NEW YORK ECONOMIC REVIEW

JOURNAL OF THE NEW YORK STATE
ECONOMICS ASSOCIATION

VOLUME LIV

Vol. 54, Fall 2024

CONTENTS

EDITORIAL

The *New York Economic Review* (NYER) is an annual publication of the New York State Economics Association (NYSEA). The NYER publishes theoretical and empirical articles, and also interpretive reviews of the literature in the fields of economics and finance. All well-written, original manuscripts are welcome for consideration at the NYER. We also encourage the submission of short articles and replication studies. Special Issue proposals are welcome and require a minimum of 4 papers to be included in the proposal, as well as a list of suggested referees.

MANUSCRIPT GUIDELINES

All manuscripts are to be submitted via https://www.nyseconomicsassociation.org/new-york-economic-review/submit-paper-to-the-editor/ to the editor in a Microsoft Word file format, with content formatted according to the guidelines at:

https://www.nyseconomicsassociation.org/NYER/Guidelines

The NYER is cited in:

- EconLit by the American Economic Association
- Ulrich's International Periodicals Directory
- Cabell's Directory of Publishing Opportunities in Business and Economics
- Australian Business Deans Council (ABDC) Journal Quality List

ISSN NO: 1090-5693

The NYS Establishment Survey as a Forecasting Tool and a Barometer of NYS Economy

Kajal Lahiri[1], Terrence Kinal[1], Zulkarnain Pulungan[1]

ABSTRACT

In this paper we test and measure the predictive performance of the NYS Establishment Surveys. Using Pesaran and Timmermann test, the null hypothesis of no predictable relationship of each establishment performance measure is rejected. This is also supported graphically by the results of the diffusion index. Goodman-Kruskal gamma coefficient is used to measure how good the forecasts predict their realizations. The coefficients range from 0.412 to 0.729.

Key Words: Directional survey; Diffusion Index; Predictive performance; Gamma coefficient.
JEL codes: C18; C25; C53; C83; D22.

1. INTRODUCTION

With rare and generous support from the Division of the Budget, the New York State Establishment Surveys were conducted by the Econometric Research Institute-SUNY Albany during 2001-2004. The type of these surveys is called a business tendency survey. The business tendency survey asks company managers/CEO about the current situation of their business and about their plans or expectations for the near future[1]. These types of surveys provide information that is valuable to the respondents themselves and to economic policymakers and analysts. The data collected in this type of survey are qualitative data and therefore they could not provide precise information on the level of economic variables. But they can be used to measure the direction of the change in economic variables and also to predict the direction changes of economic variables in the near future.

The data from a business tendency survey are usually analyzed by employing two different statistical techniques. The first technique is a transformation of qualitative data into quantitative data. The most common technique for transforming qualitative data into quantitative data is calculating the so-called "diffusion indices" or "balances". This index can be used as a measure of growth. The index is defined as the percentage of establishments indicating an increase minus the percentage of establishments indicating a decrease. Note that in the calculation of the diffusion index, the establishments indicating "remain the same" are discarded[2].

Furthermore, if we have enough observations of diffusion indices, we can explore the data using time series analysis (see for example Gourieroux and Pradel, 1986; Dasgupta and Lahiri, 1992; Pesaran and

[1]Department of Economics, University at Albany – SUNY, Albany, NY 12222. Email: klahiri@albany.edu
We are thankful to two anonymous referees and the editor for providing many comments to improve the original manuscript.

Timmermann, 1992; Granger and Pesaran, 2000). In this paper, because we do not have enough observations, the diffusion indices are only analyzed graphically.

The second technique is an analysis of a contingency table constructed from qualitative data (see for example Nerlove, 1983 for early pioneering work in econometrics). From the contingency table, we can test statistically the predictive performance of forecast variables and then measure how well the establishment's forecasts could predict their subsequent realizations. The most common way to test the predictive performance is to test the independence hypothesis of the forecasts and their realizations. The most common tests to test the independence hypothesis are χ^2 test and a test derived from a log-linear model. The discussion of χ^2 test in a more detail is provided, for example, by Agresti (1984). Nerlove (1982) used log-linear model to test the independency between forecasts and its realizations. More detail of the log-linear model can be found, for example, in Maddala (1983).

Pesaran and Timmermann (PT, 1992) propose a distribution-free procedure for testing the predictive failure when the focus of the analysis is on the correct prediction of the direction of the change in the variable under consideration. This test is appropriate when the forecasts are made available only in the form of qualitative data such as data collected in most of the business tendency surveys. The null hypothesis is there is no predictable relationship between forecasts and their realizations. This test has been used successfully by Lahiri et al. (2015).

If the hypothesis of no predictable relationships is rejected, then we need to measure how well a firm's forecasts can predict the subsequent realizations. To measure this performance, we can use a measure of association. There are several measures of association available that can be used (see Agresti, 1984). The most common measure is gamma coefficient proposed by Goodman and Kruskal (1954). The advantage of this measure is that it is easy to interpret.

The main purpose of this paper is to test the predictive performance of New York State Establishment Survey and to measure how well establishment's forecasts predict their subsequent realizations. The measures of establishment performance to be analyzed include employment, wages and salaries, profits, sales, material and energy costs, product prices, and bonuses. Unlike most establishment surveys, our survey was conducted as a panel survey. As a result, we have the unique opportunity of testing the goodness of anticipations against realizations at the level of individual establishment units.

2. STATISTICAL METHODS

2.1. Test of Predictive Performance

In this part, we discuss the PT test. Let $x_t = \hat{E}(y_t \mid \Omega_{t-1})$ be the predictor of y_t formed with respect to the information available at $t-1$, Ω_{t-1}, and suppose that there are *n* observations on (x_t, y_t). The predictive failure test is based on the proportion of times that the direction of change in y_t is correctly predicted in the sample. Suppose that the information on x_t and y_t are categorized into X_1, X_2, ..., X_m and

Y_1, Y_2, ..., Y_m, respectively. Let n_{ij} denotes the number of observations in the $X_i \times Y_j$ category cross-tabulation in the form of an $m \times m$ contingency table. Let n_{i0} and n_{0j} represent the I^{th} row and the j^{th} column of totals. The implicit null hypothesis is given by

$$H_0 : \sum_{i=1}^{m} \left(P_{ii} - P_{i0} P_{0i} \right) = 0$$

and the statistical test is

$$s_n = \sqrt{n} V_n^{-1/2} S_n \xrightarrow{a} N(0,1)$$

where

$$S_n = \sum_{i=1}^{m} \left(\hat{P}_{ii} - \hat{P}_{0i} \hat{P}_{0i} \right)$$

$$V_n = \left(\frac{\partial f(\mathbf{P})}{\partial \mathbf{P}} \right)'_{\mathbf{P}=\hat{\mathbf{P}}} \left(\hat{\mathbf{\Psi}} - \hat{\mathbf{P}}\hat{\mathbf{P}}' \right) \left(\frac{\partial f(\mathbf{P})}{\partial \mathbf{P}} \right)'_{\mathbf{P}=\hat{\mathbf{P}}}$$

$$\mathbf{P}' = \left(P_{11}, P_{12}, \ldots, P_{1m}; P_{21}, P_{22} \ \ldots; P_{m1} \ P_{m2} \ \ldots P_{mm} \right)$$

$\mathbf{\Psi} = diagonal(\mathbf{P})$ and

$$f(\mathbf{P}) = \sum_{i=1}^{m} \left(P_{ii} - P_{i0} P_{0i} \right)$$

Granger and Pesaran (2000) prove that this test is equivalent to Kuipers score that is used as a measure of forecast accuracy in the meteorology literature. This test is appropriate when the focus of analysis on predicting the overall changes in a variable rather than on the general hypothesis of statistical independence tested using the familiar χ^2 goodness-of-fit test.

2.2. Goodman-Kruskal Gamma

The Goodman-Kruskal gamma coefficient is used to measure an association between two ordered variables. The estimator of gamma is based on the number of concordant (C) and discordant pairs (D) of observations defined as

$$\hat{\gamma} = \frac{C - D}{C + D}$$

and its variance (can be found in Agresti [1984, chap. 10].) is

$$\text{var}(\hat{\gamma}) \stackrel{a}{=} \frac{16}{(C+D)^4} \sum_{i}^{m} \sum_{j=1}^{m} n_{ij} \left(DA_{ij} - CB_{ij} \right)^2$$

where $A_{ij} = \sum_{k<i}\sum_{s<j} n_{ks} + \sum_{k>i}\sum_{s>j} n_{ks}$, and $B_{ij} = \sum_{k<i}\sum_{s>j} n_{ks} + \sum_{k>i}\sum_{s<j} n_{ks}$.

Gamma has the range $-1 \le \gamma \le 1$. If the two variables are independent then the estimator of gamma tends to be close to zero. Under statistical independence, gamma will be 0, but if gamma = 0 then it is not necessarily that the two variables are independent (for example, if the number of concordant pairs is equal to the number of discordant pairs). It is a measure of rank correlation, i.e., the similarity of the orderings of the data when ranked by each of the quantities. It measures the strength of association of the cross tabulated data when both variables are measured at the ordinal level. It makes no adjustment for either table size or ties.

3. DATA

The data used in this paper are obtained from the New York State Establishment surveys collected by Econometric Research Institute- SUNY Albany. The survey was conducted biannually from April 2001 to September 2004 giving us seven rounds of very granular firm-level information. The questionnaire for the surveys is included as an appendix to the paper. The summary of the survey samples is presented in Table 1.

Table 1. Summary of Samples

Survey	Questioners	Responses	Repeated from prev. survey*
April 2001	20,000	1,905	
December 2001	41,800	3,540	456
October 2002	43,210	3,743	1,074
April 2003	43,855	4,535	1,282
October 2003	44,522	5,679	1,771
April 2004	39,880	4,729	1,803
September 2004	39,884	4,029	1,689

Note: *The repeated responses are establishments that responded to two consecutive surveys.

In these surveys, establishments were asked to give the direction of the actual changes and the expected direction of the changes of their establishment's performances (i.e. employment, wages and salaries, profits, sales, material and energy costs, product prices, and bonuses). The survey results are in the form of ordered responses. In April and December surveys of year *t*, establishments were asked to give the direction of actual changes for year *t*-2 to *t*-1 and the expected direction of the changes from *t*-1 to *t* of their establishment performances; in October and September survey for year *t*, establishments were asked to give the direction of expected change from *t*-1 to *t* and from *t* to *t*+1.

The diffusion indices are computed from all responses (including repeated and new respondents) for each survey. The summary results of the diffusion indices for all surveys were recorded. In this paper we only present the comparison of the forecasts and the realizations of their diffusion indices for the seven measures of establishment performance. There are four pairs of surveys that can be matched between the forecasts and their realizations. Those are the surveys in April and December 2001 asked the expected and actual changes[3] from 2000 to 2001, respectively; the survey in October 2002 and April 2003 asked the expected and actual changes from 2001 to 2002, respectively; the survey in April 2003 and April 2004 asked the expected and actual changes from 2002 to 2003, respectively; and the survey in October 2003 and April 2004 asked the expected and actual changes from 2002 to 2004, respectively.

For the hypothesis testing and to measure the performance of the expected changes in predicting the actual changes, we use the repeated sample from the four pairs of surveys above.

4. EMPIRICAL RESULTS[4]

Table 2 displays the joint empirical distribution for the forecasts and their realizations for the seven measures of establishment performance.

Table 2. Contingency table for expected versus actual changes in the measures of establishment performances

	Employment			Wages & salaries			Bonuses			Profits			Sales Revenue			Mat. & energy costs			Prices		
$\Delta^e \backslash \Delta$	-	=	+	-	=	+	-	=	+	-	=	+	-	=	+	-	=	+	-	=	+
Survey of April 2001 and December 2001: changes in “variable” from 2000 to 2001																					
-	6.49	4.47	0.67	1.97	3.72	1.75	9.95	5.21	1.42	16.06	4.15	2.59	11.74	3.52	2.58	0.44	0.44	0.22	2.44	2.66	0.89
=	9.40	37.36	11.63	5.47	17.72	10.28	8.53	28.91	11.85	11.40	13.73	9.33	11.50	11.74	8.69	0.44	7.47	3.74	2.88	22.39	11.31
+	3.36	10.74	15.88	4.38	14.66	40.04	3.32	15.64	15.17	8.29	10.62	23.83	8.45	11.74	30.05	5.27	21.32	60.66	3.33	18.40	35.70
Survey of October 2002 and April 2003: changes in “variable” from 2001 to 2002																					
-	10.16	5.71	1.41	6.71	4.21	1.33	14.56	7.09	2.68	26.68	6.47	4.37	21.14	8.01	3.47	0.86	1.10	1.73	3.11	2.95	0.80
=	7.82	41.59	10.24	6.01	27.77	12.25	7.28	36.40	12.64	8.56	15.03	8.65	8.26	15.94	8.75	1.49	16.24	16.24	3.27	39.03	11.09
+	1.49	7.43	14.15	2.26	9.98	29.49	1.72	6.51	11.11	4.64	5.65	19.95	3.39	6.77	24.28	1.02	12.24	49.10	0.80	14.29	24.66
Survey of April 2003 and April 2004: changes in “variable” from 2002 to 2003																					
-	8.36	6.36	2.59	5.18	4.47	2.59	13.22	7.44	3.03	19.58	8.60	3.97	17.14	7.89	3.45	0.23	0.23	0.82	1.79	3.22	1.19
=	7.89	43.23	12.60	5.29	27.65	14.59	8.54	32.78	11.02	11.51	14.42	11.38	10.48	13.19	12.08	0.94	10.43	12.08	2.26	31.47	13.23
+	1.06	6.12	11.78	1.18	9.06	30.00	4.13	6.61	13.22	7.01	7.28	16.27	6.17	9.49	20.10	0.59	12.66	62.02	1.55	16.92	28.37
Survey of October 2003 and April 2004: changes in “variable” from 2002 to 2003																					
-	9.82	6.54	1.72	6.77	4.60	2.16	14.88	4.82	2.89	26.80	7.82	4.16	23.29	7.29	3.12	0.56	0.84	1.23	2.32	3.39	0.90
=	7.76	42.71	10.59	5.60	29.01	12.87	7.30	37.05	12.53	8.76	13.81	9.90	7.06	16.12	8.94	0.95	10.98	11.37	3.11	35.40	12.54
+	1.00	5.93	13.92	1.16	8.32	29.51	1.38	7.02	12.12	3.53	6.18	19.04	3.71	6.47	24.00	1.29	12.72	60.06	0.62	13.78	27.95

Note: “variable” are the seven measures of establishment performance: employment, wages and salaries, profits, sales, material and energy costs, product prices, and bonuses.

There are three categories for expected and actual changes: “decrease” (-), “remain the same” (=), and “increase” (+). Table 2 shows that high percentage of data fall along the diagonal of each sub table. This indicates there are associations between the forecasts and their realizations.

From Table 2 we can also test the rational expectation hypothesis (REH) using the direct test[5] proposed by Gourieroux and Pradel (1986). The table indicates there is no reason to reject the REH for forecasts of employment, bonuses, profits, and sales revenue made in all surveys. The REH for expectation of wages and salaries made in April 2001 is rejected; the REH for forecasts of prices made in April 2001, April 2003, and October 2003 are rejected; and the REH for material and energy costs made in April and October 2003 are rejected. Note that rational expectation does not imply a high predictive power and vice versa as Pesaran and Timmermann point out.

