się nieciekawa i nieważna. Nie było potrzeby o niej rozmawiać, wystarczyło polegać na „niewidzialnej ręce rynku", która indywidualne wady (np. egoizm) automatycznie zamieniała w ogólne dobro (np. wzrost produktywności).

Obecnie we wszystkich dziedzinach życia ekonomicznego i społecznego coraz wyraźniej zauważa się oddziaływanie procesów informatycznych na wszystkich etapach badań statystycznych: zbierania, przetwarzania i analizy danych statystycznych ich interpretacji. Na każdym z wymienionych etapów badania statystycznego, jak i interpretacji wyników statystyka obowiązuje postawa etyczna.

Postawa etyczna - czyli dążenie do obiektywnej prawdy oraz unikanie i eliminowanie wpływów i nacisków zarówno wewnętrznych, jak i zewnętrznych na rzetelny końcowy efekt badań. Najważniejszym przesłaniem postawy etycznej jest „poszukiwanie prawdy".

Nieodzownym warunkiem utrzymania postawy etycznej statystyka jest konieczność jego wysokiego profesjonalizmu w zakresie metodologii i technik badawczych opartych na nowych technologiach.

Celem artykułu jest próba przedstawienia źródeł warunków etycznej postawy statystyka, do których zaliczamy:

- dążenie do prawdy
- profesjonalizm
- niezależność badawczą.

DĄŻENIE DO PRAWDY

Jak zauważa Stawrowski [2023 s. 72] „Już pogańscy Grecy uznawali za rzecz najważniejszą – dążenie do prawdy, co daje się sensownie pomyśleć tylko wtedy, gdy ufamy, że świat jest rozumny, że można go poznać i że stanowi to nasze zadanie".

Etyka zajmuje się tworzeniem systemów myślowych, z których można wyprowadzić zasady postępowania. Samo słowo „etyka" wywodzi się z greckiego „ethos", co oznacza obyczaj, rozumiany nie tylko jako dane zjawisko, ale także jego otoczenie społeczne i materialne. Etyka jest przestrzenią, w której ramach funkcjonują normy moralne, będące w społeczeństwach systemami normatywnymi, obok takich jak prawo czy moralność religijna" (zob. Galata [2007 s. 221]).

Informacja nie stanowi wartości sama w sobie. Wartość informacji lub mówiąc dokładniej, korzyść z jej posiadania zależy od umiejętności jej wykorzystania. W miarę dynamicznego rozwoju działalności w sektorze usług informacyjnych z najnowszymi technologiami i metodologii przetwarzania i przekazywania informacji, rośnie również coraz większe zagrożenie w związku z potencjalnym nadużyciem, prowadzonym niewłaściwym wykorzystaniem danych indywidulanych z uszczerbkiem dla jednostek respondentów, które te dane dostarczają w dobrej wierze.

Wiedza statystyczna będzie dla jednostek cennym nabytkiem pozwalającym chronić siebie i swoją rodzinę przed niebezpieczeństwem propagandy polityków i pozbawioną skrupułów reklamą przedsiębiorców, pozwoli korzystać z prognoz pogody i osiągać korzyści z innych czynników kształtujących życie ludzkie nad którym nie sprawuje ona władzy. Osoby i instytucje zajmujące się badaniami statystycznymi winny przyjąć środki, które należy podejmować w celu ochrony zapewnienia poufności i bezpieczeństwa tych danych. Ponadto powinny przestrzegać procedury zasad etyki statystycznej na wszystkich etapach: projektowania, zbierania, analizy, budowy bazy danych i przetwarzania informacji oraz interpretacji.

Francis A. Walker (1840-1897) pierwszy prezydent Waszyngtońskiego Towarzystwa Statystycznego w 1896 roku napisał „pożądane jest by ludzie którzy mają stosować statystykę – i każdy pisząc o historii, ekonomii, socjologii, którzy muszą stosować statystykę – nauczyli się przestrzegać ograniczeń dotyczących odpowiedzialności i wrażliwości liczb". Jest to dzisiaj równie ważne i aktualne jak 128 lat temu.

Ożywione dyskusje na temat „zasad etycznych" sięgają 1947 r., bowiem pierwsze sugestie na ten temat sformułował Eisenhort (por. Jowell [1981]), a które przedstawił jako <u>obowiązki statystyków wobec użytkowników informacji „Najważniejszym przesłaniem zasad etycznych jest poszukiwanie prawdy".</u>

Problem dochodzenia do prawdy

Statystyk musi nieustannie zajmować się szukaniem prawdy. Do przeciętnego odbiorcy informacji dochodzi głównie to co jest głośne, co zostało nagłośnione (media). Często podawane informacje się wykluczają (jeśli korzysta się z różnych źródeł) i się nimi manipuluje. Potem okazuje się, że informacje jawnie są nieprawdziwe, bo wyraźnie chodziło o zrobienie afery.

Tak często odbiorca ocenia różne informacje, także statystyczne. W takich warunkach pracuje statystyk.

Jak zauważył Ernest Bryl (1935-2024) „Wcześniej w Polsce istniała cenzura prewencyjna. Cenzura tworzącą atmosferę, że o czymś nie można bądź nie wypada mówić (pisać). Jeśli w społeczeństwie istnieją jakieś elementy cenzury, to właśnie w elitach, gdzie o pewnych rzeczach nie mówimy, bo nie wypada”.

„O czym nie wypada mówić? Ach, niemal o wszystkim. Dzięki nowym technologiom, smartfonom, internetowi, z jednej strony szalenie trudno jest wygasić jakąś informację, z drugiej – można ją przeinaczyć".

PROFESJONALIZM

Profesjonalista to ktoś dobrze znający swój zawód; zawodowiec, specjalista. M.G. Kendall i W.R. Buckland [1986] podają następującą definicję statystyki (s. 202).

„Statystyka, dane numeryczne dotyczące agregatów złożonych z pewnych jednostek; nauka zajmująca się zbieraniem, analizą i interpretacją tego typu danych”.

„Statystyka jest bardziej sposobem myślenia i wnioskowania niż pęczkiem recept na młócenie danych w celu odsłonięcia odpowiedzi” [Rao 1994 s. 64].

„Czy statystyka taka, jak się ją studiuje i stosuje w praktyce jest: <u>nauką, techniką czy sztuką?</u> Może jest połączeniem tego wszystkiego! Jest nauką w tym sensie, że ma tożsamość z dużym repertuarem technik wywodzących się z pewnych zasad podstawowych.

(...) Co więcej istnieją kwestie filozoficzne związane z podstawami statystyki – sposób w jaki można określić ilościowo i wyrazić niepewność – które można rozważać niezależnie od jakiejkolwiek dziedziny zastosowań. Tak więc w szerszym sensie statystyka jest odrębną dyscypliną, być może dyscypliną wszystkich dyscyplin”.

<u>Jest techniką</u> – metodologię statystyczną można włączyć w każdy działający system w celu utrzymania wymaganego poziomu i stabilności jego założeń. Np. w planach kontroli jakości produkcji przemysłowej, w systemach bankowych.

Metody statystyczne pozwalają kontrolować, redukować i uwzględniać niepewność i ryzyko, a przez to powodują maksymalizację efektywności, działań osób i instytucji.

<u>„Jest sztuką</u> – ponieważ jej metodologia zależy od rozumowania indukcyjnego, nie jest w pełni skodyfikowana ani wolna od kontrowersji”.

Statystycy mogą dochodzić do różnych wniosków korzystając z tego samego zbioru danych (np. posługując się różnymi narzędziami, niedobrze gdy celowo-wtedy ucieka prawda).

W zasadzie istniejące dane zawierają więcej informacji, niż można uzyskać dostępnymi narzędziami statystycznymi. Wydobycie z danych jak najwięcej informacji zależy od wprawy i doświadczenia – profesjonalizmu statystyka.

Profesjonalizm statystyka to połączenie nauki i techniki czyli odpowiedniej i odpowiedzialnie wybranej metodologii, najlepszej z punktu badanego problemu weryfikacji wiedzy.

Wyniki należy prezentować rzetelnie i obiektywnie uwzględniając możliwie pełne wykorzystanie informacji w interpretacji wraz z wyjaśnieniem – taki komunikat może przedstawić statystyk ujawniając umiejętność statystyka jako „artysty".

WYSOKA CENA INFORMACJI

Wyżnikiewicz [2021] przedstawił trudną drogę statystyków w procesie dążenia do prawdy. We wprowadzeniu formułuje najważniejsze przesłanie do statystyków (s. 50).

„Zadaniem statystyki jest zatem przedstawianie obiektywnego i zgodnego z rzeczywistością – czyli prawdziwego i wyrażonego w liczbach – opisu rzeczywistości gospodarczej i społecznej przy pełnej niezależności i autonomii statystyki publicznej. Takie postawienie sprawy w społeczeństwie demokratycznym powinno być czymś oczywistym, naturalnym i oczekiwanym. Tymczasem nie tylko historia, lecz także współczesność zna przypadki karania i prześladowania statystyków, których badania przedstawiały prawdę z różnych powodów niewygodną dla władzy. Faktyczna liczba takich przypadków nie jest znana".

Dalej omawia kilka przypadków: takie jak skazanie na więzienie Anar Meshimbayevej, prezes Urzędu Statystycznego w Kazachstanie. W Kanadzie w 2016 i 2020 r. dwóch kolejnych prezesów urzędu statystycznego złożyło rezygnacje ze stanowiska w proteście przeciwko ingerencji władz w niezależność statystyki.

Ocenia także sytuację w naszym kraju następująco:

„Polska statystyka w ostatnich stu latach nie doświadczyła przypadków karania statystyków za podawanie prawdy, jednakże w 1951 r., w czasach terroru stalinowskiego, kilku pracowników GUS zostało aresztowanych przez Urząd Bezpieczeństwa. Formalne powody, czas trwania i zasięg tych represji nie są znane. Aresztowania dotyczyły osób, które próbowały kontynuować działalność wydawniczą GUS. Zatrzymaniom towarzyszyła rekwizycja materiałów statystycznych przygotowywanych do publikacji w roczniku statystycznym [Łazowska 2017]. W latach 1948-1951 ukazały się cztery jego wydania. Natomiast w latach 1952-1955, aż do maja 1956 r. – co z dzisiejszej perspektywy wydaje się nieprawdopodobne - GUS nie wydawał roczników statystycznych. W przedmowie do *Rocznika Statystycznego 1955* [GUS 1956] napisano: „Pewne tematy nie zostały w *Roczniku* uwzględnione, niektóre zaś przedstawiono w sposób niewyczerpujący".

Przytoczmy za Peukerem [2008 s. 23] sytuację codziennej pracy statystyka związana z pomysłem opublikowania jawnej statystyki wypadku przy pracy.

„W wielu przypadkach ukrywanie wypadków w ogóle, ich ciężkości czy przyczyn pośrednich i bezpośrednich było wymuszane, żeby nie użyć określenia mocniejszego. W rozważaniach kuluarowych znawcy zagadnienia usiłowali operować szacunkami tych zafałszowań, ale sprawa była bardzo trudna. Zresztą prof. Stefan Szulc mawiał, że statystyk nie może być ani policjantem, ani prokuratorem, ani też sędzią. I z takimi problemami miałem do czynienia w toku swojej pracy w statystyce. Zło, o którym się wie, którego przyczyny się zna, a którego zwalczyć nie można, boli podwójnie.

Statystyk źle się czuje, jeśli wyniki jego pracy nie docierają do tych, którym są potrzebne. Opracowaliśmy więc specjalną publikację (też zresztą ograniczoną zakresowo i tematycznie w stosunku do materiałów, którymi dysponowaliśmy) i po uzyskaniu aprobaty członka kierownictwa GUS, udałem się do słynnego Urzędu Kontroli Prasy, Publikacji i Widowisk (UKPPiW), bez zgody którego nie mogła ukazać się żadna publikacja. Tamże, właściwy cenzor w randze i mundurze pułkownika (nazwisko jego zapamiętałem na całe życie), nie zajrzawszy nawet do wnętrza publikacji wybuchnął jak mina i wygłosił orację, z której wynikało, iż jest to temat wrogi w stosunku do całego obozu socjalistycznego, sprzyja zaś wrogom socjalizmu i podżegaczom wojennym. Poczułem się jak szpieg, agent, wróg socjalizmu.

Wyszedłem z wrogą publikacją pod pachą. Towarzysz pułkownik żadnych pytań nie zadawał, żadnych odpowiedzi z mojej strony nie oczekiwał. Załatwił sprawę sprawnie, po żołniersku. Tak skończył się pomysł opublikowania jawnej statystyki wypadków przy pracy. Wydawaliśmy nadal opracowania tej tematyki, ale w małych nakładach, opatrzonych różnymi klauzulami dostępności. Są w zasobach archiwalnych GUS już chyba „odpoufnione", ale potrzebne, jeśli w ogóle, to nader niewielkiej liczbie osób".