Figure 1: Forecasts and Actual of Diffusion Indices for 2000-2001

(Survey: April 2001 and December 2001)

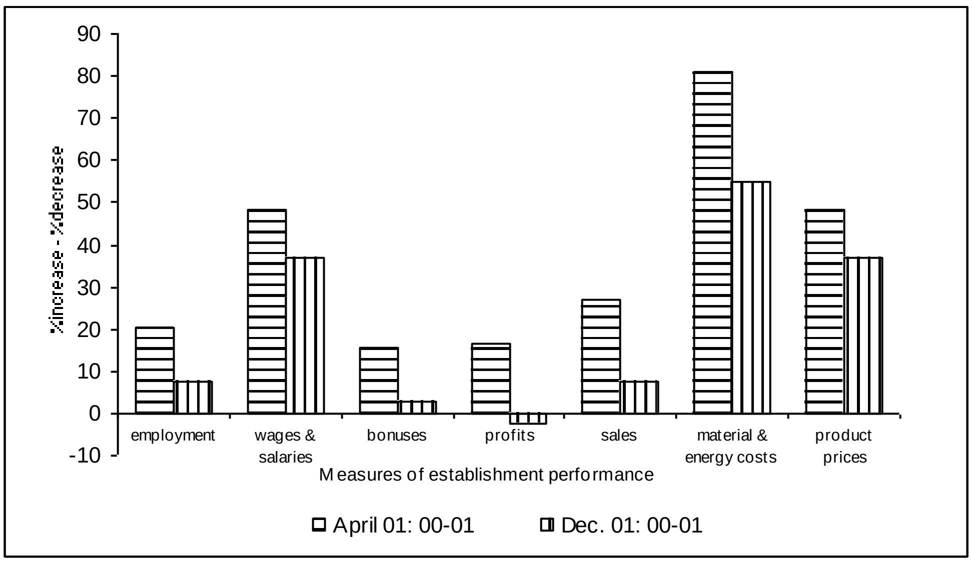


Table 3. The test statistics and gamma coefficients

Measures of Establishment Performance	**N**	**PT (s_n^2)**	**p-value**	**Chi-square (χ^2)**	**p-value**	**Goodman-Kruskal (γ)**	**Asy. std. error**
Survey of April 2001 and December 2001: changes in "variable" from 2000 to 2001							
Employment	447	62.01	0.00000	102.33	0.00000	0.580	0.059
Wages and salaries	457	58.48	0.00000	68.14	0.00000	0.560	0.055
Bonuses	211	20.84	0.00000	45.50	0.00000	0.541	0.082
Profits	386	68.08	0.00000	88.25	0.00000	0.571	0.053
Sales revenue	426	65.18	0.00000	88.49	0.00000	0.576	0.051
Material and energy costs	455	26.29	0.00000	46.94	0.00000	0.579	0.079
Prices	451	55.16	0.00000	81.57	0.00000	0.548	0.061
Survey of October 2002 and April 2003: changes in "variable" from 2001 to 2002							
Employment	1279	291.47	0.00000	488.38	0.00000	0.701	0.030
Wages and salaries	1282	336.03	0.00000	460.81	0.00000	0.702	0.027
Bonuses	522	111.72	0.00000	173.16	0.00000	0.638	0.050
Profits	1098	375.68	0.00000	410.14	0.00000	0.675	0.027
Sales revenue	1211	406.11	0.00000	475.09	0.00000	0.709	0.024
Material and energy costs	1275	124.68	0.00000	182.65	0.00000	0.567	0.039
Prices	1253	256.26	0.00000	415.64	0.00000	0.704	0.030

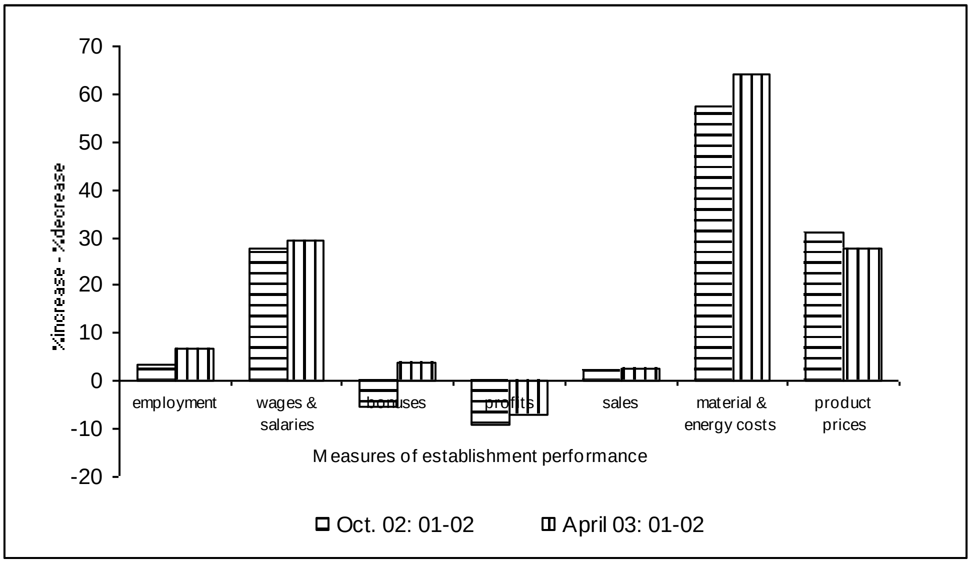
Survey of April 2003 and April 2004: changes in “variable” from 2002 to 2003							
			0.0000		0.0000		
Employment	849	138.30	0	233.45	0	0.610	0.044
			0.0000		0.0000		
Wages and salaries	850	193.95	0	256.73	0	0.674	0.035
			0.0000		0.0000		
Bonuses	363	71.84	0	94.50	0	0.525	0.067
			0.0000		0.0000		
Profits	756	88.98	0	124.56	0	0.492	0.040
			0.0000		0.0000		
Sales revenue	811	95.97	0	149.33	0	0.534	0.037
Material and energy			0.0000		0.0000		
costs	853	56.64	0	95.91	0	0.605	0.052
			0.0000		0.0000		
Prices	839	105.96	0	146.03	0	0.556	0.046
Survey of October 2003 and April 2004: changes in “variable” from 2002 to 2003							
	180		0.0000		0.0000		
Employment	3	413.64	0	699.30	0	0.714	0.025
	180		0.0000		0.0000		
Wages and salaries	3	538.32	0	705.33	0	0.729	0.021
			0.0000		0.0000		
Bonuses	726	187.65	0	288.18	0	0.412	0.038
	158		0.0000		0.0000		
Profits	6	461.82	0	549.65	0	0.657	0.023
	170		0.0000		0.0000		
Sales revenue	0	673.32	0	746.18	0	0.727	0.019
Material and energy	178		0.0000		0.0000		
costs	5	140.39	0	231.59	0	0.605	0.034
	177		0.0000		0.0000		
Prices	1	344.50	0	487.99	0	0.696	0.025

Note: The PT statistics reported as the square of s_n. We report also the χ^2 statistics as comparison.

Figure 1 shows the forecasts and their realizations in terms of diffusion indices of the seven measures of establishment performance for survey in April 2001 and December 2001. From this figure we can see clearly that the forecasts are higher than their realizations for all variables. This could be attributed to the September 11th. Although the realizations are always lower than their expected values, Table 3 shows that both PT and χ^2 tests reject the null hypothesis that there is no predictable relationship between forecasts and their realizations for all variables. Therefore, although there was September 11th shock, the establishment's forecasts can still predict their subsequent realizations quite well.

From Table 3 we can also see that the gamma coefficients range from 0.541 (bonuses) to 0.580 (employment). This means that if we pick up two establishments randomly in April 2001, then the probability that their realization order is equal to their forecast order less the probability that they have a different order, given no ties, range from 0.541 (bonuses) to 0.580 (employment). So the performance of forecasts is quite good and almost the same for all variables.

Figure 2: Forecasts and Actual of Diffusion Indices for 2001-2002
(Survey: October 2002 and April 2003)



The summary results in terms of diffusion index from the survey in October 2002 and April 2003 are presented in Figure 2. From this figure we can see that the forecasts are lower than their realizations except for product prices. But if we compare with Figure 1, the forecasts are seen to be closer to their realizations. Thus, from this figure we expect that the forecasts made in October 2002 can predict their realizations better than those made in April 2001.

Using both PT and χ^2 tests, Table 2 shows clearly that the null hypothesis that no predictable relationship is rejected for all variables. The gamma coefficients range from 0.567 (material and energy costs) to 0.709 (sales revenue). If we compare to the forecasts made in April 2001, all of these values but material and energy costs are greater. The differences can be due to two reasons. The first reason is the forecasts were conducted in two different months. In April 2001, the establishments forecasted the changes for the next 9 months, while in October 2002 the establishments forecasted only for the next 3 months. The second reason could be attributed to the September 11th. To measure the effect of the September 11th shock, we need to rule out the first cause by comparing the forecast made in April 2001 with the forecast made in April 2003.

Figure 3: Forecasts and Actual of Diffusion Indices for 2002-2003 (Survey: April 2003 and April 2004)

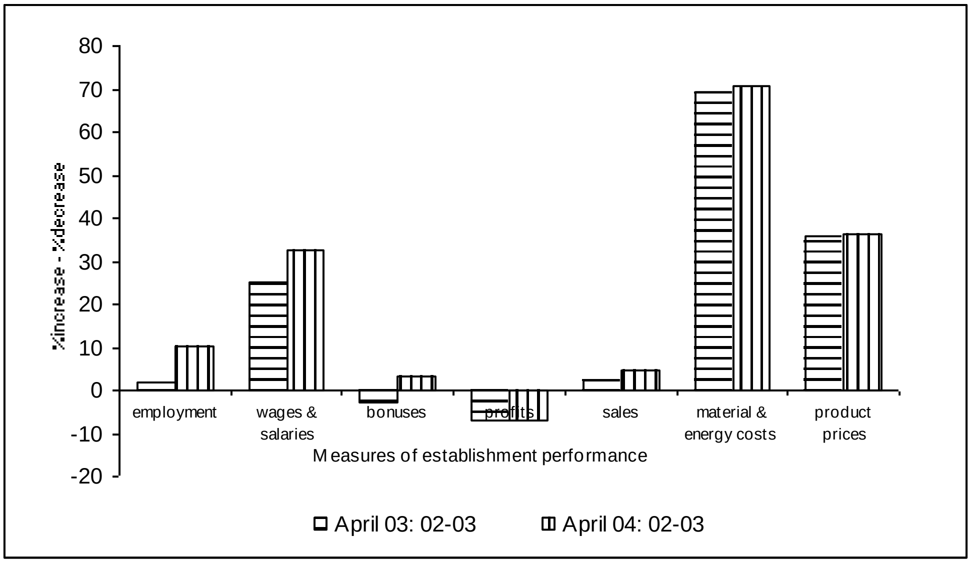


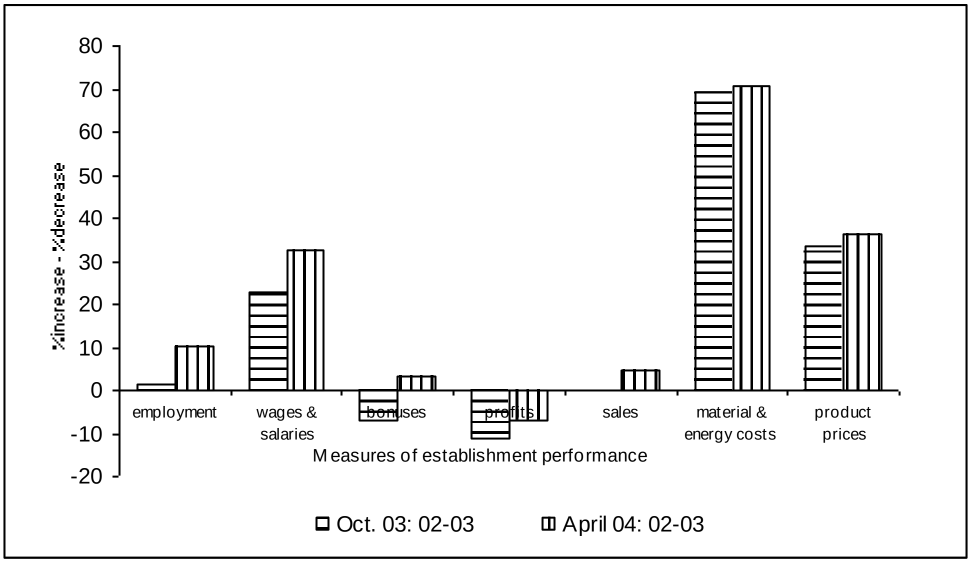
Figure 3 displays the summary of the results of the forecast made in April 2003 and their realizations. From this figure we can see that the forecasts are lower than their realizations except for profits. But if we compare to Figure 1, the forecasts in Figure 3 are closer to their realizations.

The results of the PT and χ^2 tests are presented in Table 2. Both tests reject the hypothesis that there are no predictable relationships between forecasts and their realizations for the seven measures of establishment performance. The gamma coefficients range from 0.492 (profits) to 0.674 (wages and salaries). These values are very similar to those of the forecasts made in April 2001. Using *z*-test[6], none of the variables are significantly different. Therefore, for each variable, the predictive performances of the forecasts made in April 2001 and in April 2003 have the same power statistically although there was September 11th shock in 2001.

Figure 4 displays the summary of the results of the forecasts made in October 2004 and their realizations. Similar to Figure 3 we can see that the forecasts are lower than their realizations. This figure is also very similar to Figure 2, where both figures show that the forecasts are lower than their realizations. This could indicate that the establishments were pessimistic about the economic conditions when they made forecasts for the near future.

Figure 4: Forecasts and Actual of Diffusion Indices for 2002-2003

(Survey: October 2003 and April 2004)



The results of the PT and χ^2 tests are presented in Table 2. Both tests reject the hypothesis that there are no predictable relationships between forecasts and their realizations for the seven measures of establishment performance. The gamma coefficients range from 0.412 (bonuses) to 0.729 (wages and salaries). These values are quite different from those of forecasts made in April 2003. The gamma coefficients for employment, profits, sales revenue, and prices are statistically higher than those of forecasts made in April 2003. Compared to the gamma coefficients from the forecasts made in October 2002, only bonuse variable is significantly different. Therefore, as expected, forecasts made in October can statistically improve the forecast accuracy if we compare to those made in April for most of the variables.

5. CONCLUSION

In this paper we test the predictive performance of the NYS Establishment Survey and measure how well establishment's forecasts predict the subsequent realizations of seven measures of establishment performance: employment, wages and salaries, profits, sales, material and energy costs, product prices, and bonuses. Our contribution is methodological in the sense that it shows the predictive value of directional business surveys using the actual realizations at the individual establishment level. Typically, establishment surveys are repeated cross sections and hence do not have the realizations at the level of individual firms.

The null hypothesis of no predictable relationship is rejected for the seven measures of establishment performance. These results are also supported graphically by the results of using diffusion indices. In addition, the rational expectation hypothesis is not rejected for most of variables. The predictive performance of forecasts made in October is statistically higher than those made in April for employment, profits, sales revenue, and prices. Thus the tests are capable to distinguish between surveys in terms of their predictive capabilities.

ENDNOTES

1. This definition is from Business Tendency Surveys: A Handbook
2. As OECD reported that in the OECD countries, this loss of information is unimportant for most uses of business tendency survey data.
3. We assume that survey on December 2001 asking change from 2000 to 2001 reflects the actual changes although year of 2001 has not finished.
4. All statistics in this paper are calculated using SAS/IML and SAS/Stat.
5. The criteria is that the REH (optimal prediction) is satisfied if and only if $p_{kk} \geq \max_{j \neq k} p_{kj}$ for all k=1, 2, .., m. Ivaldi (1992) used this method to check the REH and self-fulfilling hypothesis.
6. Statistic test to test the null hypothesis that $\gamma_{i1} = \gamma_{i2}$ is $z = \left(\hat{\gamma}_{i2} - \hat{\gamma}_{i1}\right) \Big/ \sqrt{s_{i1}^2 + s_{i2}^2}$.

REFERENCES

Agresti, A. (1984). Analysis of Ordinal Categorical Data: John Wiley & Son, Inc. page 186.

Dasgupta, S. and Lahiri, K. (1992): “A comparative study of alternative methods of quantifying qualitative survey responses using NAPM data,” *Journal of Business & Economic Statistics*, 10, 391-400.

Gourieroux, C. and Pradel, J. (1986). “Direct test of the rational expectation hypothesis,” *European Economic Review*, 30, 265-284.

Granger, C.W.J. and Pesaran, M.H. (2000): “Economic and statistical measures of forecast accuracy,” *Journal of Forecasting*, 19, 537-560.

Ivaldi, M. (1992): “Survey evidence on the rationality of expectation,” *Journal of Applied Econometrics*, 7, 225-241.

Lahiri, K., Peng, H. and Zhao, Y. (2015). “Testing the value of probability forecasts for calibrated combining”, *International Journal of Forecasting*, 31 (1), 113–129.

Maddala, G.S. (1983). Limited dependent and Qualitative Variables in Econometrics: Econometric Society Publication No. 3, Cambridge University Press, Cambridge, United Kingdom.

Nerlove, M. (1983): “Expectations, plans, and realization in theory and practice,” *Econometrica*, 51, 1251-1280.

OECD (2003), Business Tendency Surveys: A Handbook, Statistics.

Pesaran, M.H. and Timmermann, A. (1992): “A simple nonparametric test of predictive performance,” *Journal of Business & Economic Statistics*, 10, 461-465.

APPENDIX:

UNIVERSITY AT ALBANY - SUNY
ECONOMETRIC RESEARCH INSTITUTE
BA 110, 1400 Washington Ave, Albany, NY 12222 Ph. 518.442.4758 Fax 518.442.4736

NEW YORK STATE ESTABLISHMENT SURVEY 2004:2

1. How many employees does your establishment currently employ at its current location?

1 - 49 ☐ 50 - ☐ 100 or more ☐

2. What is the expected payroll employment outlook for your establishment at its current location?

Compared to 2003, employment in 2004 will: decrease ☐ remain the same ☐ increase ☐

Compared to 2004, employment in 2005 will: decrease ☐ remain the same ☐ increase ☐

3. What is the expected wages and salaries picture for your establishment at its current location?

Compared to 2003, wages and salaries in 2004 will: decrease ☐ remain the same ☐ increase ☐

Compared to 2004, wages and salaries in 2005 will: decrease ☐ remain the same ☐ increase ☐

4. What is the expected picture for employee bonuses for your establishment at its current location?

Our organization does not pay bonuses ☐

Compared to 2003, bonuses in 2004 will: decrease ☐ remain the same ☐ increase ☐

Compared to 2004, bonuses in 2005 will: decrease ☐ remain the same ☐ increase ☐

5. What is the expected profit picture for your establishment at its current location?

This is a not-for-profit organization ☐

Compared to 2003, profits in 2004 will: decrease ☐ remain the same ☐ increase ☐

Compared to 2004, profits in 2005 will: decrease ☐ remain the same ☐ increase ☐

6. What is the expected sales picture for your establishment at its current location?

Compared to 2003, sales revenue in 2004 will: decrease ☐ remain the same ☐ increase ☐

Compared to 2004, sales revenue in 2005 will: decrease ☐ remain the same ☐ increase ☐

7. What is the expected general picture for material and energy costs for your establishment at its current location?

Compared to 2003, material and energy costs in 2004 will: decrease ☐ remain the same ☐ increase

Compared to 2004, material and energy costs in 2005 will: decrease ☐ remain the same ☐ increase

8. What is the expected general picture for the prices of your products and services?

Compared to 2003, prices in 2004 will: decrease ☐ remain the same ☐ ☐ increase

Compared to 2004, prices in 2005 will: decrease ☐ remain the same ☐ ☐ increase

PLEASE COMPLETE OTHER SIDE

9. Relative to demand for your products and services, your production capacity is:

Too small ☐ Just right ☐ Too large ☐

10. Does your establishment have plans for new building and equipment in New York State in 2004?

Yes ☐ ☐ No

11. What most affects the performance of your establishment at its current location?

International economy ☐ national economy ☐ local economy ☐

12. Are you experiencing difficulties in obtaining supplies? ☐ Yes ☐ No

13. Are you experiencing difficulties in hiring skilled labor? ☐ Yes ☐ No

14. Are you experiencing difficulties in hiring unskilled labor? ☐ Yes ☐ No

15. What are the *three* most important factors in your business decision?

☐	Energy Costs	Interest Rates	☐
☐	Labor Costs	Availability of Credit	☐
☐	Rental Costs	Availability of Skilled Labor	☐
☐	Government Regulations	Availability of Unskilled Labor	☐
☐	State and Local Taxes	Availability of Raw Material	☐
☐	Other (please specify)____		

16. Over the next 5 years, do you expect your establishment to

Expand ☐ Contract ☐? Remain the same size ☐?

17. On a scale of 1 to 10, how do you rate the current health of the U. S. national economy (10 very robust, 1 very weak)? ____________

18. On a scale of 1 to 10, how do you rate the current health of the New York State economy (10 very robust, 1 very weak)? ____________

Please complete the information below to assist in determining the industrial classification of your firm and its location, and for a contact for future surveys. Your response will be kept confidential at all times, and no information will be released that will directly or indirectly identify your establishment.

Name of Contact Person:____

Title:____

Name of Establishment:____

Address:____

____

____

Phone: (____)_________________________ **Date:**____________________

Email address:____

In the future would you prefer to respond electronically? ☐ Yes ☐ No

Do Producer Prices Lead Consumer Prices? A Replication, and Extension Under Data Revisions

Corey J.M. Williams[*]

ABSTRACT

In 1995, Todd E. Clark of the Kansas City Federal Reserve wrote a seminal piece that asked the simple question: *do producer prices lead consumer prices*? Clark (1995) would become the foundation for the literature body on producer price inflation pass-through. Clark's principal findings illustrated that basic inflation forecasts were not *always* improved in forecast models that included producer price inflation relative to models that omitted producer prices altogether. These mixed results challenged the textbook notion that producer prices lead consumer prices. However, despite Clark's findings, over twenty-five years of new data has been made available alongside considerable revisions to the producer price index. The combination of new data and revisions to existing data makes a replication, and extension of Clark's work appealing, particularly considering the new inflation pangs plaguing the developed world currently, as well as continued strain to global value chains, which should be indirectly captured in producer price variation. Our replications results are similar in spirit to Clark (1995) while our contributions by extending Clark (1995) through the present produce mixed results, but at a lower rate of error than Clark (1995). These results suggest that the question of "when," rather than "if" producer prices lead consumer prices would be a high-potential avenue for future research.

INTRODUCTION

Open any principles of macroeconomics textbook and one would learn two stylized facts very quickly: firstly, it is a primary goal of policymakers to keep inflation stable; secondly, producer prices serve as a leading indicator of future consumer prices (Mankiw and Scarth, 2010). These two stylized facts have motivated three interrelated literature bodies that analyze pass-through of other price aggregates to consumer prices. Of these three pass-through literatures, the bulk of research tends to be oriented towards exchange rate and oil price inflation pass-through (Menon, 1995; Chen, 2009). Not unrelated, however, is the literature associated with producer prices, which while distinct from exchange rates and oil does indirectly capture variation, they could have at lower levels in the value chain on prices associated with producer raw materials and intermediate goods.

At the forefront of the producer price pass-through literature is Todd Clark's 1995 *Economic Review* paper from the Federal Reserve Bank of Kanas City. Titled simply as "Do Producer Prices Lead Consumer Prices," Clark looks at the predictive content of producer price inflation as a forecast variable for consumer price over select forecast periods. Motivated by the Bureau of Labor Statistics construction methodology

---

[*]Department of Economics, 225 DHC, Shippensburg University, Shippensburg, PA 17257. I would thank to thank Dr. Feng Yao, Dr. Alexander Cardazzi, and Dr. James Dean for early feedback on this work.

for the producer price index (PPI), Clark tests whether the links between producer prices across the production chain are as strongly tied to consumer prices as standard belief would have it. Using a parsimonious simultaneous equation forecasting model, Clark reveals that producer prices do not systematically assist in predicting consumers, contrary to popular belief. Moreover, forecast error from models that include producer prices seldom outperform forecasts models omitting producer prices altogether, particularly over shorter forecast horizons. Interestingly, however, Clark (1995) finds that over the full sample, producer prices do somewhat assist in forecasting, but in small magnitudes.

While Clark (1995) laid the foundation for most empirical examinations of producer price inflation pass-through, almost thirty years have passed since its publication. Furthermore, beyond normal data revisions, PPI itself has undergone a significant revision and change to its construction methodology in 2014. The BLS's new methodology moves farther away from the "weakly-linked" stage-of-processing characterization of producer prices and closer towards a classification based on weighted commodity groups. With improvements in construction technology and sampling procedures from the BLS, it is perhaps worthwhile to revisit the critiques put forth in Clark (1995) and extend his analysis through the present day. In essence, we ask ourselves the same simple question Todd Clark posed almost three decades ago but consider the possibility that his main results may change under data revisions and reclassification of the producer price index by the BLS.

The organization of this paper is as follows: first, we briefly highlight a handful of papers related to the producer price inflation pass-through literature that complement or are motivated in-part by Clark (1995). We then discuss the new 2014 BLS methodology for the construction of the producer price index and contrast it with its past methodology as described in Clark (1995). Next, we present an overview of the data and descriptive statistics relative to this replication effort. Finally, we replicate the methodology of Clark (1995) and compare our results to Clark (1995). Finally, we extend Clark (1995) and its methodology through the present day. Finally, we discuss the results of these exercises from a policymaker and practitioner context and conclude.

RELATED LITERATURE

The producer price index (PPI) has a storied history in the United States. As noted in Conforti (2016), PPI's largest draw is its ability to approximate and proxy for price fluctuations across the domestic value chain. As a result of this feature, PPI has typically been seen and heralded as a *leading indicator* for consumer prices. Prior to its current rendition, it was previously referred to as the wholesale price index (WPI) up until 1978 after which it was rebranded as PPI (Conforti, 2016). In documenting PPI's extensive history, Conforti (2016) notes that WPI (PPI) originally was composed from the weighting of only eight commodity groups: food, clothes, fuel, metals, lumber, chemicals, household goods, and miscellaneous goods. Presently, PPI encompasses fifteen different commodity groups and has undergone extensive revisions since its inception as WPI.

Despite the revisions to PPI over many decades, its overall utility has received mixed responses across research bodies. Belton and Nair-Reichert (2007) build off Granger et al. (1986) and stress that variance in the pass-through of producer prices to consumer prices are coming mostly from variation in food and energy commodity prices. The implication of these findings is the that pass-through of producer prices to consumers depends heavily on which measure of inflation one is examining. Furthermore, Belton and Nair-Reichert (2007) illustrate that removing food and energy prices from PPI and CPI altogether leads to a deterioration in their relationship with one another—this suggests that the cointegrating relationship (long-run relationship) between producer and consumer prices may be weakly linked to only a few highly volatile commodities. Consequently, producer price inflation's predictive content can fail in its utility regardless of whether the economy is in a low-inflation or high-inflation regime.

These statistical results notwithstanding, evidence of producer prices leading consumer tends to be robust across various countries, however, affirming the production chain view of producer price inflation pass-through. Studies like Jongwanich et al. (2019) find determinants like the food and energy prices tend to influence PPI inflation most. Alemu (2012) finds evidence of unidirectional cost-push pass-through of producer prices to consumer prices in South Africa in-line with the production chain view. Beyond, Alemu (2012), other cross-country studies like Tiwari et al. (2014) find bidirectional evidence of producer prices leading consumer prices in Mexico while Akcay (2011) finds unidirectional evidence of producer prices leading consumer prices in Turkey.

The state of the research on PPI pass-through can best be generalized as mixed. The literature in its current state suggests a few things: **1)** *which inflation* aggregates one is measuring matters for proper inference, **2)** the data environment matters in the sense that some nations may have stronger production chain linkages between producer and consumer prices than others, **3)** producer price pass-through is sensitive to the sample period, which is to say there are periods of time where producer prices lead consumer prices and other periods where the traditional production chain view fails to hold. These findings in totality make assessing the predictive content of producer price inflation a daunting task. For the scope of this study, we partially address *which* inflation aggregate matters (core CPI versus CPI, PPI for all commodities versus PPI for only industrial commodities) and touch upon *when* producer prices lead consumer prices as well following the most basic possible framework and extensions thereof outlined in Clark (1995).

A NOTE ON REVISIONS TO PPI

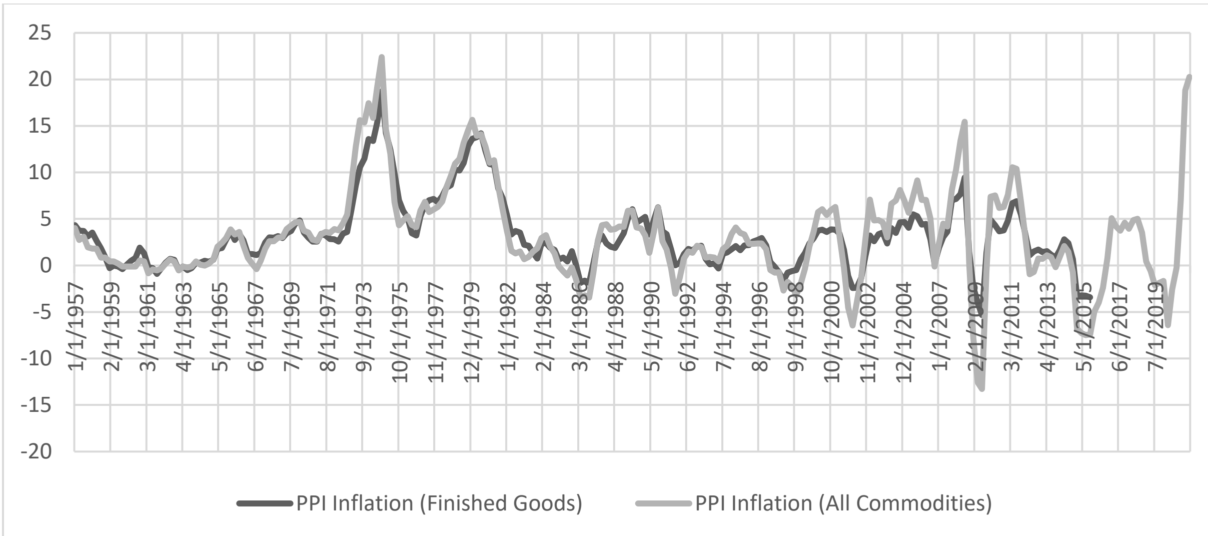
Prior to January of 2014, the producer price index published by the BLS was organized by three broad categories: crude materials, intermediate goods, and finished goods. This disaggregation by stage-of-fabrication allowed for PPI to, *in theory*, link all levels of the production chain to one another. Clark (1995) correctly points out the limitations to the organization of producer prices in this manner, namely that linking unprocessed-to-processed, and processed-to-finished goods were too broadly defined in scope to tightly emulate a true value chain. As such, weak linkages under such a broad scope could lead to instances where producer prices empirically fail to lead consumer prices.

As of today, the BLS producer price index is constructed through a sampling of goods producers, and service providers within the industries mining, forestry, utilities, construction, manufacturing, and services. The BLS notes that data coverage of mining, forestry, utilities, construction, and manufacturing is functionally 100%, while the services industry is only 70% covered. The result of this systematic sampling process, and index calculation produces roughly 3,700 commodity producer price indices for goods, and 800 for services. Beyond these, there is a minority of indices (roughly 600) that capture producer prices by final demand, and intermediate demand. The BLS notes that the current construction of the producer price index follows a relatively straightforward indexing procedure using a modified Laspeyres formula described below:

$$I_t = \left[\frac{\left(\Sigma \mathrm{Q_a} P_o \left(\frac{P_t}{P_o}\right)\right)}{\left(\Sigma Q_a P_o \left(\frac{P_{t-1}}{P_o}\right)\right)}\right] \times I_{t-1} \tag{1}$$

The index price, I_t, in the current period, t, is constructed effectively as a relative price of the sum of all shipments, Q_a, in some base period times the comparison period's price, P_o, and the sum of the same product, $Q_a P_o$, divided by last period's price relative to the comparison period. The benefits of the 2014 revision to the construction of the producer price index is twofold: firstly, by comparison to its ancestors (namely producer price index by commodity for crude materials, intermediate goods, finished goods, and the wholesale price index), the current producer price index has vastly more industry, and commodity coverage, which offers a pareto improvement in the index's ability to proxy for price variation in the production chain when compared a more limited scope index used in previous research, including Clark (1995); secondly, improvements in the sampling methodology emulate more modern probabilistic sampling approaches, rather than heuristic or "judgmental" approaches, which skewed PPI's construction towards high-volume commodities, and industries that weren't necessarily reflective of where the bulk of final demand falls for many products, and services. For comparison between the PPI inflation rate for finished goods (now discontinued) and the current PPI inflation rate for all commodities, see **Figure 1** below:

Figure 1. Historical Versus Contemporary PPI Inflation Series



Despite improvements to the methodology underlying PPI's revision, there are still *some* limitations to the current producer price index, namely that the coverage of the services industry is incomplete relative to goods producers, and goods-oriented industries. Given that the production-side of the economy is heavily composed of service-providers, it stands to reason that incomplete coverage when constructing PPI necessarily weakens the link between itself and CPI in periods where service consumption is particularly high. Secondly, despite BLS's attempts to rectify the commodity-leaning direction nature of PPI, the current index is still largely geared towards capturing, and reflecting producer commodity prices more than final demand. This once more limits the quality of linkages between end-stage production costs, and consumption of those same final goods, and services.

Furthermore, the now widely used PPI measurement as we know it is an "All Commodities" index, which is comprised of all weighted I_{jt} indices for j major commodity groups.[1] According to the BLS, there are fifteen commodity-specific indices that PPI can be broken down to such that:

$$PPI_t = \sum w_{jt} I_{jt} \tag{2}$$

The result of this weighting process is a PPI measurement that is more sensitive to variation in commodities depending on the weights, w_j, of each major commodity group. The weight assignment is based on the gross value of shipments corresponding to each commodity as reported by the US Census Bureau, thus PPI as indexed today is sensitive to changes in the measurement or revisions of shipments data by the Census Bureau.[2] This weighting procedure is a shift from the stage-of-processing techniques used to measure PPI around the time of Clark (1995), and thus will likely possess different predictive information than PPI prior to its 2014 revision. **Table 1** shows the major commodity groups that comprise PPI today along with their average relative importance from $2002—2021$.

Table 1. Weighted Components of PPI (2002—2021)

Pneumonic	Commodity Group	Avg. Relative Importance
WPU01	Farm Products	4.47%
WPU02	Processed Foods and Feeds	10.85%
WPU03	Textile Products and Apparel	1.73%
WPU04	Hides, Skins, Leather & related Products	0.14%
WPU05	Fuels and Related Products and Power	20.02%
WPU06	Chemicals and Allied Products	10.92%
WPU07	Rubber and Plastic Products	3.38%
WPU08	Lumber and Wood Products	1.73%
WPU09	Pulp, Paper, and Allied Products	5.88%
WPU10	Metals and Metal Products	9.81%
WPU11	Machinery and Equipment	12.81%
WPU12	Furniture and Household Durables	1.94%
WPU13	Nonmetallic Mineral Products	2.45%
WPU14	Transportation Equipment	10.43%
WPU15	Miscellaneous Products	3.44%

Table 1 gives a fair degree of insight into the coverage of the aggregate producer price index for all commodities (PPIACO). However useful the “relative importance” of these disaggregated indices may be, their weights vary on a yearly basis, rather than monthly or quarterly, which are standard frequencies for time series analysis of producer and consumer price inflation. Furthermore, even with the mixed frequencies between each disaggregated index and their relative importance, publicly available “weights” are only available as far back as the early 2000’s, making their use for most econometric analysis rather limited. Despite these limitations, **Table 1** ultimately provides clarity as to what commodities are covered by the BLS’s new methodology, as well as the extant of the importance of each commodity group—a marked improvement over the previous rendition of the producer price index and wholesale price index.

DATA

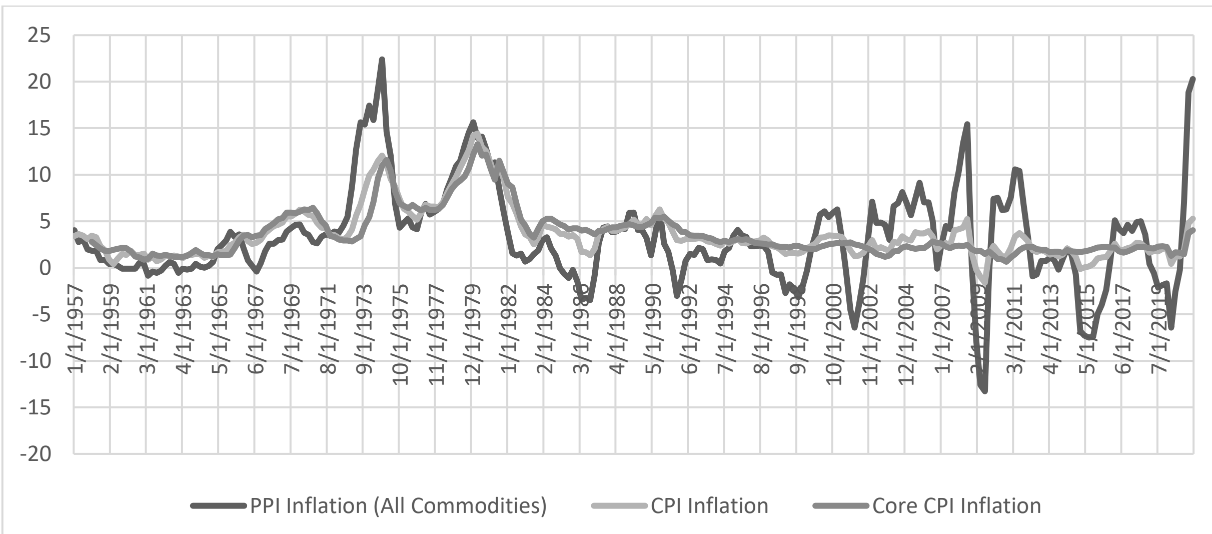
For our study, we have three main variables are interest that are all sourced from the Bureau of Labor Statistics (BLS) and retrieved via the Federal Reserve of St. Louis Economic Database (FRED). These variables are CPI inflation, core CPI inflation, and PPI inflation. This data is consistently of interest across most producer price pass-through studies. **Table 2** provides descriptive statistics for this data.