UWAGI KOŃCOWE

Przyszłość statystyki napawa optymizmem. Bez statystyki nie będziemy w stanie odnaleźć prawdy, którą systemowo często się ukrywa lub transponuje w kłamstwo Rotengruber [2001]. Jest ona wyjątkową dyscypliną nauki, o szczególnym wyjątkowym statusie wśród innych dyscyplin – tak dzięki wszechstronnej obecności w różnych obszarach badań, jak i charakterowi dostarczanych metod i narzędzi dla wyznaczania i realizacji procesów badawczych we wszystkich dziedzinach empirycznych. Wpływa w ten sposób na metodologię generowania i analizy danych zarówno w dyscyplinach obserwacyjnych, bazujących na sprawozdawczości, sondażach i spisach, jak i eksperymentalnych bazujących na manipulowalności i kontroli zmiennych i naukach przyrodniczych i technicznych oraz w sferze podejmowania decyzji publicznych i prywatnych.

Metodologia statystyki rozwinięta dzięki pracom nad teorią próbkowania i wnioskowania statystycznego i modelom statystycznym J. Spławy-Neymana (1894-1981) oraz z drugiej strony – R. A. Fishera (1990-1960), która obejmowała teorie układów eksperymentalnych, wnioskowanie, analizę wariacji, testowanie hipotez oraz teorii decyzji (w sensie A. Walda (1902-1950)), była zarazem impulsem w wielu dziedzinach ludzkiej działalności.

W szczególności inspirowała tworzenie nowych narzędzi w obliczu ryzyka i niepewności, które występowały w różnego typu sytuacjach problemowych w gospodarce i społeczeństwie i analizowanych w ekonomii, socjologii, naukach o zdrowiu itp. Stąd akcentowana już wszechobecność statystyk, przewyższającej już pod tym względem jakąkolwiek innowację technologiczną bądź naukową ubiegłego stulecia.

Trwałą wartość posiada zasługująca na przytoczenie w tym miejscu mądrość w słowach Kartezjusza, iż „jest prawdą zupełnie pewną, że gdy nie jest w naszej mocy rozpoznać mniemania najprawdziwsze, winniśmy iść za najbardziej prawdopodobnymi".

Warto zwrócić uwagę także na etyczną stronę postawy statystyków wobec dynamicznie rozwijającej się sztucznej inteligencji, która obejmuje coraz szersze obszary życia gospodarczego, społecznego i osobistego.

Zauważamy, że w największych światowych korporacjach takich, jak Google czy Amazon pracuje kilkaset tysięcy programistów, a nie ma wśród nich żadnego etyka, filozofa ani teologa.

Można z zadowoleniem przyjąć wypełnienie tej luki przez Catholic Institute of Technology (CIT) założony przez Amerykanów, w którym prowadzone są badania nad etycznymi aspektami nowych technologii. Stara się on wypracować zasady postepowania zgodnie z nauką moralną Kościoła i przenieść je na grunt technologicznego biznesu Doliny Krzemowej i innych miejsc. Instytut angażuje się w badania, edukację oraz dialog między nauką i religią starając się znaleźć wspólne punkty i tworzyć przestrzeń dla konstruktywnego współdziałania naukowców z różnych dziedzin. Co ciekawe, ta instytucja zupełnie niezależnie prowadzi badania nad etycznymi aspektami nowych technologii, takimi jak sztuczna inteligencja, biotechnologia, czy robotyka, starając się wypracować zasady postępowania zgodnie z nauką moralną Kościoła.

CIT stara się być pomostem pomiędzy światem technologii i wartościami katolickimi, promując dialog, zrozumienie i odpowiednie wykorzystanie potencjału technologicznego dla dobra wspólnego.

BIBLIOGRAFIA

Domański Cz. (1998) Etyczne i pragmatyczne przesłanki ochrony informacji statystycznych. Annales. Etyka w Życiu Gospodarczym, 1, 79-88, Salezjańska Wyższa Szkoła Ekonomii i Zarządzania.

Galata S. (2007) Biznes w przestrzeni etycznej. Motywy, metody konsekwencji. Difin, Warszawa.

Jones T. M. (1995) Instrumental Stakeholder Theory: A Synthesis of Ethics and Economics. Academy of Management Review, 20, 404-437.

Kendall M. G. Buckland W. R. (1986) Słownik terminów statystycznych. PWN, Warszawa.

Kucharczyk G. (2001) Katedra i uniwersytet. O kryzysie i nadziei chrześcijańskiej cywilizacji. Wydawnictwo Biały Kruk, Kraków.

Łazowska B. (2017) Główny Urząd Statystyczny w latach 1945-1949 [w:] Historia Polski w liczbach. Statystyka Polski. Dawniej i dziś. GUS Warszawa.

Peuker Z. (2008) Statystyka moja profesja. Wiadomości Statystyczne, 53(6), s. 17-28.

Rao C. R. (1994) Statystyka i prawda. Wydawnictwo Naukowe PWN, Warszawa.

Rotengruber P. (2001) Kłamstwo gospodarcze a interesy producenta. Annales. Etyka w Życiu Gospodarczym, 4, 117-125, Salezjańska Wyższa Szkoła Ekonomii i Zarządzania.

Sedláček T. (2012) Ekonomia dobra i zła: w poszukiwaniu istoty ekonomii od Gilgamesza do Wall Street. Wydawnictwo Studio EMKA, Warszawa.

Stawrowski Z. (2023) Przede wszystkim prawda, wartość uniwersalna [w:] Tyrania postępu, Wydawnictwo Biały Kruk, Kraków.

Wyżnikiewicz B. (2001) Wysoka cena prawdy w statystyce. Wiadomości Statystyczne, 66(1), 49-56.

THE PURSUIT OF TRUTH AS A SOURCE OF THE STATISTICIAN'S ETHICS

Abstract: The ethical attitude of a statistician is the pursuit of truth and the avoidance and elimination of internal and external influences and pressures on the reliable final effect of research. The aim of the article is to determine the sources that guarantee an ethical attitude, which include: the pursuit of truth, professionalism, and research independence. The most important message of "ethical principles" is the pursuit of truth. When discussing the problem related to the ethics of a statistician, one should take into account the dangers, because statistical information has become a "commodity" partly due to the rapid development of electronic technology, which creates opportunities for manipulating this information.

Keywords: pursuit of truth, statistician's independence

JEL classification: C18, C69, H69

Quantitative Methods in Economics
Volume XXV, No. 3, 2024, pp. 107 – 123

IF MULTILAYER PERCEPTRON NETWORK MAY HELP IN MULTIVARIATE EPS FORECASTING. EVIDENCE FROM POLAND

Wojciech Kuryłek https://orcid.org/0000-0003-0692-3300
Faculty of Management
University of Warsaw
e-mail: wkurylek@wz.uw.edu.pl

Abstract: This study examines the role of accurate earnings forecasts for publicly traded firms in achieving investment success, focusing on markets with limited analyst coverage, such as Poland. It compares the accuracy of various models, including artificial neural networks, against a seasonal random walk model applied to EPS data from Warsaw Stock Exchange companies (2008-2019). The seasonal random walk model showed the lowest error based on MAAPE, with results confirmed by statistical tests. Simpler models may outperform complex ones due to overfitting and the relatively simple dynamics of Polish companies.

Keywords: earnings per share, seasonal random walk, artificial neural network, multilayer perceptron network, financial forecasting, Warsaw Stock Exchange

JEL classification: C01, C02, C12, C14, C58, G17

INTRODUCTION

The pricing of company stocks hinges on the multiplication of earnings per share (EPS) by the Price-to-Earnings multiple, a pivotal step in investment deliberations. Accurate forecasting of these components is crucial, with EPS predictions assuming particular importance. They furnish indispensable numerical insights into a company's future trajectory, furnishing valuable data on potential market valuation and guiding auditing expectations. The estimates of Earnings Per Share (EPS) by popular financial information services like Eikon (Refinitiv), Morningstar, Bloomberg, and others are consensus forecasts, which aggregate the forecasts of multiple equity analysts. These analysts, who may be part of banks,

https://doi.org/10.22630/MIBE.2024.25.3.10

brokerage firms, or independent financial research companies, provide their EPS predictions for the current year, next year, and sometimes over longer periods. While financial analysts extensively cover companies in developed markets like the US, according to the data from the EquityRT platform (which sources the data to Morningstar), only a fraction, approximately 20%, receive similar attention in emerging markets such as Poland. Consequently, there exists a compelling necessity to employ statistical or machine learning models for EPS forecasting.

This article undertakes a comparative assessment of various models, employing distinct sets of explanatory variables, utilizing the multilayer perceptron (MLP) artificial neural network, drawing insights from Li and Mohanram's [2014] research. It encompasses quarterly EPS data for 267 companies listed on the Polish stock exchange from the 2008-2009 financial crisis through the 2020 pandemic.

Rather than relying solely on the conventional mean absolute percentage error (MAPE) metric, which is prone to extreme values when the denominator is small, an alternative measure, the mean arctangent absolute percentage error (MAAPE) proposed by Kim and Kim [2016], is computed and employed in this study.

In summary, this article pursues several objectives. Firstly, it aims to assess the performance of the multilayer perceptron network (MLP) over different sets of explanatory variables in EPS prediction. Secondly, it seeks to apply diverse error metrics, varying timeframes, and a range of statistical tests to validate the outcomes of these experiments. Thirdly, it endeavors to adapt and utilize a relative performance error metric to address scenarios where actual profits approach zero, employing MAAPE as an error metric. Lastly, it strives to elucidate the practical implications of these findings for investment strategies in Polish stocks.

LITERATURE OVERVIEW

The algorithmic forecasting of Earnings per Share began in the 1960s, sparking scholarly exploration focused on autoregressive integrated moving average (ARIMA) models (see [Ball and Watts 1972, Watts 1975, Griffin 1977, Foster 1977, Brown and Rozeff 1979]). This marked the primary class of scrutinized models, with outcomes varying across investigations: while some studies supported the simplicity of the basic random walk model, suggesting that more intricate models did not consistently surpass it, others yielded divergent conclusions. Kurylek [2023a, 2023b] conducted a similar study regarding the Polish market.

Over time, however, a consensus emerged favoring ARIMA-type models for their typically precise forecasts (see [Lorek 1979, Bathke and Lorek 1984]). This consensus lasted until the late 1980s when a prevailing belief suggested that financial analysts' forecasts surpassed those generated by time series models (see [Brown et al. 1987]). Nevertheless, Conroy and Harris [1987] noted analysts' superiority in short forecast horizons, diminishing over longer periods. This perspective endured until recent years when the superiority of analysts over time series models was once again scrutinized (see [Lacina et al. 2011, Bradshaw et al. 2012, Pagach and Warr 2020, Gaio et al. 2021]).

Furthermore, since the late 1960s, researchers explored various approaches employing exponential smoothing for EPS prediction (see [Elton and Gruber 1972, Ball and Watts 1972, Johnson and Schmitt 1974, Brooks and Buckmaster 1976, Ruland 1980, Brandon et al. 1987, Jarrett, 2008]), resulting in mixed findings.

Lorek and Willinger [1996] demonstrated the superiority of multivariate cross-sectional models over firm-specific and common-structure ARIMA models. Lev and Thiagarajan [1993] identified 12 fundamental signals from financial ratios, subsequently utilized by Abarbanell and Bushee [1997] for EPS forecasting. Cao and Gan [2009], Cao and Parry [2009], Ahmadpour et al. [2015], and Ball and Ghysels [2017] employed similar fundamental variables for EPS multivariate forecasting using neural networks, affirming their effectiveness.

Ohlson [1995, 2001] formulated a residual earnings model, while Pope-Wang [2005, 2014] established theoretical frameworks linking earnings forecasts to accounting variables and stock prices. Li [2011] developed a model for forecasting earnings for loss-making firms, demonstrating its efficacy. Lev and Souginannis [2010] provided evidence of the usefulness of estimate-based accounting items for predicting next year's earnings, albeit with limited success in subsequent years. Hou et al. [2012] achieved substantial R2 coefficients in cross-sectional regression models for earnings forecasting. Li and Mohanram [2014] compared various models, revealing the superiority of some over others. Harris and Wang [2019] found Pope and Wang's [2005] model generally less biased and more accurate.

Recent research has placed significant emphasis on the utilization of artificial neural networks in EPS forecasting, yielding ambiguous results. A pioneering study by Atiya, Shaheen, and Talaat [1997] demonstrated the superiority of a neural network based on fundamental characteristics for stock price forecasting, establishing its consistent superiority. Cao et al. [2004] conducted a comparative analysis of neural feedforward networks (MLP) and found them to outperform other forecasting models, showcasing their enhanced accuracy. However, contrary findings were presented by Lai and Li [2006], who discovered that an ANN model exhibited the worst accuracy when predicting EPS. In another study, Cao and Parry [2009] consistently demonstrated the superiority of univariate neural network models over linear regression models. They further revealed that a genetic algorithm outperformed backpropagation in estimating natural network weights. Similarly, Cao and Gan [2009] confirmed the superior performance of neural network models, especially when optimized using a genetic algorithm, for predicting the EPS of Chinese listed companies. Gupta, Khirbat, and Singh [2013] identified an optimal architecture for stock market price forecasting using multiperceptron networks, highlighting the critical role of factors like EPS and public confidence in predictions. Ahmadpour, Etemadi, and Moshashaei [2015] utilized a standard multilayer perceptron (MLP) neural network with remarkable success, with extracted rules exhibiting significantly greater accuracy than pure MLP models. Chen et al. [2020]

explored various methods for EPS prediction, including decision trees and radial basis function networks, demonstrating the superiority of the ensemble method in terms of accuracy. Elend et al. [2020] compared LSTM networks to TCNs for predicting future EPS, with LSTM outperforming the naive persistent model, showcasing a significant improvement in prediction accuracy. Additionally, Suler, Vochozka, and Vrbka [2020] successfully employed an LSTM neural network model for bankruptcy prediction in the Czech Republic. Furthermore, Xiaoqiang's [2022] article provides a concise overview of deep learning and machine learning techniques, including convolutional neural networks and decision trees, applicable to EPS forecasting. In the latest research, Dreher et al.]2024] illustrated that considering accounting information on tax loss carryforwards did not enhance EPS forecasts and often deteriorated predictions in out-of-sample tests for German listed companies, utilizing tax footnotes information.