Table 2. Variable Descriptive Statistics $(1957{:}\,Q1\text{—}2021{:}\,Q3)$

Variable	Mean	Min	Max	Std. Dev.
CPI Inflation (π_t)	0.040	0.004	0.130	0.030
Core CPI Inflation (π_t^{CORE})	0.050	0.009	0.120	0.030
PPI Inflation (π_t^{PPI})	0.040	-0.040	0.200	0.050

Graphically, **Figure 2** below shows the evolution of these inflation rates over time:

Figure 2. Inflation Rates (PPI, CPI, Core CPI)



One would also notice from **Figure 2** that the correspondence between CPI inflation, core CPI inflation and PPI inflation is quite close and tends to reflect the production chain view theory of pass-through well up until middle of the 1990's. After the sample period Clark (1995), however, producer price inflation rates grow more volatile—at least visibly. This deviation in correspondence after the mid-1990's may also call into question the validity of the production view theory pass-through, which further motivates extending the analysis set forth in Clark (1995) through to the present day.

REPLICATION

We use this space to "replicate" the forecasting models utilized in Clark (1995). We stress that we expected different results given revisions to the PPI series made in 2014. As such, our results should be interpreted as not only addressing the question as to whether PPI inflation contains information useful for predicting future consumer price inflation, but *also* whether revisions to PPI since Clark (1995) have made PPI as constructed today a "better" forecast variable relative to its prior incarnation.

For evaluating the predictive content of producer price inflation, we utilize the most parsimonious model possible: a small-scale, unrestricted vector autoregression model (VAR)—the very same utilized in Clark (1995). Clark (1995) utilized two VAR models in their exercises: a baseline four variable VAR model that *excluded* producer price inflation, but included CPI inflation, the growth rate of GDP, the 3-month Treasury bill rate, and wage growth in the manufacturing sector.[3] Formally, our VAR models can be generalized by the equation below:

$$\Delta Z_t = \beta_0 + \sum_{\rho} \beta_\rho \Delta Z_{t-\rho} + \tau_t + \epsilon_t$$

ΔZ_t is our stationary vector of endogenous variables (which may or may not include producer price inflation), β_ρ is a vector of coefficients associated with our lagged matrix of stationary endogenous variables where ρ corresponds to the lag order; β_0 is a constant and τ_t is a linear time trend. We specifically present results for two different changes of our CPI inflation variable: one model with standard CPI inflation and one with core CPI inflation.

To be precise, we begin forecasting starting in 1959:Q2 using quarterly data. One-step ahead quarterly forecasts are then computed on a rolling basis with a restricted sample. At time t, data from 1959:Q2 is used up through period $T - h$ are used for model estimation where T represents the endpoint the data is rolled back from for creating our restricted sample and h represents the rolls themselves. From this point, quarterly one-step ahead forecasts are computed “out-of-sample” and then averaged from t through $t + 3$ periods ahead to form equivalent “yearly” forecasts.[4] To evaluate forecast precision, we calculate the mean absolute error (MAE) between our forecasts for inflation and actual inflation observed over the same period.[5] **Table 3** illustrates our first set of results that roughly correspond to Table 2 in Clark (1995).[6]

Table 3. Forecast Precision, Standard CPI Inflation

Model	**VAR Without PPI Inflation**		**VAR With PPI Inflation**	
Era	**Replication**	**Clark (1995)**	**Replication**	**Clark (1995)**
1977—1980	0.62	0.96	0.98	1.13
1986—1989	1.41	1.57	1.54	1.31
1991—1994	1.27	0.57	1.46	1.12
1977—1994	1.33	1.72	1.62	1.49

We note that the VAR model without PPI inflation yields numerical forecast errors that are comparable in magnitude (except for 1991—1994) to Clark (1995). Interestingly, the forecast model with PPI inflation yields strictly worse results across the board except for 1977—1980. Notwithstanding this, results are still analytically similar in magnitude. **Table 4** illustrates the second set of results that roughly correspond to Table 1 in Clark (1995).[7]

Table 4. Forecast Precision, Core CPI Inflation

Model	VAR Without PPI Inflation		VAR With PPI Inflation	
Era	Replication	Clark (1995)	Replication	Clark (1995)
1977—1980	1.15	1.58	0.94	1.08
1986—1989	1.45	0.54	0.65	0.79
1991—1994	0.97	0.30	0.31	0.80
1977—1994	1.41	1.15	1.24	1.37

From our replication, we see that core CPI inflation forecasts from models that *exclude* PPI inflation tend to underperform relative to models that include PPI inflation. We note that in-magnitude, the results from our VAR forecasts from our model without PPI inflation are similar in spirit to Clark (1995). Clearly, there are some exceptions with the middle two eras in our replication exercise, but full sample results, and 1977—1980 are comparable.

This is not the case, however, for forecasts from our VAR model that include PPI inflation. In fact, our replication shows that PPI assists considerably in reducing the forecast error over all sample and subsample periods in Clark (1995). Clark (1995) would suggest mixed results for gains in forecast precision when using PPI inflation as a forecast variable for predicting core CPI inflation. By contrast, our findings suggest that one should *always* include PPI inflation in core CPI inflation forecasting models of this variety.

This improvement in **Table 4** is unsurprising, however, for a few reasons. To begin with, the BLS has put in substantial effort to revise the producer price index in hopes of better reflecting the linkages across the production chain to core CPI inflation; as such, one would expect tighter linkages from higher quality sampling techniques to result in higher quality data for econometric and statistical analysis. Secondly, the BLS also revises the consumer price index to adjust for seasonal factors; such adjustments can go back as far as five years in length before finalized. In theory, a consequence of this is that consumer price inflation data used in Clark (1995) from 1989 through 1994 was still subject to revision and has subsequently undergone finalized revisions at the current time. Finally, while core CPI is revised for seasonal factors on a five-year rolling basis, seasonal factors themselves are also subject to change meaning new factors considered between 1995 through the present that are within the revision window also introduce new variation that could not be accounted for at the time of the Clark (1995) analysis.

Despite this, however our result from **Table 3** is a little more surprising, but still consistent with Clark (1995) in that PPI tends to assist in forecasting core CPI inflation with higher precision than alternative, but still common, measures of CPI inflation. Overall, even with revisions to both CPI and PPI since Clark (1995), results are still largely mixed. In the case of overall CPI inflation, the inclusion of producer price inflation in forecast models does not seem to be better on par compared to a model without PPI inflation at all. Similarly, models forecasting core CPI inflation tend towards lower forecast error in absolute terms in models when PPI inflation is included in the model compared to models without PPI inflation, but even so, there are still some eras where PPI does not improve forecast precision. Furthermore, in our replication, CPI inflation forecasts over the full sample are only improved by a small amount of 0.17 percentage points in relative forecast gain compared to a model without PPI inflation. This result stands in contrast, however, to Clark (1995) who finds that models without PPI inflation do not improve core inflation forecasts at all and are less accurate by 0.12 percentage points—this number, again, is small, however.

SAMPLE EXTENSION

Our replication results do well to reflect the spirit of Clark (1995). Having said that, however, significantly more data is available to researchers and forecasters today than in 1995. Beyond the availability of new data, new periods of economic growth and decline have occurred including the Dot-Com bust in the early 2000's, as well as the Financial Crisis of late 2007, and the Covid-19 pandemic of 2020. Beyond "new data," the literature on producer price pass-through informs us that "which" producer price inflation matters as well for understanding pass-through.

Following the guidance Jongwanich et al. (2019), we utilize an alternative measure of PPI to assess sensitivity of its predictive content: the producer price index for industrial commodities (PPIIDC in the BLS). PPIIDC is the weighted sum of the WPU03 through WPU15 commodities described in **Table 1**. This aggregate effectively omits volatility that could be coming from farm products and food, but not energy (WPU05)—we argue that inclusion of energy is important to account for variation in producer prices that could be coming from oil or petroleum, which act as significant sources of shocks within the production chain. Formally, we consider additional subsamples for rolling forecasts over: $2001—2006$, $2007—2009$, $2016—2019$, and $2001—2019$. For simplicity, we maintain the same lag-lengths prescribed in Clark (1995) for models with PPI ($\rho = 3$) and our model without PPI ($\rho = 5$). **Table 5** presents forecast precision results for forecasts of CPI inflation. Underlined MAE values indicate lower error compared to a forecast model without PPI:

Table 5. Forecast Precision (Extension), CPI Inflation

Model Era	VAR Without PPI Inflation	VAR With PPIACO	VAR With PPIIDC
$2001—2006$	0.54	1.36	0.85
$2007—2009$	2.14	<u>0.46</u>	<u>1.60</u>
$2016—2019$	0.74	1.14	0.90
$2001—2019$	1.01	<u>0.91</u>	1.15

Table 5 generally tells a story like Table 1 and Table 2 in that forecast models that include PPI inflation do not do particularly well at forecasting inflation in the short-term but perform adequately over longer subsamples. PPIIDC on the other hand does not seem to provide strong gains in forecast precision relative to models using PPIACO or traditional measures for PPI. We do note that forecasts over $2001—2019$ tend to be more precise compared to $1977—1994$ from Clark (1995). These improvements are on their own not enough to make claims about the quality of data for revisions for forecasting or make statements about the long-run dynamics between CPI and PPI inflation but do serve as a benchmark for future research interested in more rigorous analysis of inflation forecasting and cointegration between producer and consumer prices. Turning towards forecasts of core CPI inflation, Table 6 provides forecast error estimates for models comparable to those in Table 3 and Table 4.

Table 6. Forecast Precision (Extension), Core CPI Inflation

Model Era	VAR Without PPI Inflation	VAR With PPIACO	VAR With PPIIDC
2001—2006	0.94	0.95	1.37
2007—2009	0.64	0.69	0.71
2016—2019	0.27	0.37	0.37
2001—2019	0.50	0.66	0.88

We observe that, unlike Clark (1995), core CPI inflation from 2001—2019 is not improved by a model's inclusion of PPI inflation of the variety examined herein. Furthermore, forecast error for core CPI inflation is considerably lower by comparison to the 1977—1994 of interest to Clark (1995). These results reinforce that PPI's ability to assist in describing the dynamics of future inflation dependent strongly on which measure of consumer price inflation is of interest to the practitioner. Furthermore, forecast precision for CPI and core CPI inflation on symmetrical over the same subsamples across all models suggesting that *if* producer prices lead consumer prices, it is not uniform over time. Given these results, future research would do well to focus on the question of *when do producer prices lead consumer prices*?

MODEL EXTENSION

While we have successfully replicated Clark (1995), a lingering question, and natural extension, is *can we do better*? Clark (1995) was motivated predominantly by the cost-push view of producer price inflation pass-through in that costs associated with production, or subsequent factors of production, put upward pressure on producers that consumers then incur farther downstream along the value chain. Notably, the two dominant components of this cost-push view from the standpoint of the firm are energy costs, and labor costs. The producer price index indirectly captures energy costs through its WPU05 component as noted in **Table 1**, while manufacturing wage growth is a reasonable approximation for labor costs. Beyond these stylized facts, more advanced econometric techniques exist that can capitalize on time series characteristics in the data that cannot be efficiently capitalized on in an unrestricted VAR—most notably, the possibility of cointegrating relationships in our data.

To this end, we offer a modest extension of Clark (1995) testing the pass-through properties, and predictive content of the cost-push components of pass-through using a vector error-correction model (VECM) described by the following equation:

$$\Delta Z_t = A_0 + \Pi Z_{t-1} + \sum_{\rho=1}^{10} \Gamma_\rho \Delta Z_{t-\rho} + \varepsilon_t \tag{4}$$

Where ΔZ_t is a matrix containing our stationary endogenous variables, specifically CPI inflation, PPI inflation, and wage growth. Π can be decomposed to obtain the reduced form rank condition of our endogenous variables in their log levels such that $\Pi = \alpha\beta^T$, where α captures our error correction term(s), while β captures the long-run relationship(s) present in our data. As is the case with most vector error corrections models, two conditions must be met pre-estimation. The first is that all model data is integrated at an order of one, or simply, $\sim I(1)$, which is to say that the data is non-stationary in their levels, and stationary in when differenced or converted into growth rates. The second condition is that there must be $0 < r < R$ cointegrating relationships where R describes the maximum rank of the Π matrix—in the case of our data, $R = 3$. Using the Johansen test for cointegration, we discover that there are likely $r = 2$ cointegrating relationships at most.[8] The lag length of $\rho = 10$ is selected via the AIC criteria. From a notation standpoint, let our endogenous variables contained in our ΔZ_t matrix be denoted respectively as π_t, π_t^{PPI}, and π_t^w for CPI inflation, PPI inflation, and wage growth. **Table 7** below illustrates our equation-by-equation results for our VECM, including estimates of our $r = 2$ lagged error-correction terms, $ECT_{1,t-1}$, and $ECT_{2,t-1}$, respectively.[9]

Table 7. VECM Regression Results

Variable	π_t	π_t^{PPI}	π_t^w
A_0	0.11 (0.11)	-0.78 (0.57)	0.02 (0.14)
$ECT_{1,t-1}$	-0.08 (0.02)**	-0.20 (0.12).	-0.03 (0.03)
$ECT_{2,t-1}$	0.00 (0.01)	-0.22 (0.06)***	-0.01 (0.01)
π_{t-1}	0.46 (0.07)***	0.28 (0.36)	0.15 (0.09).
π_{t-1}^{PPI}	0.04 (0.02)*	0.59 (0.08)***	0.02 (0.02)
π_{t-1}^w	0.02 (0.06)	0.01 (0.31)	0.11 (0.08)
π_{t-2}	0.08 (0.07)	0.34 (0.39)	0.13 (0.10)
π_{t-2}^{PPI}	-0.01 (0.02)	0.03 (0.09)	-0.02 (0.02)
π_{t-2}^w	-0.07 (0.06)	-0.62 (0.30)*	-0.02 (0.08)
π_{t-3}	0.26 (0.08)***	0.03 (0.37)	-0.02 (0.09)
π_{t-3}^{PPI}	0.03 (0.02).	0.18 (0.08)*	0.02 (0.02)
π_{t-3}^w	-0.08 (0.05)	-0.02 (0.25)	0.09 (0.06)
π_{t-4}	-0.64 (0.08)***	0.15 (0.38)	0.02 (0.10)
π_{t-4}^{PPI}	-0.02 (0.02)	-0.53 (0.08)***	0.02 (0.02)
π_{t-4}^w	-0.03 (0.05)	-0.27 (0.25)	-0.67 (0.06)***

Note: $*** \ p < 0.001$; $** \ p < 0.01$; $* \ p < 0.05$; $. p < 0.10$. $N = 244$

We observe a few things from Table 7. Firstly, the short-run pass-through of producer prices to consumer prices is relatively small (0.04) over our sample; secondly, wage inflation as a contributing factor to cost-push inflation is not statistically different from zero in our CPI inflation VECM equation; finally, the speed of adjustments captured by our error-correction terms are relatively small in magnitude and significance in our CPI inflation equation but are moderately large and significant in the PPI inflation equation. In particular, the asymmetries present in adjustment speeds imply that in periods of disequilibrium, it is producer prices that error-correct or adjust in our system to restore equilibrium, not CPI inflation. While PPI may not lead CPI in its lagged components, it does seem to possess equilibrium-restoring properties not contained in the CPI inflation equation.

Unsurprisingly, a VECM model with cost-push endogenous variables presents richer results at the surface level by comparison to an unrestricted VAR. As such, we ask ourselves if a VECM can outperform our VAR models presented in Table 5. Table 8 presents the MAE from the same set of rolling forecast exercises. Our comparison VAR models with and without PPI inflation are included for comparison.

Table 8. Forecast Precision (VECM Extension), CPI Inflation

Model Era	VECM	VAR Without PPI Inflation	VAR With PPIACO
2001—2006	<u>0.84</u>	0.94	0.95
2007—2009	<u>0.41</u>	0.64	0.69
2016—2019	0.59	0.27	0.37
2001—2019	0.68	0.50	0.66

Table 8 informs us that a cost-push VECM can offer *some* small gains in forecast precision over a higher-dimension VAR model, however, the efficacy of such precision gains is limited to specific subsamples. Once more, the eras prior to the Financial Crisis and the Financial Crisis itself seem to be relatively more precise to forecast using a VECM, but such cost-push factors do not serve to improve forecast precision in the era after the Financial Crisis recovery and over longer subsamples horizons. In fact, in the latter cases, a VECM fails to outperform both baseline models. These results suggest that the choice of forecasting models has nontrivial implications for CPI inflation forecast precision, but the nature of both the forecasting horizon, and subsample of interest play a greater role considering the results of Table 5, and Table 6 in conjunction with Table 8. Simply put, the nature of *when do producer prices lead consumer prices* may be more compelling than the question of *do producer prices lead consumer prices* on its own.

CONCLUDING REMARKS

We replicate and extend the influential piece of Clark (1995), which evaluated the forecasting properties of producer price inflation begging the question as to whether producer prices truly lead consumer prices in a meaningful way to assist practitioners like forecasters and policymakers in understanding future consumer price inflation. The results of Clark (1995) were mixed at the time of their studies and indicated that PPI inflation over short periods did not reduce forecast error compared to forecast models without producer price inflation altogether; his results did, however, provide insights into how sensitive CPI inflation forecasts are over longer periods. Clark (1995) further underscored the sensitivity of how producer prices lead different consumer prices, especially core CPI inflation.

Our replication results are reflective of the spirit of Clark (1995) even considering new data revisions to PPI since 2014. When extending our analysis to the present day, we find mixed results like Clark (1995), but lower forecast errors on average across all forecasts compared to the sample period leveraged in Clark (1995) suggesting that a larger question *when* do producer prices lead consumers should be addressed in future research. Unlike Clark (1995), however, our extended sample period yields lower forecast errors for CPI inflation, rather than core CPI inflation. Furthermore, we consider the question of *which producer prices lead consumer prices* by extending our analysis to producer prices for industrial commodities; results show that the aggregate PPI over our sample period tends to producer lower forecast errors in applicable models compared to industrial commodities. Notwithstanding this finding, future research would also do well to disaggregate PPI further into its weighted commodity groups, rather than just looking at one group in isolation.

Finally, we ask ourselves if a pareto improvement in the sophistication of our forecasting model can offer gains in forecast precision relative to the Clark (1995) forecasting models. We also consider the possibility that the cost-push nature of inflation pass-through can be distilled to two specific variables: wage growth, and producer price inflation. With this in mind, we construct a trivariate VECM forecasting model, and find that *some* small gains in forecasting precision can be achieved, but greater inference suffers from consequences of choices in forecasting horizon, and subsample. These results once more suggest that cost-push inflation is not passing-through to CPI inflation unidirectionally over time, thus investigating *when producer prices lead consumer prices* may be a promising avenue for future research.

Overall, the results of this replication and extension affirm the initial results of Clark (1995). Producer prices *can* have a small impact on reducing inflation forecast errors in simple models when forecasting over long periods of time, but seldom outperform models that omit producer price inflation altogether over short horizons. Interestingly, however, forecasting core CPI inflation was better assisted by producer price inflation in Clark (1995), but standard CPI inflation is better assisted in our extended analysis suggesting that producer prices do not lead the same consumer prices over time.

ENDNOTES

1. This pneumonic is “PPIACO” according to the BLS.
2. The BLS does not formally publish the weights of the commodity-specific I_{jt} indices, rather than make publicly available information the annual relative importance of each j commodity-specific index.

3. We utilize the same data transformation process employed by Clark (1995). To transform our non-stationary data (GDP, CPI, PPI, and manufacturing wages), we take the natural logarithm of the ratio our variable of interest to its value four quarters ago such that $\Delta x_t = \log\left(\frac{x_t}{x_{t-4}}\right)$. All growth rates are then multiplied by a factor of $\times 400$.
4. See Clark (1995) for more details on this procedure.
5. The formula for calculating MAE is: $MAE = \left(\frac{1}{n}\right) \times \sum|x_i - \tilde{x}_i|$, where x_i is the actual observed data and $\tilde{x}_i$ are the predicted values from our forecast models.
6. It should be noted that the exact pneumonic utilized in Clark (1995) for “Core Goods Inflation” is not directly specified. We utilize the percentage change relative to a year ago for the series “CPIAUCSL” from the BLS as our goods inflation measurement for the sake of comparability in this exercise given that core CPI inflation is typically defined as the single BLS series “CPILFESL.” We opt for our lag length in the models used for forecasting “CPIAUCSL” to be $\rho = 6$ for the model without PPI and $\rho = 5$ for the models with PPI as per the original models in Clark (1995).
7. We opt for our lag length in the models used for forecasting “CPILFESL” to be $\rho = 3$ for the model without PPI and $\rho = 4$ for the model with PPI as per the original models in Clark (1995).
8. Results of our stationarity tests, and Johansen tests for cointegration are available upon request.
9. For brevity, we truncate our results in **Table 7** to extend out to only $\rho = 4$ lags of our endogenous. Despite the AIC recommendation of $\rho = 10$ lags, the estimates beyond the fourth lag tend towards minimal economic and statistical significance. Results beyond the fourth lag are available upon request.

REFERENCES

Akcay, S. (2011). The causal relationship between producer price index and consumer price index: Empirical evidence from selected European countries. *International Journal of Economics and Finance*, 3 (6), 227–232.

Alemu, Z. G. (2012). Causality links between consumer and producer price inflation in South Africa. Applied Economics Letters, 19(1), 13-18.

Belton, W. J., & Nair-Reichert, U. (2007). Inflation regimes, core inflation measures and the relationship between producer and consumer price inflation. *Applied Economics*, 39(10), 1295-1305.

Clark, T. E. (1995). Do producer prices lead consumer prices?. *Economic Review-Federal Reserve Bank of Kansas City*, *80*, 25-25.

Conforti, L. (2016). The first 50 years of the Producer Price Index: setting inflation expectations for today. *Monthly Lab. Rev*., 139, 1.

Granger, C. W., Robins, R. P. and Engle, R. F. (1986) Wholesale and retail prices: bivariate time-series modelling with forecastable error variance, in *Modelling Reliability* (Eds) D. A. Belsley and K. Edwin, MIT Press, Cambridge, Massachusetts, pp. 1–17

Jongwanich, J., Park, D., & Wongcharoen, P. (2019). Determinants of Producer Price versus Consumer price inflation in emerging Asia. *Journal of the Asia Pacific Economy*, 24(2), 224-251.

Mankiw, N. G., & Scarth, W. M. (2010). *Macroeconomics (Canadian Edition)*. Macmillan.

Tiwari, A. K., Suresh, K. G., Arouri, M., & Teulon, F. (2014). Causality between consumer price and producer price: evidence from Mexico. *Economic Modelling*, 36, 432-440.

Menon, J. (1995). Exchange rate pass-through. *Journal of Economic Surveys*, 9(2), 197-231.

Impact of Entrance Fees on Disc Golf Course Attendance: Empirical Insights from Heckscher State Park

Kenneth A. Liao[*§], Jimena González-Ramírez[†], and Kevin Meyer[‡]

ABSTRACT

This study examines the effects of seasonal congestion pricing on disc golf participation at Heckscher State Park using scorekeeping app data from January 1, 2015, to December 31, 2023. By comparing daily visits on fee and non-fee days and accounting for weather, holidays, and regional disc golf activity, we isolate the impact of pricing on recreational choices. Our findings reveal a significant negative relationship between course fees and attendance, demonstrating the importance of thoughtful pricing strategies in balancing revenue generation with accessible recreation. The research provides insights for policymakers and park administrators on optimizing public access while managing financial objectives.

JEL Codes: D40, H41, Q50, Z20

1. INTRODUCTION

Disc golf is a sport that combines traditional golf with the use of discs similar to small, hard frisbees, and it has steadily grown in popularity as a recreational activity worldwide. In disc golf, players aim to complete a course—usually consisting of 18 holes—by throwing a disc into a series of baskets with the fewest throws. Disc golf offers an environmentally friendly and accessible outdoor activity. This study explores the economic aspects of disc golf, with a particular focus on how course fees impact player engagement. Using data from a popular disc golf scorekeeping app, we analyze daily visitation trends at Heckscher State Park Disc Golf Course in New York from January 1, 2015, to December 31, 2023 to explore the effects of pricing strategies on participation.

This study examines how course fees impact the demand for disc golf. The analysis is important for decision-makers considering adding disc golf to public parks, as it provides empirical insights into how pricing affects usage and demand. We use Heckscher State Park's seasonal congestion pricing as a quasi-experimental setting to explore how players respond to different fee structures.

---

[*] Department of Economics, Farmingdale State College, Farmingdale, NY. Email: kenneth.liao@farmingdale.edu
[†] Department of Economics and Finance, O'Malley School of Business, Manhattan University, Riverdale, NY.
[‡] Department of Economics, Saginaw Valley State University, University Center, MI.
[§] We extend our gratitude to UDisc and their team, especially Doeke Buursma, Steve Hill, Josh Lichti, and Anthony Rentsch, for providing access to the data and supporting this study. Special thanks to our undergraduate research assistants, Taha Choudhry, Aisha Hussain, and Omar Pineiro from Farmingdale State College for their help with intercept surveys at Heckscher State Park in the summer of 2023, and to Robert Mack (Thieke Fellow) from Manhattan University for gathering the parking fee data. We also appreciate the constructive feedback and suggestions from two anonymous referees, which have significantly improved this paper.

Our methodology uses the unique pricing structure of Heckscher State Park and includes controls for weather conditions and broader trends in disc golf's popularity. We augment and contextualize our results using supplemental intercept surveys that provide further insights into player fee sensitivity, especially concerning the Empire Pass—an annual pass that nullifies daily parking fees. By combining the scorekeeping app data with the survey data, we can more accurately assess the true impact of parking fees on disc golf demand by considering cases where a portion of disc golfers are unaffected by the fee.

This paper is structured as follows. Section 2 reviews relevant literature, placing our research within the broader discussions of sports economics and public recreation management. Section 3 describes our data. Section 4 details our econometric model and analytical techniques. Section 5 presents our findings, showing the impact of pricing strategies on disc golf usage. Finally, Section 6 summarizes our conclusions and offers directions for future research into the economics of leisure and recreational sports.

2. LITERATURE REVIEW

The intersection of economics and recreational sports has seen growing interest in recent years, with research often focusing on consumer choices, motivations, and spending patterns. Studies like Mukanjari et al. (2021), Hesseln et al. (2003), Brown et al. (2021), Latinopoulos (2014), Hanauer and Reid (2017), and Orlowski and Wicker (2019) discuss factors such as pricing, natural disasters, the COVID-19 pandemic, recessions, and perceived value and how they affect consumer behavior in outdoor sports. Within this expanding field, disc golf stands out as a unique topic, blending the spontaneity of casual activities like hiking with the structure and rules of traditional sports.

Public parks and recreational areas create unique economic challenges and opportunities. Research by Ellis and Schwartz (2016), Middle et al. (2017), and Cohen et al. (2007) emphasizes the importance of these spaces in urban planning and community development. The literature indicates that parks not only improve quality of life (Larson et al., 2016) but also boost local economies (Mayer et al., 2010). Studies on parks specifically explore the balance between keeping these spaces accessible to the public and generating revenue, often through entrance fees, to support amenities (Banzhaf and Smith, 2020; Reynisdottir et al., 2008; Buckley, 2003).

The issue of pricing in public recreational facilities is complex, as explored by researchers like Sibly (2001) and Mulwa et al. (2018). Various models have been studied, including flat-rate fees (Banzhaf and Smith, 2020; Reynisdottir et al., 2008; Buckley, 2003), ticketed events (Smith, 2018), and dynamic pricing strategies (Lewison, 2017), similar to the seasonal congestion pricing used in New York State parks. These studies offer valuable insights into how different pricing strategies affect consumer behavior and park usage. An important element, relevant to our study of Heckscher State Park Disc Golf Course, is the effect passes like the Empire Pass have in reducing the impact of daily fees.

While the existing literature offers valuable insights into consumer behavior in sports and the economics of public parks, a gap remains in understanding the unique dynamics of disc golf, particularly in the context of variable pricing models. Notable contributions include González-Ramírez et al. (2023), who use disc golf scorekeeping app data to estimate demand when players choose between two similar courses—one with occasional fees and the other without. Similarly, Meyer et al. (2024) employ intercept surveys and a travel cost model to estimate demand among disc golfers. Our study herein aims to bridge this gap by providing empirical evidence on how course fees impact player turnout. Heckscher State Park Disc Golf Course serves as a unique case, as the lack of nearby alternative courses for Long Island's disc golfers makes it one of the most frequented courses in New York. This absence of substitutes also eliminates potential confounders in our demand analysis.

Our research addresses an important gap in understanding disc golf demand while contributing to the broader discussion on the economic viability and societal benefits of integrating sports facilities like these into public spaces. Our focus on outdoor recreation economics, with an emphasis on parks and disc golf course fees, provides the motivation for the empirical analysis that follows.

3. DATA DESCRIPTION

3.1 THE UDISC SCOREKEEPING APP

The UDisc scorekeeping app is widely used within the disc golf community. Beyond its primary function of scorekeeping, UDisc also serves as a social platform, a navigator to the nearest courses, and a personal progress tracker, allowing players to compare scores, share experiences, and find local tournaments. The app includes features that cater to both casual players and seasoned veterans, such as detailed maps, statistics, and a global course directory, providing a valuable source of disc golf data (UDisc, 2023).

The UDisc dataset provides a solid basis for our research due to its size and detail. It includes 1,987,244 anonymized visits, representing the activity of 205,275 unique disc golfers across 564 courses in New York, New Jersey, and Connecticut from January 1, 2015, to December 31, 2023. We treat each user's daily score log as a 'visit.' In cases where players logged scores more than once for the same course on the same day, we count those consecutive rounds as a single visit. The dataset offers a detailed look at play patterns, course choices, and shifts in disc golf's popularity over time. With this data, we analyze the behaviors and decisions of disc golfers and connect their activity to economic models of demand, accessibility, and willingness to pay for recreational sports.

3.2 HECKSCHER STATE PARK DISC GOLF COURSE

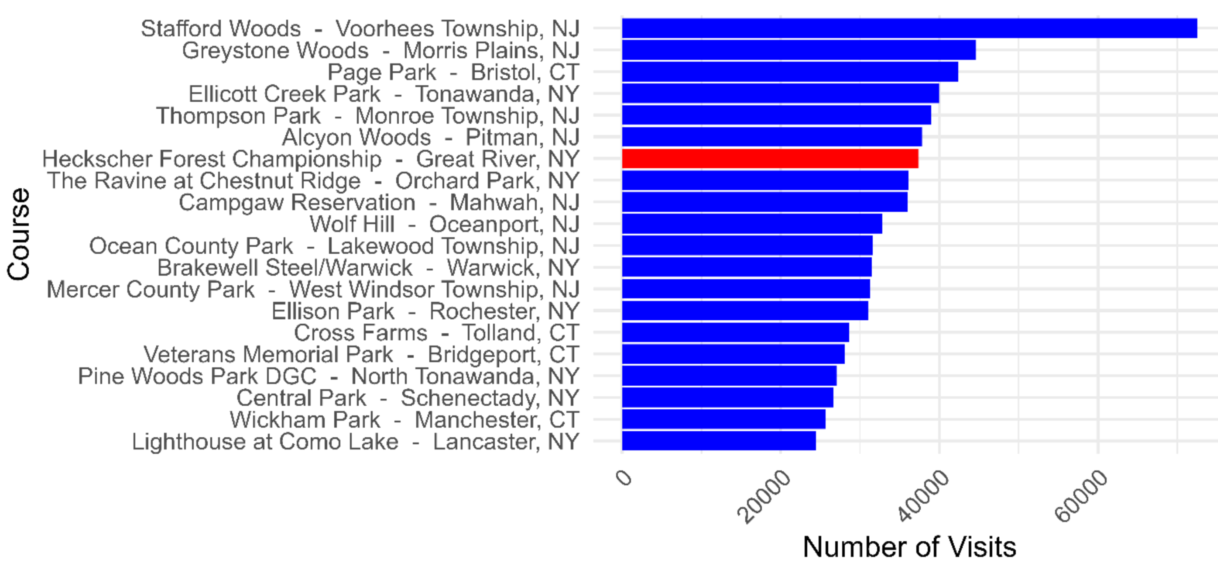
In the world of disc golf, a sport that has seen courses spread rapidly across the United States and globally, Heckscher State Park Disc Golf Course stands out for its unique location and qualities. Disc golf courses can now be found in a wide range of areas, from urban settings to more remote regions, wherever people seek recreation. However, New York City, despite its size, did not have a single disc golf course during our study period. (The first course in New York City, located in Highland Park in Brooklyn, opened in June 2024.) This shortage of courses makes Heckscher State Park especially important for Long Island residents, including those in the eastern parts of New York City. For millions of people on Long Island, Heckscher is not just a local state park; it is effectively the only nearby option for disc golf, with the next closest course requiring a one to two-hour drive.

Heckscher State Park Disc Golf Course is more than just a solitary option; it is a premier destination for disc golf enthusiasts. The course has 21 holes with a variety of playing experiences. It includes both short and long distances, a mix of wooded and open areas, and challenges suited for both beginners and seasoned players. The course is a favored site for both tournaments and casual play, featuring well-maintained facilities, top-quality baskets, clear signage, and excellent tee boxes.

The larger Heckscher State Park has a variety of amenities other than the disc golf course, including pavilions, walking and biking trails, and docks for water access. However, the main attraction for disc golfers is the disc golf course. It is rare for disc golfers to visit the park to do other activities in addition to disc golf. The course is located away from the park's busier areas and has its own parking lot. Disc golfers come to the park to play a round, and they typically leave the park afterward. While the other park facilities add to its overall appeal, they are of little importance to those visiting specifically for disc golf.

The standing of Heckscher State Park Disc Golf Course within the disc golf community is clear from its ranking in Figure 1. It is the second most visited course in New York and seventh overall among more than 500 courses across New York, New Jersey, and Connecticut. This ranking reflects both the quality of the course and its unique position as the only disc golf option on Long Island, giving it a significant advantage in drawing players compared to courses in areas with more competition. In New York, Ellicott Creek Park ranks first, and The Ravine at Chestnut Ridge ranks third, both located near Buffalo, a hub for disc golf. In Connecticut, Page Park in Bristol is notable, while in New Jersey, popular courses include Stafford Woods, Greystone Woods, Thompson Park, Alcyon Woods, Campgaw Reservation, and Wolf Hill.