DATA & METHODOLOGY

Data

The Polish stock market, which became part of the European Union after 2004, is characterized by its substantial depth, boasting a market capitalization that soared to $197 billion and accommodating 774 listed companies by the conclusion of 2021. However, it's notable that these stocks do not receive the extensive analyst coverage observed in the United States or Western Europe. In 2019, merely around 20% of the 711 listed companies garnered attention from analysts. This highlights the imperative for statistical forecasting of crucial financial data utilizing analytical methodologies. This article primarily concentrates on the earnings per share (EPS) data series and other financial explanatory variables obtained from EquityRT, a financial analysis platform. The analysis probes into EPS patterns of companies listed on the Warsaw Stock Exchange from Q1 2010 to Q4 2019, i.e. between significant structural shifts: the 2008-2009 financial crisis and the onset of the COVID-19 pandemic in 2020. For forecasting purposes, data from Q1 2010 to Q4 2018 (36 quarters) are utilized for model estimations, while Q1 2019 to Q4 2019 data are set aside for out-of-time validation testing. Forecast horizons extend from 1 to 4 quarters ahead, with additional examination incorporating the years 2017 and 2018 as validation samples. Following thorough coverage, excluding splits and reverse splits, the dataset retains 267 companies.

Explanatory variables

Redundancy among financial ratios, frequently encountered when elucidating economic phenomena, demands approaches to refine and isolate a distinct, autonomous subset of vital financial variables applicable for EPS modeling. The ensuing models were utilized:

- The seasonal random walk model (SRW)

This process can be described as:

$$EPS_t = EPS_{t-4} + \varepsilon_t \text{ where } \varepsilon_t \text{ are IID and } \varepsilon_t \sim N(0, \sigma^2) \quad (1)$$

The forecast denoted as, $\widehat{EPS}_t = EPS_{t-4}$ relies on the value delayed by 4 quarters as the prediction, thereby obviating the necessity for parameter estimation. This model acts as a benchmark, based on research by Kuryłek [2023a, 2023b], demonstrating its superiority over time series models in the context of Poland.

- The Laurent model (L)

Laurent [1979] condensed a set of 45 financial ratios into 10 factors using principal component analysis, collectively explaining nearly 90% of the observed variance. Through a meticulous selection process, ten specific financial ratios were identified to represent these factors. These ratios offer a comprehensive insight into a firm's financial performance from various angles. Selection criteria included a strong correlation with the represented factor, independence from each other, and widespread acceptance in usage. The resulting set of ratios covers diverse aspects, such as Profit before interest and tax to total assets (R1), Long-term debt to total assets (R2), Revenue to working capital (R3), Revenue to fixed assets (R4), Revenue to shareholders' funds (R5), Revenue to inventory (R6), Revenue to debt (R7), Quick liquidity ratio (R8), Profit before interest and taxes to interest (R9), and Reserves to net income (R10). These ratios serve as explanatory variables for EPS. Consequently, the estimated equation can be formulated as follows:

$$EPS_{t+4} = f(R1_t, R2_t, .., R10_t) + \varepsilon_t \quad (2)$$

- The Lev and Thiagarajan model (LT)

The research conducted by Lev and Thiagarajan [1993] identified 12 fundamental signals extracted from various practical texts dedicated to utilizing financial ratios for delineating different facets of a firm's performance. This approach stands as one of the most frequently referenced methodologies in the literature, as evidenced by Ahmadpour et al. [2015], Ball and Ghysels [2017], and others. Key variables that proved useful include Inventory (I), Accounts receivable (AR), Capital expenditure (CAPEX), Gross margin (GM), Sales and administrative expenses (SAE), Provision for doubtful receivables (PROV), Effective tax (ET), and Labor intensity of sales (LP). However, it's important to note that some variables that were mentioned in the initial research are missing in Poland. Polish listed companies do not disclose data on R&D expenditures and order backlog. Furthermore, the database lacks information on the chosen FIFO or LIFO accounting standard, and it is impossible to discern auditor qualifications. Therefore, the equation akin to Lev and Thiagarajan's, structured in quarterly terms for making one-year-ahead predictions, can be articulated as follows:

$$E_{t+4} = f(I_t, AR_t, CAPEX_t, GM_t, SAE_t, PROV_t, ET_t, LP_t) + \varepsilon_t \quad (3)$$

- The Residual Income model (RI)

Ohlson [1995, 2001], Olhson and Juettner-Nauroth [2005] developed a theoretical model of residual earnings grounded on the clean surplus assumption. This assumption posits that the market value of equity equals the book value plus the net present value of future abnormal returns. Residual income, or abnormal earnings, is quantified as the disparity between actual earnings and book value, multiplied by the cost of capital, and is presumed to follow an autoregressive stochastic process. As a result, future earnings are influenced by current earnings, book value, dividends paid, and the cost of capital, leading to the incorporation of the following accounting variables: Delayed earnings (E), a Dummy variable for firms with delayed negative earnings (NegE), a Polynomial term combining two preceding variables (NegE * E), Book value (BV), and Accruals (ACC). The specification of this model is provided by the research conducted by Li and Mohanram [2014]. With quarterly data and the application of this approach, the equation assumes the following structure:

$$E_{t+4} = f(E_t, NegE_t, BV_t, ACC_t) + \varepsilon_t \quad (4)$$

- The Pope and Wang model (PW)

Based on the assumptions of linear valuation, no-arbitrage, dividend irrelevance, and clean surplus accounting, Pope and Wang [2005, 2014] expounded upon the theoretical linkage between earnings forecasts and six observable accounting variables. These variables include Earnings (E), Book value (BV) lagged by one and two years, Accruals (ACC), and non-accounting variables such as Stock price (P) lagged by one and two years. The model's equation formulation was detailed by Harris and Wang (2019). For quarterly data aiming to predict earnings one year ahead, the expression is presented as follows:

$$E_{t+4} = f(E_t, BV_t, BV_{t-4}, ACC_t, P_t, P_{t-4}) + \varepsilon_t \quad (5)$$

- The Earnings Persistence model (EP)

Li [2011] formulated the earnings persistence model to account for earnings growth, integrating explanatory variables including delayed Earnings (E), the indicator for negative earnings (NegE), and the interaction term between NegE and earnings (NegE*E). This facilitates the discernment of profit and loss persistence. The equation specification for this model is outlined in the research by Li and Mohanram (2014). It can be represented in the following concise functional form:

$$E_{t+4} = f(E_t, NegE_t) + \varepsilon_t \quad (6)$$

- The Hou, van Dijk and Zhang model (HDZ)

Hou, van Dijk, and Zhang [2012] performed a cross-sectional regression analysis aimed at predicting earnings. They utilized data spanning the previous decade and included variables such as delayed Earnings (E), Total assets (A), Dividend payout (D), Accruals (ACC), a Dummy variable for dividend payers (DD), and a Dummy variable for firms with delayed negative earnings (NegE). Their efforts resulted in a substantial R2 coefficient of around 0.8. The equation introduced by Li and Mohanram (2014) is applied in the following manner:

$$E_{t+4} = f(E_t, NegE_t, ACC_t, A_t, D_t, DD_t) + \varepsilon_t \tag{7}$$

To forecast earnings per share (EPS), one divides future earnings by the constant number of shares. Because in the Pope Wang models 4 and 8 lags are used, 7 476 observations (28 quarters x 267 companies) observations are used for training the artificial network in 2019.

The multilayer perceptron network (MLP)

The artificial networks presented in this study were trained using the TensorFlow module in Python. These networks are of the feedforward type, meaning that data flows uni-directionally from the input layer to the output layer. Artificial neural networks (ANNs) are commonly employed for analyzing cause-effect relationships within complex systems, often in the context of big data frameworks. However, they can also be applied to small datasets, as demonstrated in fields like health sciences by Pasini [2015], as is the case in this article. Hyperparameters, such as the width and depth of networks (i.e., the number of neurons in each layer and the number of layers), were optimized using the hyperas library in Python. The models were trained using the backpropagation algorithm based on gradient descent, employing only 20 epochs (where one epoch constitutes one complete run-through of the training set). Backpropagation, popularized by Werbos [1988] in the late 80s, is a standard method for learning neural networks, involving the backward propagation of errors. It fine-tunes the weights of a neural network based on the error rate obtained in the previous epoch. The error, calculated as the difference between the predicted and actual output, is then fed back through the network. The weights are adjusted accordingly to minimize the error, with the learning rate determining the rate of adjustment. Proper tuning of the weights aims to reduce error rates and enhance the model's generalization ability. After a certain number of epochs, the algorithm converges to a state where there's minimal change in loss over subsequent epochs, typically reaching a local optimum of the defined loss function. Input parameters are usually standardized for ANNs when dealing with multivariate data. In all analyzed models, the hyperbolic tangent (tanh) activation function, a popular choice, was used in all layers. Additionally, the weights between layers were initialized using the glorot_uniform initializer, proposed by Bengio and Glorot [2010], which generates initial weight values from a uniform distribution. Further insights into different network architectures and parameters can be found in the book by Bengio et al. [2017].

A multilayer perceptron (MLP) represents a form of artificial neural network structured with multiple layers of interconnected nodes. Each layer's nodes establish connections with those in the subsequent layer, with the connection weights learned during training. Typically, an MLP comprises three layers or more: an input layer, one or more hidden layers, and an output layer. Within each hidden layer, the output of each node results from a weighted sum of the preceding layer's node outputs, augmented by a bias term. The inception of MLP neural networks dates back to 1958 when Rosenblatt [1958] introduced a layered network of perceptrons. It featured an input layer, a hidden layer with fixed weights that didn't adapt, and an output layer with learning connections. Rosenblatt drew inspiration from the brain's functionality. The number of layers and neurons constitutes the network's hyperparameters, subject to fine-tuning. While deeper neural networks excel in data processing, excessively deep layers can engender challenges like vanishing gradients and overfitting. Empirical rules of thumb guide the determination of the optimal number and size of hidden layers, as detailed in Heaton's [2008] book. According to this source, a single hidden layer suffices to approximate any function. Consequently, the network in this study was designed with one hidden layer. Additionally, a widely endorsed guideline suggests that the hidden layer's optimal size should lie approximately between that of the input and output layers. In this instance, the hidden layer's size equates to the mean of the sizes of the input and output layers.

Mean Arctangent Absolute Percentage Error (MAAPE)

Denoting $A_1^i, \dots, A_4^i$, as the actual earnings per share (EPS) from the first to the fourth quarter of 2019 for a specific firm I, and $F_1^i, \dots, F_4^i$ as the corresponding forecasts (i.e. $\hat{Q}_t$, where t=37,..,40 for i-th company), the absolute percentage error (APE) of such prediction during the j-th quarter of 2019, for any firm i, can be expressed as:

$$APE_j^i = \left| \frac{A_j^i - F_j^i}{A_j^i} \right| \tag{8}$$

However, the absolute percentage error (APE) exhibits a significant limitation: it may result in infinite or undefined values when the actual figures approach or reach zero, a situation frequently encountered in earnings forecasts. Moreover, extremely low actual figures, typically below one, can result in substantial percentage errors (outliers). Furthermore, when actual values are zero, APE becomes infinite. To address this issue, Kim and Kim (2016) proposed the arctangent absolute percentage error as a novel solution in the domain.

$$AAPE_j^i = \begin{cases} 0 & if\ A_j^i = F_j^i = 0 \\ arctan\left(\left|\frac{A_j^i - F_j^i}{A_j^i}\right|\right) & otherwise \end{cases} \tag{9}$$

This reasoning stems from the characteristic of the arctan function, which maps values from the range of [-∞,+∞] to the interval [-π/2,π/2]. As a result, the Mean Arctangent Absolute Percentage Error (MAAPE) for the j-th quarter among all I companies in the dataset can be expressed as

$$MAAPE_j = \frac{1}{I}\sum_{i=1}^{I} AAPE_j^i = \frac{1}{I}\sum_{i=1}^{I} arctan\left(\left|\frac{A_j^i - F_j^i}{A_j^i}\right|\right) \quad (10)$$

The decision to opt for MAAPE over MAPE (Mean Absolute Percentage Error) was intentional because of the inclusion of companies with actual profits very close to zero in the analyzed sample. When only one observation approaches zero while others are substantially distant from it, the MAPE of this specific observation can become extremely large (almost infinite), potentially dominating the mean calculation and making the rest of observations negligible.

The statistical test

To assess the statistical significance of MAAPE variations among multiple models, a nonparametric two-sided Wilcoxon test, as detailed by Wilcoxon (1945), is employed. This test serves as a paired difference test for two matched samples. It's important to highlight that this test does not require specific assumptions regarding a probability distribution, except for the symmetry of the difference in scores and the independence of these differences. Ruland (1980) extensively explained the application of the Wilcoxon test in validation, particularly in determining whether errors generated by different EPS models display statistical differences. Separate tables containing p-values are generated for each quarter, ranging from one to four, as well as for all quarters collectively.

$$H_o: medians\ of\ AAPEs\ of\ a\ pair\ of\ models\ are\ the\ same \quad (11)$$

If the p-values of each test fell below the predetermined significance threshold of 0.05, the null hypothesis for each test was considered invalid. This principle, widely utilized, draws from various sources, including Ruland (1980).

RESULTS

Empirical findings

The seasonal random walk (SRW) model, as outlined in Table 1, consistently outperforms all other models estimated within the multilayer perceptron (MLP) framework across every quarter, demonstrating superior overall performance. Conversely, the Residual Income model (RI) displays the poorest performance, while the Laurent model (L) achieves the second-best performance. The MAAPEs of all other models fall in between, at comparable levels.

To assess whether the errors of the top-performing model differ significantly from those of the other models, the Wilcoxon nonparametric test was utilized to compare the AAPE medians between the SRW model and all other models. As depicted in Table 2, the findings reveal that the seasonal random walk (SRW) model consistently exhibits statistically lower errors compared to the other models across all analyzed periods. However, in the fourth quarter of 2019, the significance levels of the Laurent model (L), the Lev and Thiagarajan model (LT), and the Pope Wang model (PW) are in proximity to the 0.05 threshold.

Table 1. Summary statistics on forecast errors for 2019 quarters

model	Q1 MAAPE	Q2 MAAPE	Q3 MAAPE	Q4 MAAPE	Total MAAPE
SRW	0.658	0.702	0.653	0.736	0.687
L	0.785	0.782	0.785	0.785	0.785
LT	0.785	0.786	0.788	0.791	0.787
RI	1.016	0.965	0.986	0.917	0.971
PW	0.788	0.788	0.791	0.790	0.789
EP	0.930	0.880	0.898	0.859	0.892
HDZ	0.923	0.871	0.891	0.874	0.890

Source: own calculations

Table 2. P-values of the Wilcoxon test of forecast errors for SRW and respective models in 2019

quarter	model	L	LT	RI	PW	EP	HDZ
1	SRW	0.000	0.000	0.000	0.000	0.000	0.000
2	SRW	0.004	0.005	0.000	0.004	0.000	0.000
3	SRW	0.000	0.000	0.000	0.000	0.000	0.000
4	SRW	0.046	0.046	0.000	0.049	0.000	0.000
ALL	SRW	0.000	0.000	0.000	0.000	0.000	0.000

Source: own calculations

Robustness checks

Robustness assessments were undertaken across different years and various prevalent error metrics. Notably, across all scrutinized years—2017, 2018, and 2019—the seasonal random walk model (SRW) consistently produced superior forecasts compared to alternative models, as depicted in Table 3. In both 2017 and 2018, the least effective model was the simplest among all multivariate models, the Earnings Persistence model (EP). However, in 2019, as previously noted, the Residual Income model (RI) exhibited the highest forecast errors. Furthermore, the Wilcoxon test was employed to compare all model pairs alongside the seasonal random walk model, and the corresponding p-values for each year are delineated in Table 4. Across each of these years, the seasonal random walk model (SRW) demonstrated statistically superior outcomes compared to alternative methods. Hence, the consistent dominance of the seasonal random walk model becomes apparent over time.

Table 3. Summary statistics on forecast errors for all quarters 2017–2019

		2017	2018	2019
		MAAPE	MAAPE	MAAPE
model	SRW	0.686	0.711	0.687
	L	0.785	0.790	0.785
	LT	0.790	0.784	0.787
	RI	0.784	0.785	0.971
	PW	0.785	0.785	0.789
	EP	0.852	0.953	0.892
	HDZ	0.809	0.829	0.890

Source: own calculations

Table 4. P-values of paired Wilcoxon test of forecast errors for all quarters 2017–2019 and SRW model

year	model	L	LT	RI	PW	EP	HDZ
2017	SRW	0.000	0.000	0.000	0.000	0.000	0.000
2018	SRW	0.000	0.001	0.001	0.000	0.000	0.000
2019	SRW	0.000	0.000	0.000	0.000	0.000	0.000

Source: own calculations

Table 5 illustrates an assessment of the performance of the examined models using alternative error metrics: Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). This assessment encompasses all quarters aggregated for the year 2019. To facilitate a fair comparison, these metrics underwent adjustment for Consumer Price Index (CPI) inflation. This adjustment ensures parity in the present value of future errors in nominal terms with current errors. Consistent with previous observations in 2019, the seasonal random walk model demonstrated the lowest errors across all metrics, encompassing both RMSE and MAE.

The p-values derived from Table 6, per the Wilcoxon test, highlight significant disparities between the outcomes of the SRW model and other model pairings. This indicates that the forecasts produced by the seasonal random walk (SRW) model, in terms of both RMSE and MAE, exhibit superior performance and statistical distinctiveness compared to all other models implemented via the multilayer perceptron (MLP) methodology.

Table 5. Summary statistics on forecast errors for RMSE and MAPE in all quarters 2019

	SRW	L	LT	RI	PW	EP	HDZ
RMSE	0.937	1.334	1.327	1.501	1.346	1.352	1.362
MAE	0.705	1.105	1.097	1.247	1.116	1.117	1.125

Source: own calculations

Table 6. P-values of paired Wilcoxon test of forecast errors for RMSE and MAE in 2019

measure	model	L	LT	RI	PW	EP	HDZ
RMSE	SRW	0.000	0.000	0.000	0.000	0.000	0.000
MAE	SRW	0.000	0.000	0.000	0.000	0.000	0.000

Source: own calculations

Discussion

The relatively inferior performance of more intricate models employing artificial neural networks can be attributed to overfitting, which leads to unstable relationships among variables contingent on the pertinent test dataset. The utilization of such relationships in making predictions is reasonable only if the statistical relationship is sufficiently robust (see [Lev and Souginannis 2010]). This assertion is consistent with the findings of Dreher et al. [2024], who also demonstrated for German listed companies that complex deep learning approaches, which optimize explanatory power within the sample, do not fare well for out-of-sample prediction. These sophisticated models risk overparameterizing the market's straightforward behavior, resulting in larger forecast errors.

The rationale behind the superior performance of simpler models may align with the Polish scenario, as advanced models often tend to be overly intricate, possessing an excess of parameters to describe relatively straightforward economic phenomena. This observation further corroborates the research conducted by Kurylek [2023a, 2023b], which showed that even basic models like ARIMA and exponential smoothing, effective for the US market, were outperformed by the simple seasonal random walk model in Poland. This reinforces the hypothesis that the inherent simplicity of the Polish stock market likely underpins the effectiveness of the seasonal random walk (SRW) model, or alternatively, additional calibration for out-of-sample predictions might be necessary.

Hence, straightforwardly applying any of these sophisticated techniques beyond the conventional seasonal random walk in Poland for EPS forecasting in investment contexts appears impractical. Furthermore, considering that EPS behavior follows a seasonal random walk and acknowledging that stock prices are derived from the multiplication of the P/E multiple by EPS, one might infer that stock prices exhibit at least as much randomness as EPS. Since EPS behavior, characterized by a random walk, is inherently challenging, accurately predicting stock prices for a period extending at least one quarter ahead becomes even more daunting.

In shorter timeframes, when EPS remains constant, stock price forecasting behaves similarly to P/E multiples. Consequently, exploring methods to forecast P/E multiples for periods shorter than one quarter, occurring between the publication of quarterly financial reports, could be of significant interest from an investment perspective. The forecast generated by the seasonal random walk (SRW) essentially represents a value from the corresponding quarter of the previous year. This implies that for predicting future prices, even over extended horizons, the P/E multiple might carry more significance than next year's earnings of companies (EPS).

This aligns with economic theory, which suggests that the P/E multiple is influenced by expected future earnings growth, future interest rates, and market sentiment or premium reflecting investor risk appetite (i.e., market sentiment), whereas EPS forecasts pertain only to near-future earnings. In both short-term and long-term contexts, the conclusion is clear: for investment, the P/E multiple holds greater importance than EPS prediction.

Conclusions

This study assesses the predictive performance of seven models: the seasonal random walk (SRW), the Laurent model (L), the Lev and Thiagarajan model (LT), the Residual Income model (RI), the Pope and Wang model (PW), the Earnings Persistence model (EP), and the Hou, van Dijk, and Zhang model (HDZ). The forecasting of EPS holds significant value in emerging markets, where coverage of listed firms by financial analysts is sparse, as evidenced by Poland's case. When applied to quarterly EPS data from 267 Polish companies spanning 2010 to 2019, the SRW model consistently demonstrated the lowest error, offering a more accurate portrayal of the Polish market compared to other models. Furthermore, the SRW model consistently surpassed other models across different periods and error metrics like RMSE or MAE. This trend is supported by Wilcoxon tests and can be attributed to the over-parameterization of complex models, their tendency to overfit, and the relatively simple nature of the Polish stock market.

The practical implication of this research suggests that utilizing techniques more sophisticated than the standard seasonal random walk for EPS forecasting in Poland lacks practical merit. However, relying on the seasonal random walk for EPS modeling implies that forecasted stock prices may exhibit significant randomness, posing challenges for prediction. Hence, forecasting the P/E multiple might be more critical than predicting EPS for future stock price forecasts, especially in shorter investment horizons when EPS remains constant.

Future research could explore the relationship between forecasting accuracy and firm size, with industry sector analysis potentially influencing the choice of the most suitable model for EPS forecasting. Investigating time series transformations to normalize EPS distributions could offer valuable insights. Additionally, a broader set of explanatory variables warrants exploration. Comparing the predictive accuracy of the best algorithmic model with forecasts from market analysts presents an intriguing avenue. Furthermore, evaluating the performance and accuracy of various predictive models and financial analysts' projections during economic downturns, such as the 2008-2009 financial crisis or the COVID-19 pandemic, could yield valuable insights. Identifying seasonal patterns through the SRW model may offer insights into investment strategies, potentially challenging the "weak form" of the Efficient Market Hypothesis (EMH).

REFERENCES

Abarbanell J., Bushee B. (1997) Fundamental Analysis, Future EPS, and Stock Prices. Journal of Accounting Research, 35(1), 1-24.

Ahmadpour A., Etemadi H., Moshashaei S. (2015) Earnings per Share Forecast using Extracted Rules from Trained Neural Network by Genetic Algorithm. Computational Economics, 46(1), 55-63.

Atiya A., Shaheen S., Talaat N. (1997) An Efficient Stock Market Forecasting Model using Neural Networks. IEEE International Conference on Neural Networks - Conference Proceedings.

Ball R., Ghysels E. (2017) Automated Earnings Forecasts: Beat Analysts or Combine and Conquer? Management Science, 64(10), 4936-4952.

Ball R. Watts R. (1972) Some Time Series Properties of Accounting Income. The Journal of Finance, 27(3), 663-681.

Banerjee P. (2020) A Guide on XGBoost Hyperparameters Tuning, Accessed February 14, 2024. https://www.kaggle.com/code/prashant111/a-guide-on-xgboost-hyperparameters-tuning.

Bathke Jr. A. W., Lorek K. S. (1984) The Relationship between Time-Series Models and the Security Market's Expectation of Quarterly Earnings. The Accounting Review, 59(2), 163-176.

Bradshaw M., Drake M., Myers J., Myers L. (2012) A Re-Examination of Analysts' Superiority over Time-Series Forecasts of Annual Earnings. Review of Accounting Studies, 17(4), 944-968.

Bengio Y., Courville A., Goodfellow I. (2017) Deep Learning. Cambridge, Massachusetts: The MIT Press.

Bengio Y., Glorot X. (2010) Understanding the Difficulty of Training Deep Feedforward Neural Networks. Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, 9, 249-256.

Brandon Ch., Jarrett J. E., Khumawala S. B. (1987) A Comparative Study of the Forecasting Accuracy of Holt-Winters and Economic Indicator Models of Earnings Per Share For Financial Decision Making. Managerial Finance, 13(2), 10-15.

Brooks L. D., Buckmaster D. A. (1976) Further Evidence of the Time Series Properties of Accounting Income. The Journal of Finance, 31(5), 1359-1373.

Brown L. D., Griffin P. A., Hagerman R. L., Zmijewski M. E. (1987) Security Analyst Superiority Relative to Univariate Time-Series Models in Forecasting Quarterly Earnings. Journal of Accounting and Economics, 9(1), 61-87.

Brown L. D., Rozeff M. S. (1979) Univariate Time-Series Models of Quarterly Accounting Earnings per Share: A Proposed Model. Journal of Accounting Research, 17(1), 179-189.

Cao Q., Gan Q. (2009) Forecasting EPS of Chinese Listed Companies using a Neural Network with Genetic Algorithm. 15th Americas Conference on Information Systems 2009, AMCIS 2009, 2791-2981.

Cao Q., Parry M. (2009) Neural Network Earnings per Share Forecasting Models: A Comparison of Backward Propagation and the Genetic Algorithm. Decision Support Systems, 47 (1), 32-41.