Figure 1. Top 20 Disc Golf Courses in NY, NJ, and CT by Number of Visits (2015-2023)

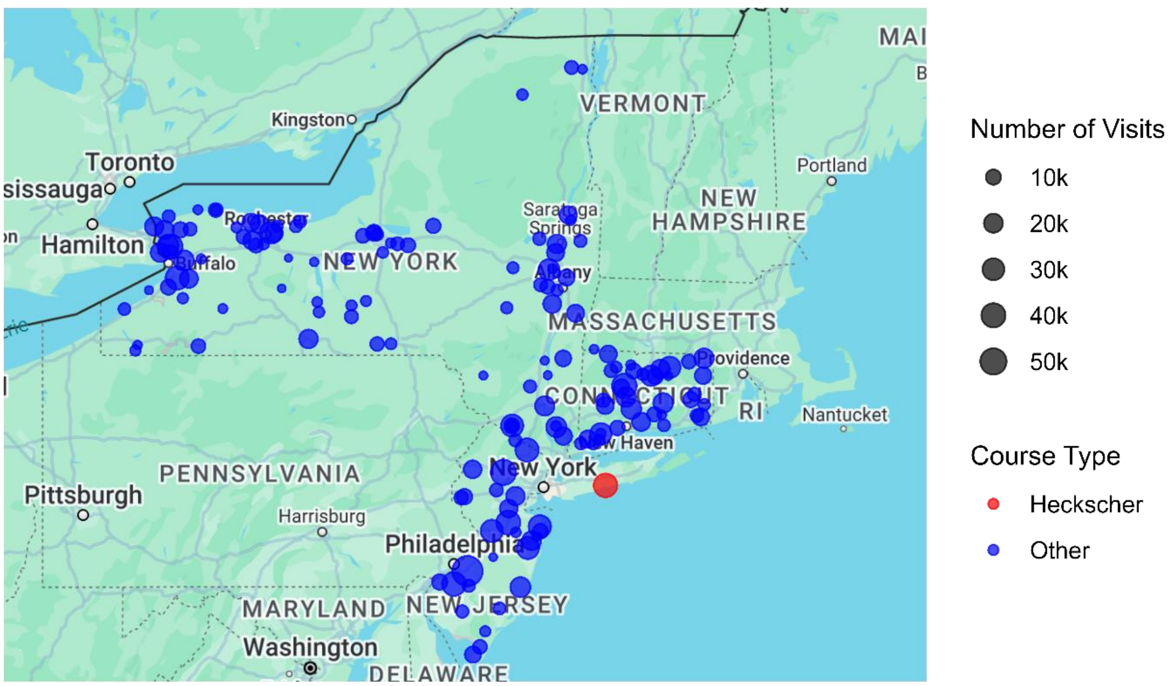


Heckscher State Park Disc Golf Course ranks seventh in visits among over 500 courses in New York, New Jersey, and Connecticut, reflecting both its quality and its unique status as Long Island's only disc golf course.

Not to be confused with the famous Central Park in Manhattan, Central Park in Schenectady, NY, also appears on the list. Schenectady is about a 3-hour drive north of New York City, further demonstrating the sport's reach beyond the immediate urban areas. The distribution of courses in Figure 1 shows disc golf's presence across the tri-state area, reinforcing Heckscher's importance to a large and underserved population of disc golfers on Long Island.

The spatial distribution of disc golf courses in our dataset is shown in Figures 2 and 3. The maps include all courses with significant activity, defined as those receiving at least 2,000 visits between 2015 and 2023. Figure 2 provides a broad view of the NY-NJ-CT region, showcasing 155 courses that account for 1,868,325 visits—representing 94.1% of the total visits in our dataset. The concentration of visits to certain popular courses confirms the representativeness of the courses included in the map.

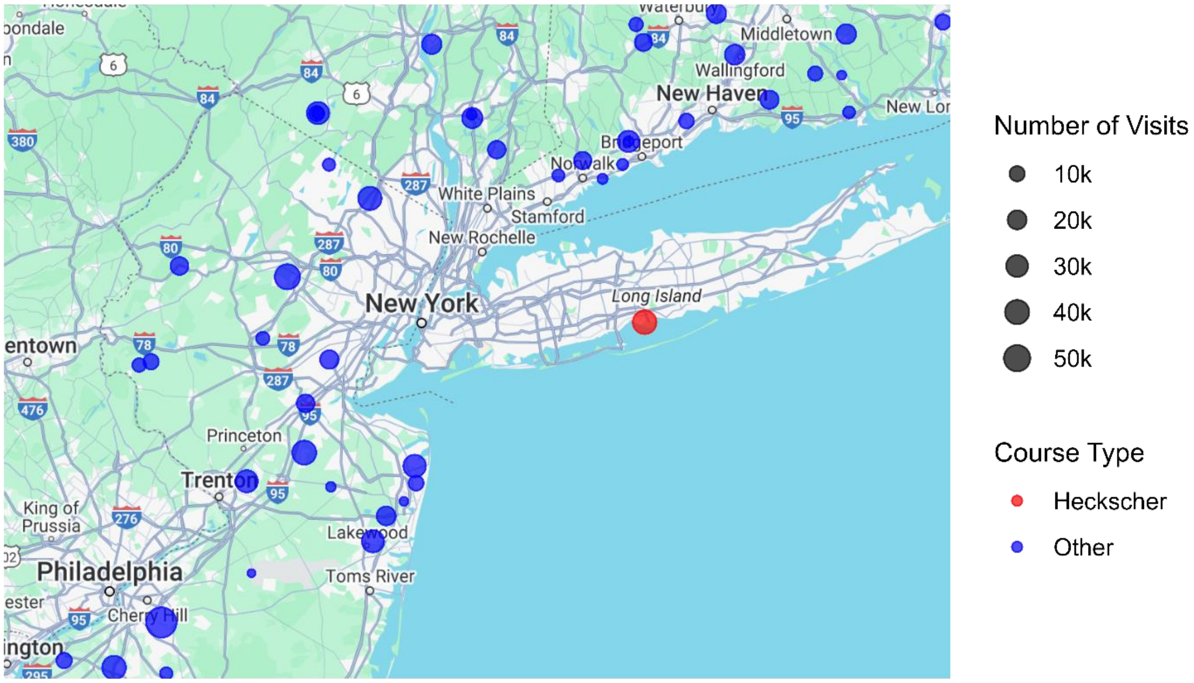
Figure 2. Spatial Distribution of Disc Golf Courses in NY, NJ, and CT (2015-2023)



The distribution of disc golf courses with at least 2,000 visits from 2015 to 2023 across the New York, New Jersey, and Connecticut region shows the widespread presence of courses near urban centers and densely populated areas.

Figure 3 provides a closer look at the New York City area, revealing the absence of disc golf courses within the city's limits. The map shows Heckscher State Park Disc Golf Course on Long Island, about an hour's drive east of New York City, along with other nearby courses in New Jersey and New York State, including Thompson Park, Oak Ridge Park, Greystone Woods, Campgaw Reservation, and FDR State Park. The lack of courses within the city stands in contrast to the relatively dense cluster in its surrounding areas. Data from the UDisc app and intercept surveys shows that players who visit Heckscher tend to play most of their rounds there. In contrast, players at other regional courses visit multiple locations. This further underscores Heckscher's unique position as a course with no close substitutes, as shown on the map.

Figure 3. Spatial Distribution of Disc Golf Courses in the Greater NYC Area (2015-2023)

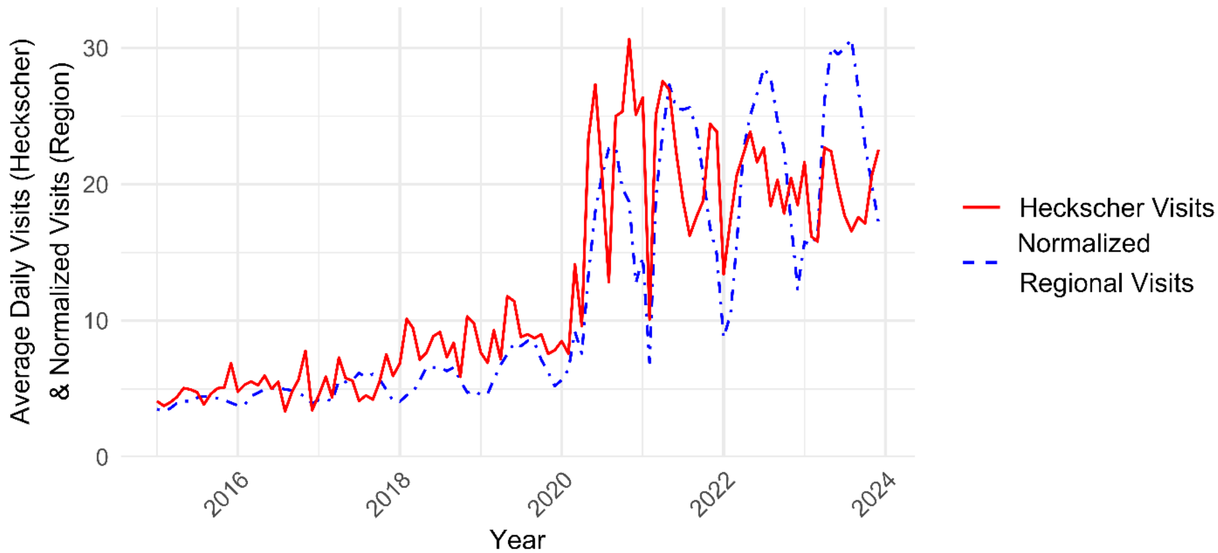


The focused map of the New York City region shows the lack of disc golf courses within the city and Long Island, except for Heckscher State Park, emphasizing the resultant reliance on Heckscher for the local disc golf community.

Figure 4 shows the visits to Heckscher State Park Disc Golf Course over time, compared with normalized regional visits from all other disc golf courses (excluding Heckscher) in New York, New Jersey, and Connecticut. The solid red line represents the average daily visits to Heckscher each month. For example, if a 30-day month had 600 total visits, the average daily visits would be 600 ÷ 30 = 20. The figure plots the average daily visits for each month from January 2015 through December 2023. The dashed blue line represents the normalized average daily visits to all other courses. Normalization adjusts the regional data so that the maximum and minimum values of the regional data align with the maximum and minimum values of the Heckscher data, allowing for easier comparison of trends.

The data shows a cyclical pattern in visits, with declines during colder months and rebounds in warmer seasons. The seasonality reflects the preference for milder weather, as disc golf is an outdoor and active sport. Both Heckscher and the broader tri-state region display an overall upward trend. Heckscher's growth was gradual from 2015 through 2019, with only a few months toward the end of 2018 and 2019 seeing more than 10 average daily visits. In the earlier years of the dataset, particularly between 2015 and 2017, monthly averages were often fewer than 5 daily visits, translating to less than 150 total visitors in a 30-day month. The regional data, by comparison, shows a more pronounced upward trend, especially in later years, reflecting both the growing popularity of the sport and an expanding disc golf community.

Figure 4. Average Daily Visits over Time: Heckscher vs. Normalized Regional (2015-2023)



The average daily visits for each month at Heckscher and across the tri-state region grew steadily until 2019, then sharply increased in 2020 with continued seasonal patterns, more evident in the normalized regional data.

A significant increase in activity occurred in 2020, with both Heckscher and regional average daily visits rising sharply, possibly due to the impact of the COVID-19 pandemic on recreational choices. Disc golf's compatibility with social distancing guidelines likely contributed to its appeal as a leisure activity during this time.[1] Additionally, the rise in visits may have been attributed to the sport's growing popularity and the increasing use of the UDisc app.

From February 2022 to December 2023, the data suggests that Heckscher's popularity stabilized, with average daily visits consistently ranging between 15 and 25 per month. The steady level of engagement is an indication of the course's established reputation and the lasting appeal of disc golf, even beyond the unique conditions brought about by the pandemic.

The figure also provides a broader comparison by juxtaposing Heckscher's trends with normalized regional visits. The regional data shows more pronounced seasonality, with sharper peaks and troughs compared to Heckscher. Between 2015 and 2019, both Heckscher and the broader region experienced similar growth rates, though Heckscher slightly outpaced regional growth, particularly in the lead-up to 2020. Both regions saw a significant increase in 2020, but the growth at Heckscher was notably stronger. After this surge, while regional attendance continued to rise gradually, Heckscher's growth slowed and slightly declined from 2020 to 2023.

The divergence in attendance after 2020 suggests that growth at Heckscher has plateaued, while visits across the broader region continued to rise. This trend shows the sustained popularity of disc golf and potentially the increasing use of the UDisc app within the community. In our empirical model, we use regional disc golf visits as a control variable for Heckscher's daily visits and include a time trend to capture other gradual shifts in attendance that might evolve linearly over time.

3.3 REFINED ANALYSIS TIMEFRAME: 2020-2023

Given the patterns and trends observed in Figure 4, we decided to narrow the focus of our research sample to the period from January 1, 2020, through December 31, 2023. This refinement is supported by several compelling factors in our initial broad analysis of the available nine-year dataset.

Concentration of Visits: A substantial proportion of the recorded visits to Heckscher—27,872 out of 37,367 (75%)—occurred within these four years. As a result, although the observation window is reduced from nine years to four, the dataset remains both rich and representative. At the regional level, 1,695,322 of the total 1,949,877 visits (87%) took place during the 2020-2023 period. This temporal concentration reflects the significant rise in disc golf activity, likely driven by the sport's growing popularity and increased use of the UDisc scorekeeping app.

Data Reliability and Relevance: The reliability of the data from 2020 onward is reinforced by several factors. Increased use of scorekeeping apps has paralleled the sport's growing popularity, providing a more accurate and comprehensive view of current disc golfing behavior. In contrast, the earlier years in the dataset—particularly 2015 and 2016—capture a time when both the sport and digital scorekeeping were less widespread.

Stability of Recent Data: After 2020, visitation data at Heckscher shows a more consistent pattern, with fluctuations becoming more predictable. This stability allows us to analyze the impact of variables like parking fees with greater precision. Changes in visits can be assessed against a relatively steady baseline, offering clearer insights into the effects of pricing policies.

Comprehensive Appraisal in Appendix A: For thoroughness and a comparative analysis, Appendix A presents an extended model using the full nine years of data, covering 3,287 observation days compared to 1,461 days in the focused four-year period. When we apply our model to the entire nine-year dataset, the results remain consistent with the findings from the four-year analysis, showing similar trends and effects. This reinforces the decision to concentrate on the 2020-2023 period, while the appendix offers validation of our approach and a broader perspective on disc golf activity.

In conclusion, focusing on the 2020-2023 period ensures that our analysis is grounded in a rich, reliable, and representative dataset, capturing the growth in disc golf participation and the increased use of the UDisc scorekeeping app. Earlier years reflect a time when disc golf and digital scorekeeping were far less common, with low and erratic daily visit numbers. Including those years could distort our understanding by giving equal weight to less reliable data. By narrowing our focus to the 1,461 observation days in the four years from 2020-2023, we ensure our results are based on more stable and high-quality data. This approach allows us to draw conclusions that are not only reflective of the sport's recent growth but are also supported by a robust and substantial body of reliable data from a period of higher engagement.

3.4 THE ENTRANCE FEE AT HECKSCHER STATE PARK

Unlike many disc golf courses that offer free access, Heckscher State Park charges a parking fee during certain times of the year. The fee is collected at booths near the park's entrance, where visitors are required to stop and pay per vehicle before entering. In New York, most state parks, including several on Long Island such as Belmont Lake, Robert Moses, and Sunken Meadow, follow a similar seasonal fee structure. However, Heckscher is the only park on Long Island with a disc golf course. Given the park's location and the absence of public transportation, nearly all visitors must arrive by car, making the fee virtually unavoidable. (As described in Section 3.7, the purchase of the Empire Pass can substitute for the daily fee.) The disc golf course is located deep within the park, so alternative access is impractical. Importantly, this parking fee is exogenous to our study—it affects the number of visitors to the park, but it is not affected by disc golf attendance or course management. The fee is a broader policy set by the New York State Parks Department, and all park visitors, regardless of their activities, are subject to it.

Although the park sometimes refers to this seasonal charge as a "congestion fee," it differs significantly from dynamic congestion pricing models used in some urban areas. Dynamic congestion pricing typically adjusts fees in real-time based on current demand to manage congestion, improve the user experience, and generate revenue. In contrast, the fee at Heckscher State Park is fixed by the calendar and applied on predictably high-demand days, such as weekends and during the summer months. As a result, even on these designated "congestion days," the disc golf course might be underutilized or completely empty, yet visitors are still required to pay the parking fee. The fixed seasonal pricing does not account for real-time congestion or actual course usage.

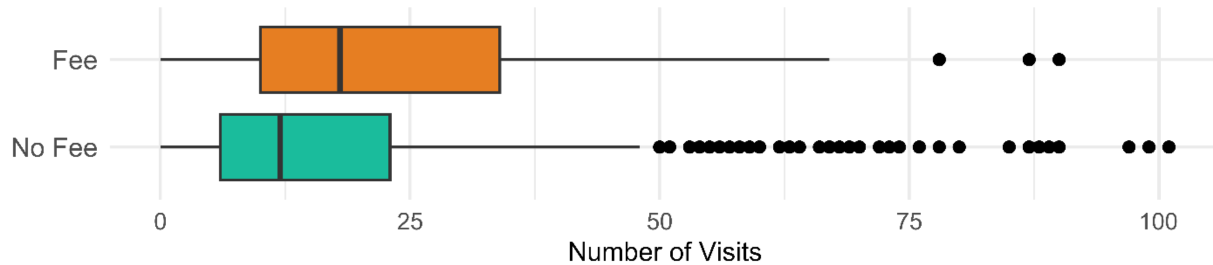
We obtained the fee schedule for Heckscher State Park from its official website and confirmed the details through direct communication with the park administration. The fee structure follows a seasonal pattern designed to manage visitor access and congestion. In the winter, parking is free every day. During spring, typically from early April to late May, a fee of $8 is charged on weekends. From late May to late June, the $8 fee applies every day. In the peak summer season, from late June through Labor Day, the parking fee increases to $10 per vehicle and is charged daily. After Labor Day, the fee reverts to $8 and is only charged on weekends through mid-October, after which parking is free again. In 2020, the fee structure was adjusted slightly, with fees not beginning until June, but otherwise following the same schedule.