Cao Q., Schniederjans M. J., Zhang W. (2004) Neural Network Earnings per Share Forecasting Models: A Comparative Analysis of Alternative Methods. Decision Sciences, 35 (2), 205-237.

Chen Y., Chen S., Huang H., Sangaiah A. (2020) Applied Identification of Industry Data Science using an Advanced Multi-Componential Discretization Model. Symmetry, 12(10), 1-28.

Chen T., Guestrin C. (2016) XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785-794.

Conroy R., Harris R. (1987) Consensus Forecasts of Corporate Earnings: Analysts' Forecasts and Time Series Methods. Management Science, 33(6), 725-738.

Dreher S., Eichfelder S., Noth F. (2024) Does IFRS Information on Tax Loss Carryforwards and Negative Performance Improve Predictions of Earnings and Cash Flows? Journal of Business Economics, 94(1), 1-39.

Elend L., Kramer O., Lopatta K., Tideman S. (2020) Earnings prediction with deep learning. German Conference on Artificial Intelligence (Künstliche Intelligenz), KI 2020: Advances in Artificial Intelligence, 267-274.

Elton E. J., Gruber M. J. (1972) Earnings Estimates and the Accuracy of Expectational Data, Management Science, 18(8), B409-B424.

Foster G. (1977) Quarterly Accounting Data: Time-Series Properties and Predictive-Ability Results. The Accounting Review, 52(1), 1-21.

Gaio L., Gatsios R, Lima F., Piamenta Jr. T. (2021) Re-Examining Analyst Superiority in Forecasting Results of Publicly-Traded Brazilian Companies. Revista de Administracao Mackenzie, 22(1), eRAMF210164.

Griffin P. (1977) The Time-Series Behavior of Quarterly Earnings: Preliminary Evidence. Journal of Accounting Research, 15(1), 71-83.

Gupta R., Khirbat G., Singh S. (2013) Optimal Neural Network Architecture for Stock Market Forecasting. Proceedings - 2013 International Conference on Communication Systems and Network Technologies, CSNT 2013, 557-561.

Harris R. D. F., Wang P. (2019) Model-Based Earnings Forecasts vs. Financial Analysts' Earnings Forecasts. British Accounting Review, 51(4), 424-437.

Heaton J. (2008) Introduction to Neural Networks for Java, 2nd Edition. Heaton Research Inc.

Hou K., van Dijk M., Zhang Y. (2012) The Implied Cost of Capital: A New Approach. Journal of Accounting and Economics, 53(3), 504-526.

Jarrett J. E. (2008) Evaluating Methods for Forecasting Earnings per Share. Managerial Finance, 16, 30-35.

Johnson T. E., Schmitt T. G. (1974) Effectiveness of Earnings per Share Forecasts. Financial Management, 3(2), 64-72.
Kim S., Kim H. (2016) A New Metric of Absolute Percentage Error for Intermittent Demand Forecasts. International Journal of Forecasting, 32(3), 669-679.
Kuryłek W. (2023a) The Modeling of Earnings per Share of Polish Companies for the Post-Financial Crisis Period using Random Walk and ARIMA Models. Journal of Banking and Financial Economics, 1(19), 26-43.
Kuryłek W. (2023b) Can Exponential Smoothing Do Better than Seasonal Random Walk for Earnings per Share Forecasting in Poland? Bank & Credit, 54(6), 651-672.
Lacina M., Lee B., Xu R. (2011) An Evaluation of Financial Analysts and Naïve Methods in Forecasting Long-Term Earnings. [In:] K. D Lawrence, R. K. Klimberg (Eds.), Advances in business and management forecasting (pp. 77-101), Bingley, UK, Emerald.
Lai S., Li H. (2006) The Predictive Power of Quarterly Earnings per Share based on Time Series and Artificial Intelligence Model. Applied Financial Economics, 16(18), 1375-1388. http://dx.doi.org/10.1080/09603100600592752.
Laurent C. (1979) Improving the Efficiency and Effectiveness of Financial Ratio Analysis. Journal of Business Finance & Accounting, 6(3), 401-413.
Lev B., Souginannis T. (2010) The Usefulness of Accounting Estimates for Predicting Cash Flows and Earnings. Review of Accounting Studies, 15(4), 779-807.
Lev B., Thiagarajan S. (1993) Fundamental Information Analysis. Journal of Accounting Research, 31(2), 190-215.
Li K. K. (2011) How Well Do Investors Understand Loss Persistence? Review of Accounting Studies, 16(3), 630-667.
Li K. K., Mohanram P. (2014) Evaluating Cross-Sectional Forecasting Models for the Implied Cost of Capital. Review of Accounting Studies, 19(3), 1152-1185.
Lorek K. S. (1979) Predicting Annual Net Earnings with Quarterly Earnings Time-Series Models, Journal of Accounting Research, 17(1), 190-204.
Lorek K. S, Willinger G. L. (1996) A Multivariate Time-Series Model for Cash-Flow Data. Accounting Review, 71, 81-101.
Ohlson J. A. (1995) Earnings, Book Values, and Dividends in Equity Valuation. Contemporary Accounting Research, 11(2), 661-687.
Ohlson J. A. (2001) Earnings, Book Values, and Dividends in Equity Valuation: An Empirical Perspective. Contemporary Accounting Research, 18(1), 107-120. https://doi.org/10.1092/7tpj-rxqn-tqc7-ffae.
Ohlson J. A., Juettner-Nauroth B. E. (2005) Expected EPS and EPS Growth as Determinants of Value. Review of Accounting Studies, 10(2-3), 349-365.
Pagach D. P., Warr R. S. (2020) Analysts Versus Time-Series Forecasts of Quarterly Earnings: A Maintained Hypothesis Revisited. Advances in Accounting, 51, 1-15.
Pasini A. (2015) Artificial Neural Networks for Small Dataset Analysis. Journal of Thoracic Disease, 7(5), 953-960.
Pope P. F., Wang P. (2005) Earnings Components, Accounting Bias and Equity Valuation. Review of Accounting Studies. 10(4), 387-407.
Pope P., Wang P. (2014) On the Relevance of Earnings Components: Valuation and Forecasting Links. Review of Quantitative Finance and Accounting. 42, 399-413.

Rosenblatt F. (1958). The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. Psychological Review, 65(6), 386-408.

Ruland W. (1980) On the Choice of Simple Extrapolative Model Forecasts of Annual Earnings. Financial Management, 9(2), 30-37.

Suler P., Vochozka M., Vrbka J. (2020) Bankruptcy or Success? The Effective Prediction of a Company's Financial Development using LSTM. Sustainability, 12(18), 2299-2314. https://doi.org/10.3390/su12187529.

Watts R. L. (1975) The Time Series Behavior of Quarterly Earnings, Working paper, Department of Commerce, University of New Castle, April 1975.

Werbos P. (1988) Backpropagation: Past and Future. IEEE International Conference on Neural Networks, 343-353.

Wilcoxon F. (1945) Individual Comparisons by Ranking Methods. Biometrics, 1, 80-83.

Xiaoqiang W. (2022) Research on Enterprise Financial Performance Evaluation Method based on Data Mining. [In:] 2022 IEEE 2nd International Conference on Electronic Technology, Communication and Information (ICETCI). Accessed February 14, 2024. https://doi.org/10.1109/icetci55101.2022.9832404.

QUANTITATIVE METHODS IN ECONOMICS
Volume XXV, No. 3, 2024, pp. 124 – 133

APPLICATION OF ITRANSFORMERS TO PREDICTING PRETERM BIRTH RATE. COMPARISON WITH THE ARIMA MODEL

Marek Karwański https://orcid.org/0000-0001-5192-7920
Urszula Grzybowska https://orcid.org/0000-0001-7342-5382
Institute of Information Technology
Warsaw University of Life Sciences, Poland
e-mail: urszula_grzybowska@sggw.edu.pl; marek_karwanski@sggw.edu.pl
Vassilis Kostoglou
Department of Informatics and Electronic Engineering
International Hellenic University, Greece
e-mail: vkostogl@ihu.gr
Ewa Mierzejewska https://orcid.org/0000-0003-0822-4781
Katarzyna Szamotulska https://orcid.org/0000-0003-1045-7584
Department of Epidemiology and Biostatistics
Institute of Mother and Child, Warsaw, Poland
e-mail: katarzyna.szamotulska@imid.med.pl; ewa.mierzejewska@imid.med.pl

Abstract: In this paper, we study the differences between forecasts obtained with the classical seasonal ARIMA model and forecasts obtained with the neural network model called iTransformers. The analysis is done on Polish data concerning preterm birth from 2015 to 2020. We compare the results and calculate the effect size to conclude the impact of the obtained differences.

Keywords: time series forecasting, seasonal ARIMA, iTransformers, preterm birth

JEL classification: C530, I0, C630

INTRODUCTION

Time series forecasting is important in various domains, including economics or environmental and health policy. Classical methods, with their foundations laid in

https://doi.org/10.22630/MIBE.2024.25.3.11

the 70th by Box and Jenkins[1] [Box, Jenkins 1970], are still widely applied. However, they face increased competition from neural network models, especially those based on attention mechanisms [Vaswani 2017; Ahmed 2023].

We aim to compare the traditional seasonal ARIMA model used for time series forecasting with a recent approach based on Transformers. Namely, we apply a neural network architecture called iTransformers which is suitable for long-term predictions [Liu et al. 2023]. To compare the models, we calculate the differences between the predictions obtained in the statistical and machine learning approaches, test the hypothesis about the predictions, and compare the results by calculating the effect size. The analysis is done on Polish preterm birth data. The problem is important as preterm birth is responsible for a high rate of infant deaths [Brandon, McGrath 2016]. We have built forecasts for 4, 6, and 12 months. Calculations were performed using the GluonTS package with torch ver. 2.3.0+cu114.

LITERATURE REVIEW

Neural networks have been successfully used for decades to classify objects or recognize images, text, and speech. Nowadays, their applications cover virtually all fields, including time series modeling and forecasting. A relatively new survey of deep learning methods for time series forecasting, including Transforme-based architectures, can be found in, e.g., [Lim, Zohren 2021]. A list of the most important statistical and machine learning models with references can be found in a survey by [Miller et al. 2024]. This survey also contains a short description of the newest Transformer-based models and other architectures and forecast quality metrics. Traditional regression or ARIMA models have not lost their relevance for time series forecasting. Also hybrid models are applied [Aijaz, Agarwal 2019]. The data used in our analysis concerns the preterm birth rate. There has not been much research on predicting preterm birth in the last decade. Some examples of statistical modeling for preterm birth rates can be found in [Priya et al. 2024; Gemmill et al. 2021; Sefidkar et al. 2021]. On the other hand, the problem of predicting preterm birth with the help of machine learning methods can be found in [Dench, Joyce 2022; Tzitiridou-Chatzopoulou, Zournatzidou Kourakos 2024; Yu et al. 2024; Zhang et al. 2024; Zhang et al. 2023; Borboa-Olivares et al. 2023].

DATA DESCRIPTION

According to the WHO (World Health Organization), premature birth is a condition in which birth occurs before the 37th week or 259th day of its duration. It is known that the shorter the duration of pregnancy, the lower the infant's chances of healthy development and survival. In Poland, nearly 7% of pregnancies end in

[1] The famous book with many reissues summarises previous research. Historical details can be found in e. g., [Miller 2024].

premature birth in one calendar year. The data set includes information about the proportion of preterm births for the duration of pregnancy less or equal to 36 weeks in Poland for single live births from 22 weeks of gestation. We have used monthly records from 2015-2020 for the analysis[2]. Cases of stillbirth and multiple pregnancies were excluded from the study. In individual months, births ranged from 20,000 to 40,000, including 1,000 to 2,000 premature births. We have used the percentage rate of premature births in the analysis. Data exhibits seasonal effects with the highest rate in December. There is no visible trend.

METHODS

We compare the traditional ARIMA model applied to seasonally adjusted data with a Transformer-based neural network model called iTransformer. Conventional statistical models, such as regression or ARIMA models, require the data to fulfill certain assumptions. The ARIMA model requires the data to be stationary (constant mean and variance, no trend, no seasonality). Therefore, the data must be prepared before modeling, e.g., by removing stochastic and deterministic trends and seasonality. Neural network models do not require the data to meet similar assumptions. However, the data must be transformed before being fed into the network. It requires positional encoding and model architecture design [Ahmed et al. 2023]. We calculate 4, 6, and 12-month forecasts using ARIMA for X13 adjusted data and ITransformers applied to raw data. To compare the results, we calculate differences in the obtained forecast. The differences were obtained based on MIXED models with autocorrelated errors. We use the effect size measure [Lankens 2013] to better understand the differences' significance.

The classical tool for time series forecasting is the autoregressive–moving-average (ARMA) model and its generalizations which were first published in [Box, Jenkings 1970]. ARMA(p,q) is expressed as the sum of two polynomials, autoregression AR and moving average (MA).

Model $ARMA(p,q)$ can be formulated in B-operator form as:

$$W(B)X_t = V(B)\varepsilon_t$$

or

$$X_t = \frac{V(B)}{W(B)}\varepsilon_t$$

where W() and V() are polynomials of degree p and q, respectively.