The park's fee schedule offers an opportunity for economic analysis, as the price variation allows us to assess the impact of parking fees on attendance at the disc golf course. Although the differences in fees are modest, they still provide a basis for examining how slight changes in cost might affect visitor behavior. In Section 4, we extend our analysis beyond a simple fee vs. no fee comparison to include the actual amounts charged. This approach enables us to estimate how sensitive disc golf course visits are to changes in parking fees, offering valuable insights into demand within this setting.

The boxplot in Figure 5 provides an initial depiction of visitation patterns based on the parking fee data from Heckscher State Park. On both fee and non-fee days, there were instances of zero attendance.

However, the median number of visits was higher on fee days (18) compared to non-fee days (12). Similarly, the third quartile for fee days was 34 visits, while for non-fee days, it was 23.

Figure 5. Daily Visits to Heckscher State Park Disc Golf Course on Fee and Non-Fee Days (2020-2023)



The boxplots display the distribution of daily visits on fee and non-fee days from January 2020 to December 2023. The higher visitation on fee days corresponds to high-demand periods such as weekends and the summer season.

At first glance, the raw data might suggest that fees increase attendance. However, the higher numbers are primarily due to fees being charged on popular days, such as summer weekends. In Section 4, our econometric analysis accounts for variables such as weather, day of the week, and regional disc golf activity to control for these biases. Our econometric approach provides a better understanding compared to the initial, naive view in Figure 5, which simply compares attendance on fee versus non-fee days. Once adjusted, the analysis shows that the parking fee has a negative effect on attendance, underscoring the importance of controlling for external factors to accurately assess the impact of fees.

The distribution of visits skews towards higher numbers on both fee and non-fee days, as shown by the long upper whiskers and outliers. These outliers—representing days with significantly higher attendance—could be due to special events, ideal weather, or other irregular factors. The maximum number of visits was 90 on fee days and 101 on non-fee days.

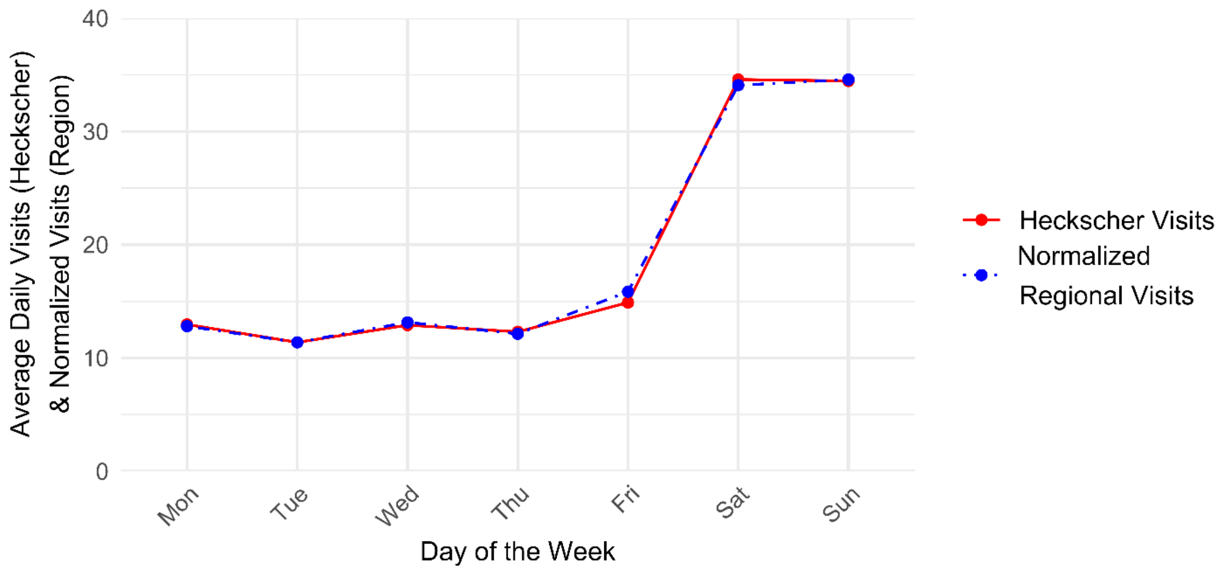
Given the four-year period from January 1, 2020, to December 31, 2023, which includes 1,461 total days (including dates with zero visitors), the data outliers represent the variability inherent in daily park attendance. These outlying observations are characteristic of distributions with right-skewed tendencies often encountered in visitor count data.

The outliers shown in Figure 5 are included in our regression analysis. These data points, identified by the standard boxplot rule (1.5 times the interquartile range above the 75th percentile), represent real fluctuations in attendance. Excluding them could bias the results by overlooking these important variations. On a related note, the daily visits data shows that capacity constraints are not an issue at Heckscher State Park—the expansive 21-hole course can handle dozens of players per hour, and even on the busiest days, attendance remains well within manageable limits. If capacity constraints were present, the observed effect of the parking fee would likely be smaller, as the implicit cost of playing at a popular, crowded course could be the time spent waiting at hole 1, rather than just the fee itself. We discuss the extrapolation of our results to other courses with different local dynamics further in Section 5.

3.5 DAY OF WEEK VARIATIONS IN DISC GOLF COURSE VISITATION

Figure 6 shows the daily visitation trends at Heckscher State Park, compared to normalized regional visits. Weekends see a marked increase in activity, consistent with the expectation that people typically have more leisure time. From Monday to Thursday, visits remain fairly low and stable, averaging between 12 and 13 visits per day. There is a slight uptick on Fridays, with an average of 15 visits, followed by a sharp rise on Saturdays and Sundays, where average visits are around 34 on both days. The figure aligns with common leisure behavior, where outdoor activities, including disc golf, are more frequent on weekends.

Figure 6. Average Daily Visits by Weekday: Heckscher vs. Normalized Regional (2020-2023)



The average daily visits to Heckscher by weekday alongside normalized regional visits from January 2020 to December 2023 show higher visits on weekends, with parallel trends across the tri-state region.

The figure also shows remarkably similar weekly changes in regional visitation, indicating that Heckscher's trends are consistent across the tri-state area. From Monday to Thursday, visits remain low, with a modest increase on Fridays, followed by a significant increase on weekends. The parallel across New York, New Jersey, and Connecticut reinforces the idea that the higher attendance on fee days at Heckscher reflects a broader regional trend in disc golf participation, tied to the typical workweek structure.

The main challenge with the observational data is separating the effect of the parking fee from the natural increase in weekend visits. To measure the fee's impact more accurately, our econometric model compares days with similar conditions—such as weather, day of the week, and regional disc golf activity—where the only difference is whether a fee was charged. For example, we compare two Saturdays or two Tuesdays with similar conditions, one with a fee and one without. This controlled comparison helps us isolate the fee's effect, avoiding the misleading assumption that higher weekend visits indicate a preference for paid entry.

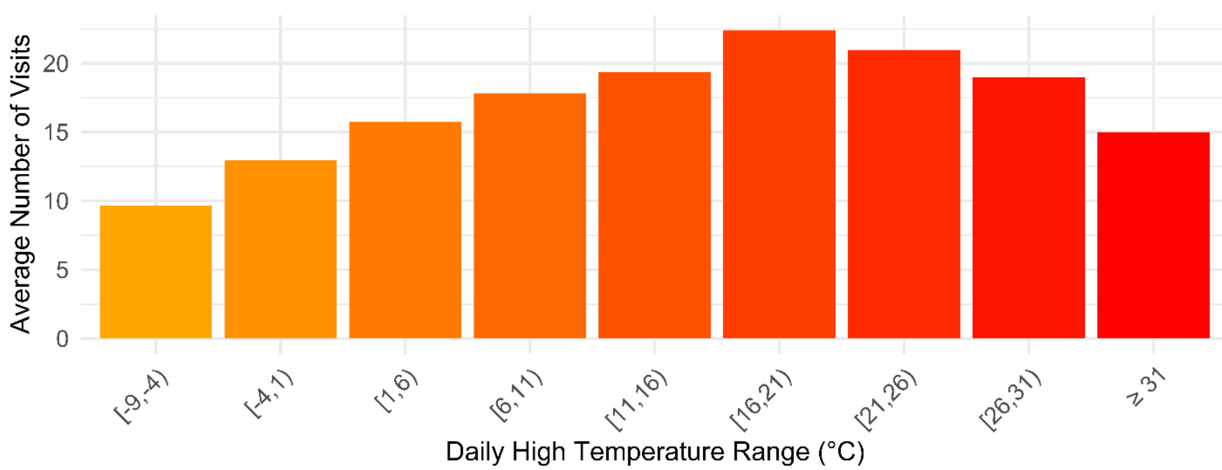
3.6 WEATHER DATA

To account for the effect of weather on disc golf visits, our analysis includes data on both temperature and precipitation. We used daily weather data from Oregon State University's PRISM dataset, which covers the entire contiguous United States with high-resolution data from over 800,000 grid points, downloaded from Dr. Aaron Smith's website (Smith, 2024). We combined the dataset with a shapefile representing U.S. zip codes. Finally, we calculated daily averages of the high temperatures and precipitation values specifically for the zip code encompassing Heckscher State Park, ensuring that our analysis reflects the local weather conditions experienced by disc golfers at the course.

The weather data shows a wide range of temperatures throughout the year at Heckscher State Park, with daily highs ranging from -8.6°C to 34.3°C during our study period. The median daily high temperature was 16.75°C, reflecting Long Island's temperate but distinctly seasonal climate. For precipitation, measured in centimeters, the median value is 0, indicating that over half of the days saw no rainfall. The mean precipitation was 0.34 cm, with a maximum recorded value of 11.30 cm. While rain is relatively infrequent, occasional heavy rainfall events do occur.

Figure 7 shows the relationship between daily high temperatures and visits to the disc golf course. As temperatures rise, the number of visits initially increases, reflecting more favorable conditions for outdoor activity. However, the trend reverses at high temperatures, presumably due to discomfort or health risks associated with extreme heat. The resulting inverse-U shape in the graph suggests a quadratic relationship, where moderate temperatures encourage more visits, but attendance drops off as the temperature exceeds certain thresholds.

Figure 7. Average Visits to Heckscher Disc Golf Course by Daily High Temperature (2020-2023)

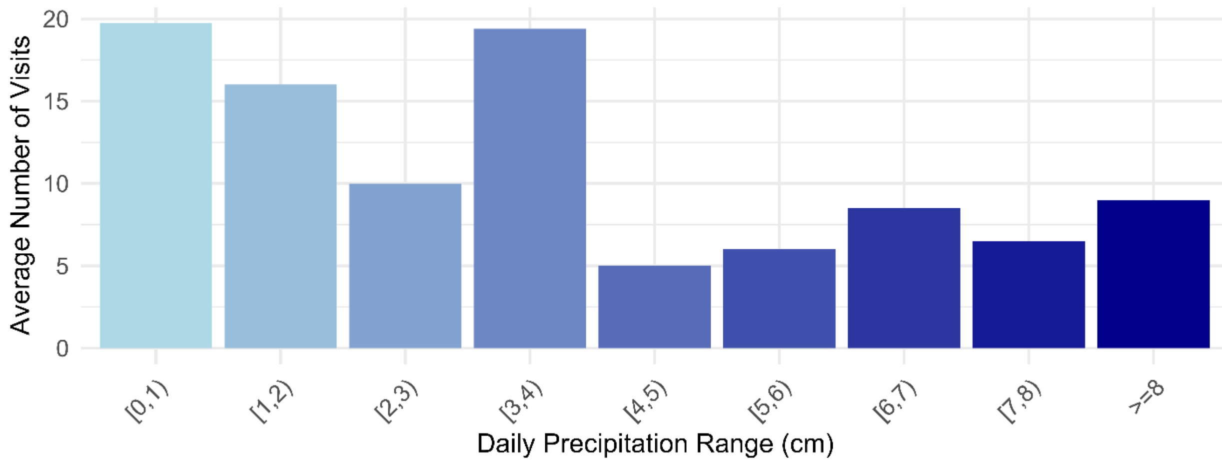


The average number of daily visits to Heckscher Disc Golf Course across different temperature ranges from Jan 2020 to Dec 2023 shows the increase in visits at moderate temperatures and the decline at higher temperatures.

Figure 8 shows the relationship between precipitation and course visits. As expected, visits decrease on days with higher rainfall, reflecting how unfavorable weather discourages outdoor activities like disc golf.

The data shows a significant drop in attendance on days with heavy rain, although such days were relatively uncommon during the period covered by our study.

Figure 8. Average Visits to Heckscher Disc Golf Course by Daily Precipitation (2020-2023)



The average number of daily visits to Heckscher Disc Golf Course across different precipitation ranges from Jan 2020 to Dec 2023 shows a decline as rainfall increases, with fewer visits on days with heavier precipitation.

In our regression model, we include both temperature and its square to capture the non-linear effects of weather, along with precipitation, to control for how these factors might affect daily visits. By accounting for these variables, we can more accurately isolate the impact of parking fees on attendance, ensuring that weather-related fluctuations do not obscure the true effects of the fee. The detailed results of this analysis are discussed in the following sections.

3.7 INTERCEPT SURVEY: CONTEXTUAL INSIGHTS AND MODEL CALIBRATION

In the summer of 2023, we conducted an intercept survey at Heckscher State Park as part of a broader research effort examining the economics of disc golf. The purpose of the survey was to assess how representative the UDisc app's visitation data was and to understand how parking fees affect player behavior. The survey was conducted over several days, including weekends and weekday afternoons, to capture a representative sample of disc golf activity. While the comprehensive survey informs a separate study on contingent behavior and travel costs (see Meyer et al. (2024) for a detailed discussion), the present paper uses a subset of relevant questions. The data enhances our current econometric analysis and aids in the transferability of our insights to the wider disc golfing community.[2]

The survey was conducted near the parking area and the first tee, where we anticipated the highest likelihood of engagement. We, along with our team of student research assistants, asked a series of structured questions directly to the disc golfers. We captured responses in real-time via the Qualtrics survey platform on various mobile devices, ensuring immediate and accurate data entry. The nonresponse rate was virtually zero, as almost all the disc golfers we approached were willing to participate, aided by the interspersed nature of their arrivals and departures and the availability of multiple devices for data collection. The survey was often administered while participants were warming up for their round, waiting for friends to arrive, or hanging around their vehicles after completing their round. We conducted the survey over several days to minimize selection bias and compile a robust dataset that we use for further calibration and contextual analysis.

Disc Golf Scorekeeping App Usage: In our survey, we asked disc golfers about their use of scorekeeping apps during play. Out of 104 respondents, the majority indicated that they used some form of digital scorekeeping. Specifically, 84 players reported using a scorekeeping app at varying frequencies above 0%, while 10 said they did not use any app. The remaining 10 participants did not answer the question.

For the 84 disc golfers who reported using a scorekeeping app, we followed up by asking which app they preferred, offering options like UDisc, Disc Golf Course Review, Disc Caddy, and others. An overwhelming 83 out of 84 users identified UDisc as their primary app. The strong preference for UDisc shows its prominence within the disc golf community. The detailed usage statistics for UDisc among the 83 respondents are shown in Table 1.

Table 1. Summary of UDisc Usage Percentages among Heckscher Disc Golfers Who Use the App

Statistic	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
UDisc Usage (%)	3	50	90	75.56	100	100

“What percentage of the time do you use a scorekeeping app when you play disc golf?”

The statistics show the app's adoption rates and how integral it is to the routines of disc golfers. With a median usage rate of 90%, the typical UDisc user uses the app for most of their rounds, and the mean usage rate of 75.56% suggests widespread preference for scorekeeping. As a result, the UDisc data provides a strong representation of overall disc golf activity, albeit a small percentage of rounds may go unrecorded. Additionally, we have no evidence to suggest that the two populations of golfers—*viz.* those who use the UDisc app and those who do not—would have different sensitivities to parking fees.

Use of the Empire Pass: One component of our study examining how parking fees affect the behavior of disc golfers involved asking survey respondents about their use of the Empire Pass, an annual pass that costs $75 and provides unlimited access to most New York State parks, effectively nullifying the daily parking fees. We also included questions about whether respondents were aware of the parking fees beforehand when deciding to visit a disc golf course. The questions were important for understanding how parking fees might impact visitation rates and contextualizing our overall analysis.

The responses to the Empire Pass question revealed an interesting distribution: 59 out of 104 respondents reported regularly using the Empire Pass, suggesting a potential indifference to daily parking fees at state parks. On the other hand, 42 respondents did not have the pass, indicating they may be more sensitive to parking fees. Three respondents did not provide an answer to this question.

Regarding parking fee awareness, 65 out of 104 respondents stated that they were generally aware of parking fees at disc golf courses, which may affect their decision to visit certain locations. In contrast, 37 respondents were unaware of the fees (most of whom were Empire Pass holders), and 2 did not answer the question. Table 2 summarizes the responses to questions regarding Empire Pass possession and awareness of parking fees at the disc golf course.

Table 2. Summary of Responses to Questions on Empire Pass and Parking Fee Awareness

Question	Yes	No	NA
Empire Pass Possession	59	42	3
Parking Fee Awareness	65	37	2

"Do you typically use an Empire Pass (regional parking pass for NY State Parks)?" and "When you make the choice to visit a disc golf course, do you typically know when/whether the park has a vehicle/parking fee?"

The responses reveal significant variation in the potential economic behavior of park visitors. Disc golfers with an Empire Pass may be largely unaffected by parking fees, while those without the pass might adjust their visits based on fee schedules. The distinction offers valuable insight into the elasticity of demand concerning parking fees and helps refine our econometric model. In Section 5, we examine how the existence of the Empire Pass affects our interpretation of the econometric results and discuss its broader implications for applying the findings to other courses.