In the case of series that do not satisfy the assumption of stationarity but can be reduced to stationary, the integrated moving average ARIMA(p,d,q) model is used.

$$W(B)(1-B)^d X_t = V(B)\varepsilon_t$$

[2] Data is the property of the Institute of Mother and Child. It can be obtained on request.

The parameters that refer to the non-seasonal form of the ARIMA model are: p-the number of autoregressive components (lags), d-the number of differencing applied, q-the number of lagged error components. A seasonal ARIMA model is built by including additional seasonal terms in the ARIMA model with parameters $(P,D,Q)_m$ describing the seasonal part of the model. The terms of the seasonal part of ARIMA are similar to the non-seasonal components of the model, but involve backshifts of the seasonal period. The parameters are: m-length of season, Q-the number of seasonal moving average components, P-the number of seasonal autoregressors. As a result the ARIMA requires adjusting 7 parameters. The combinations are selected based on AIC criterion.

X13 seasonal adjustment

X-13ARIMA-SEATS, is a successor to X-12-ARIMA and X-11. It is a set of statistical methods for seasonal adjustment and other descriptive analysis of time series data that are implemented in the U.S. Census Bureau's software package and are widely used to meet the requirements of forecasting models.

The X13 model assumes that a time series is a sum of components with different rates of change. In particular, it is described by four components:

- Seasonal component
- Trend component
- Cyclical component
- Random component (error)

The general form of the multiplicative model (used in this article) to describe time series looks like:

$$Y_t = Tr_t \cdot Sn_t \cdot \varepsilon_t,$$

where

Y_t - forecasted variable at time t

Tr_t - tendency (trend)

Sn_t - seasonal fluctuations at time t

ε_t - random component at time t (in the analysis of time series the structure of the random component is not considered)

In the calculations we use ARIMA(3,0,0)(0,0,1)[12] model.

Transformers based time series forecasting

Since the famous paper "Attention is All You Need" was published [Vaswani et al. 2017], the focus of the neural network community was paid solely to the new mechanism. Transformers have become the key element in almost all LLM. Numerous models using attention mechanisms have been proposed in time series forecasting, among them Informer, Autoformer, and Pyraformer. A survey on transformers for time series can be found in [Wen et al. 2023; Miller et al. 2024]. GluonTS provides various neural network models used for time series[3].

Much attention was given to the work "Are transformers effective for time series forecasting?" [Zeng et al. 2022], in which transformers application for long-term forecasting has been questioned. The recent model iTransformers is considered the latest breakthrough in long-run time series forecasting. [Liu et al. 2023]. It is regarded to be the best for long-term forecasting[4]. The main difference between this model and the common transformer approach lies in the encoding process, which is quite different; see G3 part of [Liu et al. 2023]. General Transformers' architecture has been explained in many papers. Therefore, we consider its description here unnecessary.

Effect size

Effect sizes are measures that allow the determination of sample size or to examine how big the investigated effect or difference is [Lakens 2013]. To determine if the calculated RMSE indicates the difference in models' predictions, we use Cohen's eta squared comparison given by:

$$\eta^2 = \frac{SS_{effect}}{SS_{total}},$$

where SS is the total sample (corrected) sum of squares, and SS_{effect} is the observed sum of squares due to the effect being tested.

In practice, one often adjusts the model on the training data and tests it on the test data. In this case, the predicted values can be compared with the actual data and the model performance can be assessed using goodness of fit measures (MAE) [Miller et al. 2024]. We aim not to test the model's performance but to compare the forecasts. Therefore, we calculate the differences between the obtained forecasts and infer the significance of means of differences. We test the hypothesis and estimate the effect size to draw conclusions.

RESULTS

Both ARIMA() and iTransformer models were used to build the forecasts. The ARIMA model was built on a deseasonalized series with the X13 procedure. A forecast of the seasonal component was generated using this procedure. Finally, both forecasts, ARIMA and X13, were combined to obtain the final forecast. The ARIMA hyperparameters were selected using an automatic search strategy for acceptable models. The Akaike Information Criterion (AIC) statistic was used to select the optimal model. Ultimately, the ARIMA(3,0,0)(0,0,1)$_{12}$ model proved the best. Admittedly, the seasonal component appeared here despite the earlier removal of

[3] https://ts.gluon.ai/stable/getting_started/models.html.

[3] https://github.com/thuml/Time-Series-Library?tab=readme-ov-file

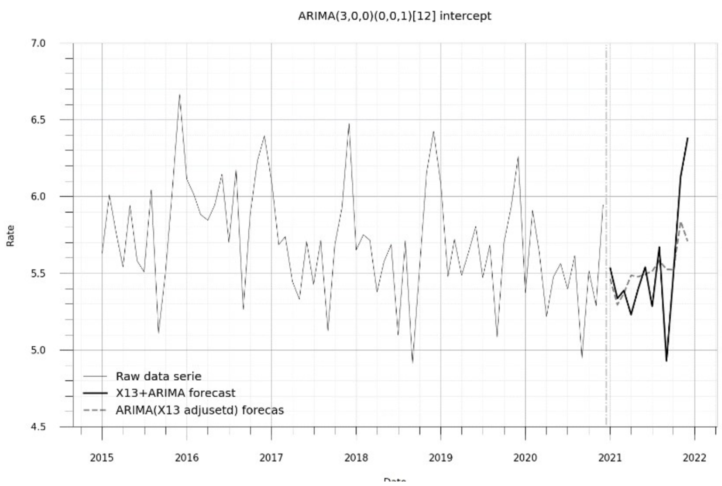
seasonality by the X13 method. Still, the authors left it out because the often-used automatic model selection procedure is not corrected in practice. The paper aimed to test the hypothesis of model equivalence, including the methodologies for using them.

In the data used, an effect is very often encountered in practice: the seasonal component is pronounced, and it determines the short-term forecasts. If this component is removed, we will have a slow-moving time series that ARIMA models can forecast.

In the case of the iTransformers model, the series was divided into a training and a test set, which were used to build the model. The testing set was 12 months, and the training set was 24 months. The training/testing window was shifted by 6 months to obtain further scenarios. The choice of size of the training and testing series was dictated by the expectation of a seasonal period. This choice made it unnecessary to pre-season the data.

Notably, the pattern obtained in this approach did not differ from that obtained in the X13 procedure - the forecasts did not differ within error limits.

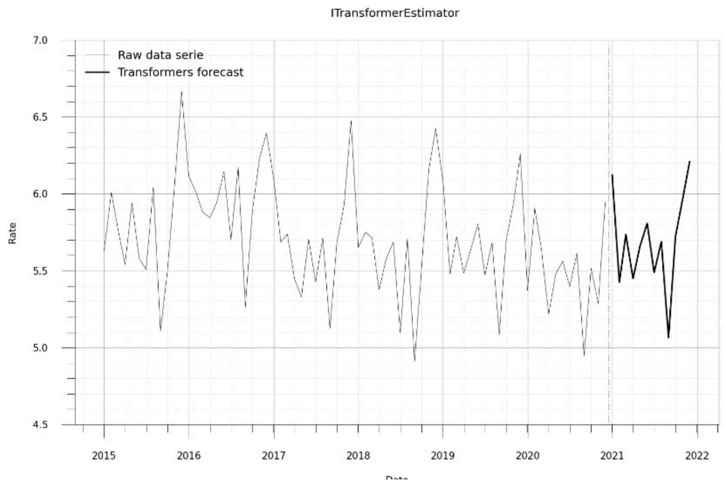
Figure 1. Preterm birth rate. iTransformers forecast for 12 months.



Source: own calculations

Fig. 1 presents the graph of the original time series (dashed line) and the forecast for the sequential 12 months (solid line) obtained for iTransforemrs. The forecast exhibits seasonal behavior with a similar pattern to the original data.

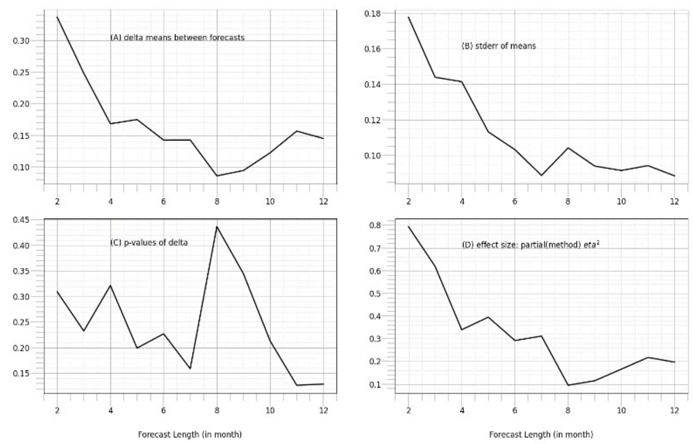
Figure 2. Seasonal ARIMA forecast for 12 months



Source: own calculations

Figure 2 shows the graph of the original time series (dashed line) and the forecast for the sequential 12 months (solid line) obtained for ARIMA X13 adjusted model. The forecast exhibits seasonal behavior with a similar pattern to the original data.

Figure 3. Comparison of forecasts for forecast horizon from 2 to 12 months. (A) differences of means calculated for different horizon values, (B) standard errors of means, (C) p values for hypothesis, (D) effect size for different horizon values.



Source: own calculations

Fig. 3 shows the results of comparing the forecasts obtained from the two models. The Generalized Linear Mixed Model (GLMM) was used for the comparison, which considers that the observations are not independent - the model uses combined observations for individual values of the forecast horizon length. In addition, each observation has a different variance, which leads to the need to introduce weights for each observation. The weights are inversely proportional to the variance.

As the forecast horizon length increases, the standardized distance between models decreases, forecast error increases, and significance tests have less and less power. Therefore, the last graph (D) shows the so-called effect size partial η^2 for the factor: the difference between forecasts.

Table 1. P-values calculated for testing mean differences

Model	Difference between the means	p-value
4 months	0.17	0.32
6 months	0.14	0.225
12 months	0.145	0.12

Source: own calculations

The values shown in Table 1 indicate that it no longer mandates the rejection of hypothesis H_0 of equality of model results.

Table 2. The effect size for different horizon values

Model	η^2	Effect
4 months	0.32	small/medium
6 months	0.3	small/medium
12 months	0.2	small

Source: own calculations

The values shown in Table 2 indicate that the measured difference between the forecast values of the two models can be considered small.

CONCLUSIONS

Our results indicate that both approaches yield similar results. The model based on iTransformer captures well seasonality without any data preparation.

Seasonal variations often affect time series of sub-annual observations, e.g., monthly, quarterly, and weekly. The presence of such variations introduces an essential component to analytical models. Indeed, seasonality usually accounts for most of the total variation within a year. Seasonality results from the fact that some months or quarters of the year are more important in activity or level. The seasonal pattern measures the relative importance of the months of the year and affects both the interpretation and projections of medium- and short-term results. The constant

100% represents the average month. The peak month is December, in which the prematurity rate is almost 15% higher than in the average month; the trough months are September and October, in which the rate is nearly 10% lower than in the average month. The seasonal amplitude, the difference between the peak and trough months of the seasonal pattern, is almost 40%. Research is ongoing to attempt to link the seasonal pattern to biological factors. Forecasting models that take seasonality into account are essential for macro-level management purposes. A methodology based on the X3-X13 procedures has been developed for several decades. It is widely used in economics, although it consists of steps based on moving averages of different lengths. In this article, we wanted to check whether deep learning methods allow for building time series forecasts in the case of seasonality. Based on the obtained results, it can be stated that iTransformer neural networks allow for obtaining results comparable to the X13 procedure. This means that iTransformer neural networks are suitable for modeling seasonal data. It is worth noting that the results are obtained from an analysis based on a self-attention mechanism that examines the global properties of the series - including seasonality, and not the search for local patterns, as in the case of autocorrelation

Comparing these approaches can help find the causes of strong seasonality in the studied problem of preterm births.

We hope our study will contribute to discussing Transformers' application in time series forecasting.

REFERENCES

Ahmed S., Nielsen I.E., Tripathi A. et al. (2023) Transformers in Time-Series Analysis: A Tutorial. Circuits Syst Signal Process 42, 7433-7466, https://doi.org/10.1007/s00034-023-02454-8.

Aijaz I., Agarwal P. (2019) A Study on Time Series Forecasting using Hybridization of Time Series Models and Neural Networks. Recent Patents on Computer Science, 12(5), 827-832, https://doi.org/10.2174/1573401315666190619112842.

Borboa-Olivares H., Rodríguez-Sibaja M.J., Espejel-Nuñez A., Flores-Pliego A., Mendoza-Ortega J., Camacho-Arroyo I., González-Camarena R., Echeverría-Arjonilla J.C., Estrada-Gutierrez G. (2023) A Novel Predictive Machine Learning Model Integrating Cytokines in Cervical-Vaginal Mucus Increases the Prediction Rate for Preterm Birth. International Journal of Molecular Sciences, 24(18), 13851, https://doi.org/10.3390/ijms241813851.

Box G., Jenkins G. (1970) Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco.

Brandon D., McGrath J. (2016) Infant Mortality Rates and Preterm Birth: A Challenge for Advocacy. Advances in Neonatal Care, 16(5), 323-324, https://doi.org/10.1097/ANC.0000000000000344.

Dench D., Joyce T., Minkoff H. (2022) United States Preterm Birth Rate and COVID-19. Pediatrics, 149(5), https://doi.org/10.1542/peds.2021-055495.

Gemmill A., Casey J. A., Catalano R., Karasek D., Margerison C. E., Bruckner T. (2022) Changes in Preterm Birth and Caesarean Deliveries in the United States during the SARS-CoV-2 Pandemic. Paediatric and Perinatal Epidemiology, 36(4), 485-489, https://doi.org/10.1111/ppe.12811.

Lakens D. (2013) Calculating and Reporting Effect Sizes to Facilitate Cumulative Science: a Practical Primer for T-Tests and ANOVAs. Frontiers in Psychology, 4, 863. https://doi.org/10.3389/fpsyg.2013.00863.

Lim B., Zohren S. (2021) Time-Series Forecasting with Deep Learning: a Survey. Phil. Trans. R. Soc. A, 379: 20200209, https://doi.org/10.1098/rsta.2020.0209

Liu Y., Hu T., Zhang H., Wu H., Wang S., Ma L., Long M. (2023) Itransformer: Inverted Transformers Are Effective for Time Series Forecasting. arXiv preprint arXiv: 2310.06625.

Miller J. A., Aldosari M., Saeed F., Barna N. H., Rana S., Arpinar I. B., Liu N. (2024) A Survey of Deep Learning and Foundation Models for Time Series Forecasting. arXiv preprint arXiv:2401.13912.

Priya T., Sarkar B. K., Sahana S. K. (2024) Regression based Machine Learning Models for Forecasting Preterm Birth Cases. Procedia Computer Science, 235, 830-839, https://doi.org/10.1016/j.procs.2024.04.079.

Sefidkar R., Zayeri F., Kazemi E., Salehi M., Dehnad A., Hafizi M. (2021) A Trend Study of Preterm Infant Mortality Rate in Developed and Developing Countries over 1990 to 2017. Iranian Journal of Public Health, *50*(2), 369-375, https://doi.org/10.18502/ijph.v50i2.5353.

Tzitiridou-Chatzopoulou M., Zournatzidou G., Kourakos M. (2024) Predicting Future Birth Rates with the Use of an Adaptive Machine Learning Algorithm: A Forecasting Experiment for Scotland. International Journal of Environmental Research and Public Health, 21(7), 84, https://doi.org/10.3390/ijerph21070841.

Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A., Kaiser L., Polosukhin I. (2017) Attention Is All You Need. NeurIPS, 2017.

Wen Q., Zhou T., Zhang C., Chen W., Ma Z., Yan J., Sun L. (2022) Transformers in Time Series: A Survey. arXiv preprint arXiv:2202.07125.

Yu Q. Y., Lin Y., Zhou Y. R., Yang X. J., Hemelaar J. (2024) Predicting Risk of Preterm Birth in Singleton Pregnancies using Machine Learning Algorithms. Frontiers in Big Data, 7, https://doi.org/10.3389/fdata.2024.1291196.

Zeng A., Chen M., Zhang, L., & Xu, Q. (2023). Are Transformers Effective for Time Series Forecasting?. Proceedings of the AAAI Conference on Artificial Intelligence, 37(9), 11121-11128, https://doi.org/10.1609/aaai.v37i9.26317.

Zhang F., Tong L., Shi C., Zuo R., Wang L. Wang Y. (2024) Deep Learning in Predicting Preterm Birth: A Comparative Study of Machine Learning Algorithms. Maternal-Fetal Medicine, 6(3), 141-146, https://doi.org/10.1097/FM9.0000000000000236.

Zhang Y., Du S., Hu T., Xu S., Lu H., Xu C., Li J., Zhu X. (2023) Establishment of a Model for Predicting Preterm Birth based on the Machine Learning Algorithm. BMC Pregnancy and Childbirth, 23(1), 779, https://doi.org/10.1186/s12884-023-06058-7.

QUANTITATIVE METHODS IN ECONOMICS
Volume XXV, No. 3, 2024, pp. 134 – 143

PROBLEMS OF EXCISE TAX ON THE EXAMPLE OF IMPORT OF PLASTIC LUBRICANTS

Jarosław Gorący https://orcid.org/0009-0009-6650-8046
Doctoral School
University of Szczecin, Poland
e-mail: jaroslaw.goracy@phd.usz.edu.pl
Iwona Markowicz https://orcid.org/0000-0003-1119-0789
Institute of Economics and Finance
University of Szczecin, Poland
e-mail: iwona.markowicz@usz.edu.pl

Abstract: The paper attempts to present the issue of excise taxation of goods - plastic greases. The specificity of the product such as plastic greases and the gap in existing tax regulations will be explained. Excise tax and the method of calculating it will be discussed. The dynamics of the volume of imports of goods to Poland will be presented and problems related to separating the volume of imports of plastic greases will be indicated. The paper will propose an attempt to identify the lubricants in question by CN codes (EU "Combined Nomenclature" - tariff and statistical nomenclature for goods). The volume of import of plastic greases will be assessed against the background of imports of other lubricants and fuel products.

Keywords: excise tax, plastic lubricants, imports, data analysis

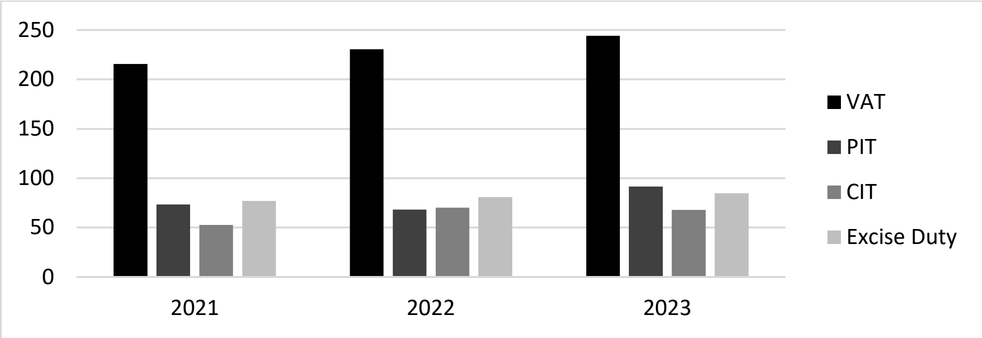
JEL classification: F14, H26, C13

INTRODUCTION

The issue of excise tax is discussed much less frequently in scientific literature than another common indirect tax, which is the value-added tax (VAT). This is due, among other things, to the size of the collection of both taxes. According to available data, VAT has been the largest revenue for the state budget for many years, and excise tax was the second largest revenue, but many times lower (Figure 1).

https://doi.org/10.22630/MIBE.2024.25.3.12

Figure 1. Tax revenue in Poland in 2021-2023 (in PLN billion)



Source: own study based on data: www.mf.gov.pl, www.stat.gov.pl

It is worth noting that in the last year, excise tax fell to third place, after personal income tax, in terms of revenue from this source. For years, Polish state authorities have been trying to curb the grey zone, both in the case of VAT and excise tax, especially in the fuel and tobacco industries. The problems with determining the volume of trade in goods from individual product groups in EU countries (for the purpose of calculating VAT) were written about by, for example, Markowicz, Baran [2021], Javorcik, Narciso [2008].

Excise duty is an indirect tax, but it is not applied universally like VAT, but on selected categories of products. It is imposed on selected goods to limit their consumption (such as alcohol or tobacco) and on infrastructure goods (such as the liquid fuel market). It is paid by the final recipient and has different rates, depending on the type of goods [Excise Duty Act]. Excise duty is divided into harmonized and non-harmonized. Harmonization consists in common rules for the production, movement and storage of goods, describes the amount and method of collecting excise duty and provides rates that cannot be lower than those given in EU directives. The group of goods subject to harmonized excise duty are motor fuels, heating oils, tobacco products, alcohol and electricity. The group of non-harmonized products (greater freedom in levying excise duty) includes passenger cars, cosmetics, weapons, leather and fur [Excise Duty Act].

RESEARCH PROBLEM

Plastic greases are "a group of lubricants used in cases where there is a need for good sealing of the friction node against the access of water and mechanical impurities, as well as a requirement for good adhesion to metal surfaces. The plastic nature of greases means that under normal conditions they maintain the shape given to them, without flowing - like oils. Greases begin to flow only after they are subjected to shear stresses exceeding the plastic limit" [Zajezierska 2016]. Plastic greases are usually a combination of three components: 1) liquid, usually base oil (70-90% content) - determines the properties (mineral oils, synthetic hydrocarbon oils, polyglycol oils, synthetic esters, silicone oils, vegetable or animal fats, fatty acid esters, ethers), 2) thickener (5-30% content) - soaps, organic and inorganic agents, naphthenic waxes, 3) refining additives (0.5-5% content) - antioxidants, anticorrosives, adhesives and metal deactivators [Skibińska 2021; Krawiec 2011]. The practical division of plastic greases based on their intended use is as follows: antifriction agents, preservatives to prevent corrosion, friction greases increasing the coefficient of friction, sealing greases, greases for special applications. According to the EU regulation, lubricants are "hydrocarbons produced from distillation by-products, mainly used to reduce friction between load-bearing surfaces. This item includes all types of finished lubricating oils, from spindle oil to cylinder oil, and used in greases, engine oils, as well as all types of base oils" [Commission Regulation (EU) 2017/2010]. In European countries, the ISO 6743-9:2003 standard is most commonly used for the classification of plastic lubricants, and in Poland its equivalent is PN-ISO 6743-9:2009. The quality requirements for plastic lubricants are also included in European national standards: German - DIN 51825:2004 and Swedish - SS155470:2003. The international standard combining the classification system of plastic lubricants and specifying their quality requirements is ISO 12924:2010 [Sik 2020]. It should be mentioned here that DIN 51825:2004 is a standard that divides plastic greases according to their consistency and resistance to deformation. It is most often used by manufacturers of goods purchased within the Community to Poland, mainly from Germany. On the other hand, the basis for the classification of greases according to the PN-ISO 6743-9.2009 standard is, among others, the minimum and maximum working temperature, the ability to properly lubricate in the presence of water, the ability to lubricate under pressure [Ptak 2012; Błaszkiewicz, Moskała 2017].

In the excise law, applied to taxation at the effective excise duty rate, no statutory definition of the concept of "plastic greases" was used. The legislator did not provide in the Excise Duty Act any procedures to be applied to the movement of excise goods not listed in Annex No. 2 to the Excise Duty Act subject to a zero excise duty rate (applies to plastic greases). It was necessary to submit an excise duty registration application - AKC-R - by 30.06.2021 (in accordance with Art. 16 sec. 1 of the Act, an entity conducting business activity was obliged, before the date of the first activity subject to excise duty or the first activity using excise goods exempt from excise duty or taxed at a zero rate, to submit a registration application to the relevant head of the tax office). Additionally, taxpayers must be registered in the Central Register of Excise Entities (CRPA). This registration replaced the submission of AKC-R registration forms (from 01.07.2021) [Excise Duty Act]. RPA is a nationwide database of entities that have submitted registration applications for excise duty purposes. It allows you to check whether contractors who participate in the trade of excise goods are registered and thus reduces the risk of cooperation with a dishonest business partner. In addition, from the third quarter of 2021, taxpayers acquiring intra-Community excise goods with a zero rate must submit the AKC-UAKZ declaration (from the first quarter of 2023, the form will change to AKC-

KZ), which is a new excise declaration template that has not yet functioned on the basis of excise duty. The AKC-KZ declaration shows the quantities of excise goods subject to a zero rate. However, suppose the conditions for applying the zero rate are not met, and it is necessary to use the positive rate and pay the tax. In that case, the taxpayer submits a simplified AKC-UA declaration. At this point, the entity will also pay the excise duty due. Unfortunately, the regulation regarding the submission of AKC-KZ declarations for each quarter for products with a zero rate applies only to products included in Annex No. 2. Plastic greases included in CN codes 27101999 and 3403 are not listed in this Annex, and taxpayers registered in the CRPA as purchasing plastic greases from the EU are not obliged to submit an excise duty declaration. In addition, the taxpayer is obliged to keep records of excise goods acquired within the Community. In February 2023, the trade in excise goods with paid excise duty listed in Annex No. 2 was regulated by introducing the obligation to register the e-SAD document. On its basis, goods are moved, within the framework of intra-Community acquisition or delivery, outside the excise duty suspension procedure. In principle, e-SAD concerns products harmonized within the EU with excise duty paid or subject to a zero rate, as well as ethyl alcohol completely denatured with permitted agents (based on Commission Regulation (EC) No. 3199/93 of 22 November 1993), but it concerns the circulation of excise goods listed in Annex No. 2. Again, this means that it does not cover the obligation to register intra-Community movements for plastic greases, which are excluded from effective excise duty taxation.

Thanks to this regulation, in the SENT system, in addition to the EMCS PL 2 and ALINA 4 systems, an IT system was created in which the transport of lubricating oils, lubricating preparations and plastic greases is recorded. However, the limitation in the use of the system of only four-digit CN codes makes it impossible to conduct a practical analysis of the transport of lubricating products (including plastic greases) because it is impossible to determine the exact commodity classification of the transported grease. The data collected in the SENT system is burdened with statistical errors due to the arbitrariness of filling in the fields (lack of system validation). The above systems often use a non-uniform unit of measurement of the transported product (litres, kg) and the use of trade names of products in the descriptions and name of the product (which makes it difficult to check the identity of the goods). In addition, the SENT system does not cover the transport of products classified under CN codes 2710 and 3403 provided that they are transported in small unit packages, which is the norm in the case of plastic greases. The Combined Nomenclature (CN) is a commodity coding system, generally consisting of eight digits, covering the Harmonized System (HS) codes with further subdivisions. It serves the EU's common customs tariff and also provides statistics on trade within the EU and between the EU and the rest of the world. The CN nomenclature groups goods into sections, chapters and subdivisions [Regulation 2018]. Therefore, there is no single definition of plastic lubricants in Polish law or one selected CN code to describe them. They can be included in two CN codes - 27101999 or 3403. Currently, plastic lubricants with CN code 27101999 are subject to a zero excise duty rate pursuant to Article 89 paragraph 1 item 11 of the Excise Duty Act. The Act was amended in 2019 and introduced similar provisions for plastic greases from item 3403. At the same time, lubricating oils and lubricating preparations taxed with effective excise duty (PLN 1,180/1,000 l) may be subject to the excise duty suspension procedure in domestic trade and when exported from a tax warehouse. Given that plastic greases are excluded from effective excise duty taxation, the excise duty suspension procedure does not apply to them. The difference is also that lubricating oils and lubricating preparations may be exempt from excise duty due to their intended use. Regarding the movement of products using the excise duty suspension procedure, taxpayers rightly apply to the Director of the National Tax Information for issuing Binding Excise Information (WIA) due to the lack of precision in the regulations. In accordance with Art. 7 of the Excise Duty Act WIA is a decision issued for the purposes of taxing an excise product with excise duty, organizing the trade in excise products or marking these products with excise stamps. WIA determines the classification of an excise product in a system corresponding to the Combined Nomenclature (CN) or the type of excise product by describing this product as sufficient to determine the taxation of excise duty, organizing the trade in excise products or marking these products with excise stamps.

Since plastic greases have the same CN code as oils and lubricating preparations and due to the lack of clearly defined definitions of plastic grease in Polish regulations, this gives rise to many problems. The basic form of challenging an incorrectly applied excise duty rate in the trade of plastic greases is to verify the correct classification of the grease according to industry standards and the physicochemical composition of the product, i.e. laboratory testing. Only effective negation as a result of sampling and recognition of the incorrect CN classification of the product or verification of the lack of properties of the plastic grease and determination of the properties of the product as a product in another CN code resulting in the creation of a tax liability will result in effective determination of the moment of emergence of the mandatory excise duty. In particular, CN item 3403 is used declaratively by economic entities. The lack of a processed Central Register of Excise Products (CEWA) makes it impossible to distinguish, on the basis of the records of excise goods in entities without physical control in the entity, which products are plastic greases and which are lubricating preparations in entities with excise permits. The names "lubricant preparation" or "plastic grease" used in the trade names of products often cause interpretational discrepancies in the application of excise tax law provisions, especially in the scope of resale of plastic greases after purchase on the domestic market, import or acquisition or intra-Community delivery. Secondly, it is possible to use plastic greases as an admixture or use these products for the production of liquid or heating fuels, which, taking into account the physicochemical properties of plastic greases, should not take place due to, for example, their density and the use of improvers in their production. Mixing, reclassifying, adding other excise goods to the production in the form of oils classified in the CN 2710 group (mainly liquid and heating fuels, or various types of lubricants subject to excise duty) taxed at an effective rate without declaring the use or consumption for the production of these products, means that the products "produced" in this way do not include the full amount of excise duty due on the products in their sales price and are more competitive in terms of price. Such irregularities are also indicated by industry organizations such as the Polish Organization of Oil Industry and Trade, which postulates the need to establish a definition of plastic greases in the Excise Duty Act or to regulate the issue of entry into the BDO database.

RESEARCH RESULTS

Taking into account the described difficulties in demonstrating which specific product may be a plastic grease, an attempt was made to show the volume of products imported to Poland in the last three years with CN codes: 3403 and 2710 19 99, which contain plastic greases (Tables 1-4).

Table 1. Imports of CN 3403 goods from non-EU countries in 2021-2023 (thousand kg)

Countries	2021	2022	2023
Total non-EU	**3 399**	**3 504**	**3 012**
In it:			
United Kingdom	1 604	1 741	1 346
Switzerland	702	657	659
United States	367	381	329
Japan	421	413	311
China	24	75	175
Countries and territories not specified within the framework of trade with third countries	94	99	68
Turkey	80	78	43
South Korea	32	19	31
India	13	7	18
Ukraine	12	15	16
Canada	10	11	7
Russia	18	1	0

Source: own study based on data: https://trade.ec.europa.eu/access-to-markets/pl/statistics [access: 10 May 2024]

Table 2. Intra-Community acquisition of goods CN 3403 in 2021-2023 (thousand kg)

Countries	2021	2022	2023
EU27	**50 517**	**47 965**	**33 493**
In it:			
Germany	19 866	18 368	16 102
Netherlands	6 604	6 785	5 844
France	8 997	10 071	2 713

Table 2. continued

Countries	2021	2022	2023
Italy	3 067	2 985	2 708
Belgium	3 414	3 303	2 665
Spain	1 155	1 352	1 293
Czech Republic	873	1 061	893
Slovakia	361	709	359
Sweden	369	248	234
Austria	452	137	222
Ireland	473	412	216
Slovenia	25	43	52
Hungary	38	25	48
Romania	119	90	33
Lithuania	4 556	2 243	30
Denmark	20	24	30
Greece	23	34	20
Finland	41	56	20
Portugal	50	13	2
Estonia	10	5	2

Source: own study based on data: Eurostat database

Table 3. Imports of CN 2710 19 99 goods from non-EU countries in 2021-2023 (thousand kg)

Countries	2021	2022	2023
Total non-EU	**4 039**	**3 965.1**	**9 505.6**
In it:			
China	1	1 292.5	6 172.9
Japan	283	272.2	1 381.5
South Korea	189	417.5	714.6
United Kingdom	407	567.8	501.1
United States	269	295.3	385.3
Switzerland	46	56.6	117.4
Canada	3	27.7	70.3
Countries and territories not specified within the framework of trade with third countries	83	79.0	55.2
Turkey	97	163.1	45.2
Serbia	29	40.0	14.2
Singapore	20.8	0.5	1.0
Norway	1.2	3.4	0.9
Australia	1.6	1.9	0.8
Russia	2183.0	743.2	0
Belarus	412.2	0	0
Malaysia	13.6	0	0

Source: own study based on data: https://trade.ec.europa.eu/access-to-markets/pl/statistics [access: 10 May 2024]

Table 4. Intra-Community acquisition of goods CN 2710 19 99 in 2021-2023 (thousand kg)

Countries	2021	2022	2023
EU27	**116 947.0**	**113 194.5**	**109 639.2**
In it:			
Germany	44 388.6	40 914.2	49 537.0
Belgium	23 135.9	26 644.2	27 119.5
Sweden	9 712.7	8 521.6	7 483.5
France	11 855.5	10 601.1	5 336.1
Hungary	2 987.2	3 515.2	4 502.8
Netherlands	4 141.1	4 732.2	3 550.0
Italy	2 183.1	2 086.7	3 144.1
Spain	615.2	1 761.5	2 211.8
Czech Republic	9 288.4	3 587.5	1 541.9
Denmark	2 346.8	201.2	1 478.1
Lithuania	2 108.4	4 779.3	1 437.5
Finland	647.6	1 651.9	1 206.2
Latvia	2 015.3	3 666.4	552.5
Greece	10.0	16.9	318.1
Austria	58.2	115.3	75.9
Slovakia	914.8	152.5	50.4
Romania	168.7	16.5	36.2
Cyprus	0	178.4	22.4
Luxembourg	12.6	25.5	14.4
Portugal	223.6	13.0	2.3
Bulgaria	127.3	0.1	0

Source: own study based on data: Eurostat database

Data on the CN code 3403 clearly indicate a decreasing trend in the import of this type of goods to Poland in the last three years, both from outside the EU and from EU countries (tables 1 and 2). As for products with the CN code 2710 19 99, we notice a decreasing trend in the scope of intra-Community acquisition of goods and an increasing trend in the case of import from outside the EU. Important in the context of the outbreak of the war in Ukraine is the reduction and then cessation of import of goods from Russia and Belarus, with a simultaneous huge increase in import from China. It can be assumed that some of the goods previously imported from Russia were sold by China due to sanctions, but this is only an attempt to explain this state of affairs.

The Polish Organization of Oil Industry and Trade [2023, 2024] periodically publishes reports on the structure of the industrial oils market in Poland. In recent years (2021-2022), the volume of trade in plastic lubricants in Poland is estimated at approximately 6.5% of this market. Considering that the volume of goods imported in 2023 with CN codes 3403 and 2710 1999 amounted to a total of about 156 million kg, and considering that plastic greases constitute about 6.6% of the industrial oil market, the volume of import of plastic greases to Poland can be estimated. It is about

10.3 million kg. Of course, we do not include production in the territory of the country, which is difficult to determine for the purposes of scientific research. The reasons include: the lack of precise specification in the regulations of what constitutes a plastic grease, the specification of production data of individual enterprises and failure to inform about the production volume of individual goods due to trade secrets.

SUMMARY

This article aimed to introduce the subject of plastic greases in the context of importing goods to Poland and excise duty. The most important conclusion is that in the current legal situation it is almost impossible to determine whether a product imported to the country is actually a plastic grease, because we rely on manufacturers' declarations. Plans to introduce the Central Register of Excise Products may change this state of affairs, but at the moment state authorities, including the National Revenue Administration, have no possibility of distinguishing based on the records of excise products in entities with excise permits, without physical inspection in the entity, which products are plastic greases and which are lubricants. Indicating the differences is possible only after a possible analysis of the product data sheets of the lubricant product. Therefore, an attempt was made to bring the issue closer to determining the volume of goods imported with CN codes 3403 and 2710 19 99 to Poland over 2021-2023. This shows that this volume is significant, which may translate into increased tax revenues in the country. This gives rise to further research in this area. In addition, the excise tax itself is extremely important for a balanced Polish budget and there will undoubtedly be an increasing emphasis on its collection.

REFERENCES

Błaszkiewicz Z., Moskała J. (2017) Normalizacja środków smarowych w Polsce. Nafta-Gaz, 12, 980-987.

Eurostat, baza danych, https://ec.europa.eu/eurostat/web/main/data/database.

Javorcik, B. S., Narciso, G. (2008) Differentiated Products and Evasion of Import Tariffs. Journal of International Economics, 76(2), 208–222.

Krawiec S. (2011) Kompozycje smarów plastycznych i stałych w procesie tarcia stalowych węzłów maszyn. Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław.

Markowicz, I., Baran, P. (2021) Mirror Data Asymmetry in International Trade by Commodity Group: the Case of Intra-Community Trade. Oeconomia Copernicana, 12(4), 889–905.

Podatek akcyzowy – wewnątrzwspólnotowe nabycie smarów i olejów smarowych. Dyrektor Krajowej Informacji Skarbowej, interpretacja indywidualna 30.11.2018 r., 0111-KDIB3-3.4013.199.2018.3.WR, Wolters Kluwer, https://sip.lex.pl [dostęp: 10.04.2024].

Polska Organizacja Przemysłu i Handlu Naftowego (2023) Raport Roczny POPHIN, Przemysł i handel naftowy2022, Warszawa, https://popihn.pl/raporty-i-konferencje/.

Polska Organizacja Przemysłu i Handlu Naftowego (2024) Raport Roczny POPHIN, Przemysł i handel naftowy2023, Warszawa, https://popihn.pl/raporty-i-konferencje/.

Ptak S. (2012) Klasyfikacja jakościowa i charakterystyka przemysłowych środków smarowych. Nafta-Gaz, 7, 454-461.

Rozporządzenie wykonawcze Komisji (UE) 2018/1602 z dnia 11 października 2018 r. zmieniające załącznik I do rozporządzenia Rady (EWG) nr 2658/87 w sprawie nomenklatury taryfowej i statystycznej oraz w sprawie Wspólnej Taryfy Celnej.

Skibińska A. (2021) Stabilność termooksydacyjna smarów plastycznych. Nafta-Gaz, 7, 471-479.

Sik M. (2020) Stawka podatku akcyzowego dla nabywanych wewnątrzwspólnotowo wyrobów energetycznych. Akcyza i Podatek Akcyzowy – Serwis ISP, 23/2020.

Rozporządzenie Komisji (UE) 2017/2010 z dnia 9 listopada 2017 r. zmieniające rozporządzenie Parlamentu Europejskiego i Rady (WE) nr 1099/2008 w sprawie statystyki energii w odniesieniu do aktualizacji rocznych i miesięcznych statystyk dotyczących energii, Załącznik A pkt 3.4.20.

Zajezierska A. (2016) Biodegradowalne smary plastyczne. Prace Naukowe INiG-PIB, 197.

Ustawa o podatku akcyzowym [Excise duty act], Dz. U. 2009 Nr 3 poz. 11, 6 grudnia 2008 r.

https://trade.ec.europa.eu/access-to-markets/pl/statistics [access: 10.05.2024]

www.mf.gov.pl [access: 16.05.2024].

www.stat.gov.pl [access: 16.05.2024].