3.8 SUMMARY STATISTICS

Our main dataset brings together UDisc course visit data, seasonal pricing information, localized weather conditions, and insights from our intercept survey, creating a comprehensive foundation for analyzing the impact of park fees on disc golf course attendance. While the intercept survey responses are not directly included in our regression model, they provide important context that informs the broader discussion and helps refine our analytical framework. Table 3 presents summary statistics for the variables in our study.

Table 3. Summary Statistics for Regression Variables

Variable Description	Min	Median	Mean	SD	Max
Daily Visits to Heckscher	0	14	19.08	17.03	101
Daily Visits to All Other Courses (NY-NJ-CT)	19	1098	1160.38	754.12	3628
Parking Fee (1 = Yes, 0 = No)	0	0	0.34	0.47	1
Parking Fee Amount ($)	0	0	3.14	4.43	10
Parking Fee Amount in Dec 2023 $	0	0	3.37	4.77	11.50
Daily High Temperature (°C)	-8.56	16.75	16.48	8.90	34.27
Daily Precipitation (cm)	0	0	0.34	0.87	11.30

Summary statistics are based on $N = 1461$ observation days from Jan 1, 2020, through Dec 31, 2023.

The summary statistics show that the average number of daily visits to Heckscher State Park Disc Golf Course was 19.08 during our observation period. This is naturally much lower than the average of 1,160.38 daily visits recorded for all other courses in New York, New Jersey, and Connecticut combined. While Heckscher ranks as the seventh most popular course across these three states, the sheer number of courses means that Heckscher ultimately accounts for only about 19.08 ÷ (1160.38 + 19.08) ≈ 1.6% of the region's total disc golf activity.

Parking fees were in effect on approximately 34% of the observation days, primarily during the summer and on weekends in the spring and fall. We differentiate between the nominal and real parking fee amounts to account for the gradual decrease in the fee's real value over time; the \$8 or \$10 charge has remained unchanged for many years, effectively decreasing in real terms due to inflation. Real fees are adjusted to December 2023 dollars using the monthly Consumer Price Index for Urban Consumers (CPI-U) for the New York-Newark-Jersey City, NY-NJ-PA area from the Bureau of Labor Statistics (BLS, 2024). Although our main analysis focuses on the binary presence of a fee, we also explore the impact of varying fee amounts. We exploit the slight differences in parking fees, and we estimate how these small variations affect park visits. The approach enhances our understanding of disc golfers' price sensitivity and adds to discussions about park revenues and the economic value of public recreational spaces.

4. MODEL SPECIFICATION AND ECONOMETRIC APPROACH

To measure the impact of parking fees on the number of visits to the Heckscher State Park Disc Golf Course, we compare daily visits on fee days with those on non-fee days, while carefully controlling for other factors that could affect attendance. The econometric model we use is outlined in the following equation:

$$\begin{aligned}\mathrm{Visits}_t = \beta_0 &+ \beta_1 \mathrm{Fee}_t + \beta_2 \mathrm{RegionalCourseVisits}_t + \beta_3 \mathrm{Temperature}_t + \beta_4 \mathrm{Temperature}_t{}^2 \\ &+ \beta_5 \mathrm{Precipitation}_t + \beta_6 \mathrm{Precipitation}_{t-1} + \gamma_t + \delta_w + \zeta_h + \phi_y + \epsilon_t \qquad (1)\end{aligned}$$

In the model, Visits_t is the total number of visits to Heckscher State Park Disc Golf Course on day t, while Fee_t is a binary variable indicating whether a parking fee was charged that day. $\mathrm{RegionalCourseVisits}_t$ is the total number of visits to all other disc golf courses in New York, New Jersey, and Connecticut (excluding Heckscher) and serves as a proxy for the general popularity of disc golf and UDisc app usage in the tri-state area. The model includes both $\mathrm{Temperature}_t$ and its square to account for the non-linear relationship between temperature and visits, and $\mathrm{Precipitation}_t$ alongside its lagged value ($\mathrm{Precipitation}_{t-1}$) to account for immediate and delayed rain effects like wet or muddy conditions the next day. We also include γ_t (daily time trend), δ_w (day-of-the-week fixed effects), ζ_h (holiday fixed effects), and ϕ_y (year fixed effects) to control for any temporal effects that could impact visits. Lastly, ϵ_t is the error term, capturing any unobserved factors impacting daily visits.

The control setup ensures that the days being compared are as similar as possible in terms of the day of the week, weather conditions, regional disc golf activity, and other important variables, allowing us to isolate the specific impact of the parking fee on attendance. Equation (1) outlines our primary model, hereafter referred to as Model 1, which examines the binary impact of the parking fee—that is, whether a fee is charged or not—on the number of daily visits to the disc golf course.

To gain a more detailed understanding of how parking fees affect disc golfer behavior, we introduce a second model, Model 2. The model follows the same estimating equation as Model 1 but replaces the binary Fee_t variable with the actual parking fee amount, adjusted to December 2023 dollars. The nominal parking fees during the observation period were $0, $8, or $10. By adjusting for inflation, we can examine the fee in real terms to see how different fee levels affect visits. Formally, Model 2 can be expressed as follows:

$$\begin{aligned}\mathrm{Visits}_t = \beta_0 + \beta_1 \mathrm{FeeReal}_t + \beta_2 \mathrm{RegionalCourseVisits}_t + \beta_3 \mathrm{Temperature}_t + \beta_4 \mathrm{Temperature}_t{}^2 \\ + \beta_5 \mathrm{Precipitation}_t + \beta_6 \mathrm{Precipitation}_{t-1} + \gamma_t + \delta_w + \zeta_h + \phi_y + \epsilon_t \qquad (2)\end{aligned}$$

In Model 2, $\mathrm{FeeReal}_t$ is the parking fee adjusted to December 2023 dollars. The coefficient associated with this variable, β_1, captures the effect of each additional dollar of the parking fee on the number of visitors, providing additional insight to the binary analysis in Model 1.

To estimate the models, we use a fixed effects Ordinary Least Squares (OLS) approach. We model the daily visits to Heckscher State Park Disc Golf Course based on the fee, daily regional visits, and the local weather, accounting for factors that vary across days but are consistent within similar periods. For example, we expect more disc golfers on Saturdays than Tuesdays, and this approach allows us to compare Saturdays with other Saturdays and Tuesdays with other Tuesdays to isolate the impact of the parking fee. We include fixed effects for day of the week, holidays, and years, as well as a linear time trend, to control for systematic variations. The method minimizes omitted variable bias by comparing similar days over time.

Our analysis uses heteroskedasticity-robust standard errors to address the potential issue of non-constant variance in the error terms. When heteroskedasticity is present, it can distort estimates and affect the accuracy of standard error calculations, which might lead to incorrect conclusions about the significance of the variables. We adjust for the issue by applying heteroskedasticity-robust standard errors, ensuring that our inferences remain reliable even when errors vary across observations.

5. RESULTS

5.1 ECONOMETRIC RESULTS

Our econometric analysis, consisting of both Model 1 and Model 2, provides valuable insights into how parking fees affect visitation to the Heckscher State Park Disc Golf Course. Table 4 summarizes our findings and shows the impact of parking fees along with the other factors that affect daily visits.

The main result from Model 1 is that the parking fee at Heckscher State Park reduces average daily visits by approximately 4.756, holding all other factors constant. For example, if a non-fee day typically attracts 25 disc golfers, a comparable fee day—where weather, day of the week, and regional disc golf activity are similar—would draw around 20 golfers. The decrease in visits indicates that additional costs reduce participation in recreational activities, consistent with theoretical expectations.

Table 4. Econometric Analysis Results

Variable Description	Model 1 (Fee Presence)	Model 2 (Fee Amount)
Fee Presence (1=Yes, 0=No)	–4.756*** (1.033)	
Fee Amount (Dec '23 $)		–0.443*** (0.104)
Daily Regional Visits (NY-NJ-CT)	0.018*** (0.001)	0.018*** (0.001)
Daily High Temperature (°C)	–0.186 (0.130)	–0.210 (0.132)
Daily High Temperature Squared	–0.008** (0.004)	–0.007* (0.004)
Daily Precipitation (cm)	–1.046*** (0.342)	–1.056*** (0.345)
One-Day Lagged Precipitation	–0.268 (0.281)	–0.268 (0.283)
Daily Time Trend	0.001 (0.003)	0.001 (0.003)
Day of Week Fixed Effects	✓	✓
Holiday Fixed Effects	✓	✓
Year Fixed Effects	✓	✓
Observations	1461	1461
Adjusted R-squared	0.621	0.619
Akaike Information Criterion (AIC)	11032.3	11037.1
Root Mean Square Error (RMSE)	10.43	10.44

The dependent variable is Daily Visits to Heckscher. Standard errors are heteroskedasticity-robust and are reported in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

In Model 2, examining the impact of the parking fee's amount shows that each additional dollar leads to a reduction of about 0.443 visits. For example, an $8 fee would reduce visits by roughly 3.54 (0.443 visits per dollar × 8 dollars), while a $10 fee would reduce visits by about 4.43. The figures align with the findings from Model 1. In terms of percentages, the decreases represent 18.6% and 23.2% of the 19.08 average daily visits, respectively. Also note that the fees are measured in December 2023 dollars, so a $10 fee in 2021 was roughly equivalent to an $11 fee in 2023, accounting for inflation.

The revenue implications of these findings are important for park management. Based on the observed sensitivity to parking fees, setting the fee at $10 instead of $8 would increase revenue by $2 per visitor, but it would also reduce the number of visitors. To fully understand the trade-off, we can calculate the overall revenue by multiplying the fee by the expected number of visitors. For example, suppose the park attracts an average of 19.08 visitors per day when there is no fee. Increasing the fee by $8 reduces visits by 3.54, bringing the average down to 15.54 visitors per day, which generates $124.32 in daily revenue. Increasing the fee by \$10 reduces visits by 4.43, resulting in 14.65 visitors per day and \$146.50 in daily revenue. Table 5 shows that, over a typical 128-day season, the differences in daily revenue add up significantly, providing valuable insight into how park management might optimize the fee structure to balance between maximizing visitor numbers and maximizing revenue.

Table 5. Predicted Visitors and Revenue for Various Parking Fees Over a 128-Day Season.

Fee	Predicted # Visitors		Predicted Revenue	
	Per Day	Per Season (128 days)	Per Day	Per Season (128 days)
\$0	19.08	2,442	\$0.00	\$0.00
\$8	15.54	1,989	\$124.32	\$15,913
\$10	14.65	1,875	\$146.50	\$18,752

Predicted visitors and revenue are derived from Model 2, which estimates the impact of parking fee amounts on daily visits. Calculations assume a 128-day season, with fees adjusted to December 2023 dollars.

The analysis shows a positive and statistically significant relationship between the number of daily visits to disc golf courses across New York, New Jersey, and Connecticut (excluding Heckscher) and daily visits to Heckscher. Specifically, the coefficient estimate of 0.018 indicates that for every 100 additional visits to other courses in the region, Heckscher gets an increase of 1.8 visits. The fact that the relationship between regional visits and Heckscher visits is positive and significant at the *** $p<0.01$ level indicates that Heckscher's visits are closely linked to regional disc golf activity. Thus, the number of regional disc golf visits serves as a highly effective explanatory control for Heckscher visits. The link between Heckscher visits and regional visits could be due to broader shared trends like favorable weather, increased awareness or usage of the UDisc app, or a general rise in the popularity of disc golf.

The weather variables, particularly temperature and its squared term, are slightly more complicated with how they affect visits to Heckscher. The squared temperature term is negative and significant, as expected, but the linear temperature term is negative and not statistically significant at any conventional level. The results indicate that overall temperature may not have a straightforward impact, but there is a non-linear relationship, where visits drop off at higher temperatures. The result holds across both models, and while the lack of significance for the linear temperature effect is somewhat unexpected, we hypothesize that the weather impact may be indirectly captured through daily regional visits. For instance, if it is a pleasant 70-degree day on Long Island, it is likely a similar 70-degree day across the Long Island Sound in Connecticut, and the broader regional weather trend is reflected in the regional visitation data. The correlation between local and regional conditions may explain why the temperature effect is not significant in the regression analysis but still affects visits through its effect on regional disc golf activity. Additionally, in Appendix A, where we run the analysis using the full nine years of data, the temperature terms show the expected results: the linear term is positive, the squared term is negative, and both are significant.

The effect of precipitation is consistently negative across both models, which is what we expect, because rain generally reduces participation in outdoor activities, including disc golf. However, the effect of lagged precipitation—reflecting the impact of rainfall from the previous day—while negative, was not significant. This result may not be as surprising as it initially seems. Unlike traditional golf, where wet ground conditions can hinder play, disc golf primarily involves walking and throwing, making it less dependent on ground conditions. Additionally, many disc golf courses tend to be located in parks or areas prone to flooding that are not suitable for other types of infrastructure and development. Thus, playing in muddy or damp conditions is often expected and does not seem to deter players, showing that players are willing to play in less-than-ideal conditions.

5.2 GOODNESS OF FIT AND MODEL SELECTION

The adjusted R-squared values from our econometric analysis indicate that both models have relatively strong explanatory power. Model 1 has an adjusted R-squared of 0.621, indicating that about 62.1% of the variability in daily visits to Heckscher State Park Disc Golf Course is explained by the variables in the model. Model 2 has a similarly close fit with an adjusted R-squared of 0.619.

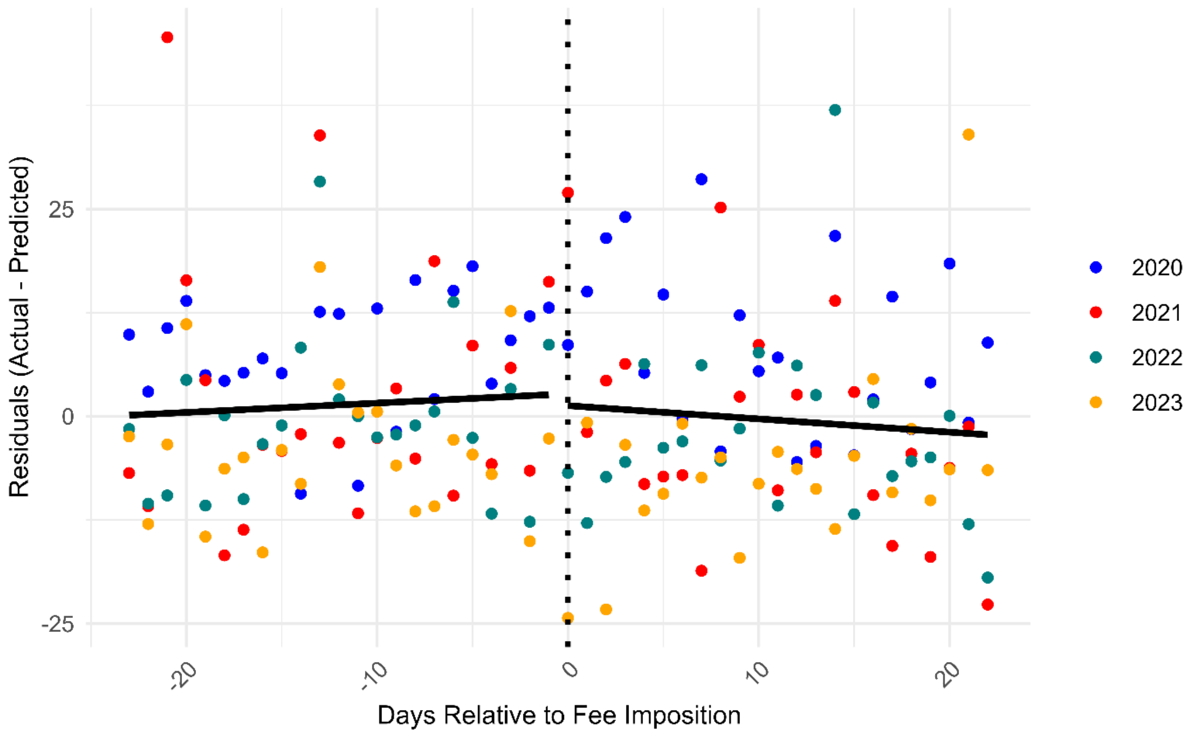
The Akaike Information Criterion (AIC) offers additional insights for model selection. Model 1 has a slightly lower AIC of 11032.3 compared to 11037.1 for Model 2, indicating a marginally better fit to the data, even with the simplicity of using fee presence as a binary variable. While the difference is subtle, it leans in favor of Model 1, particularly when considered in combination with the slightly higher R-squared value.

In terms of predictive accuracy, the Root Mean Square Error (RMSE) provides a clear measure of each model's performance in forecasting daily visits. Model 1 has an RMSE of 10.43, and Model 2's RMSE is 10.44, meaning both models are typically around 10 visits off from the actual observed numbers. Given that the average daily visitation is roughly 19, this might seem like a significant error at first glance. However, when considering the high variability in the data and the presence of potential outliers, these RMSE values reflect a fairly reasonable level of predictive accuracy.

Overall, while both models offer valuable insights into the factors influencing visitation at Heckscher State Park, the slightly stronger goodness of fit of Model 1 make it the preferred choice. Model 1's simplicity—focusing on the binary presence of a fee—provides an intuitive and straightforward interpretation of how fees impact visitation. The analysis suggests that disc golfers' decisions are more strongly affected by the existence of a fee rather than its specific amount, supporting the idea that the simple decision of whether to pay a fee is a more powerful deterrent than the exact cost. Additionally, as a robustness check, we replicated the regression analysis using a comprehensive dataset spanning nine years. The extended analysis, which is detailed in the appendix, reinforces the primary findings, with all significant variables maintaining their direction and significance levels, although the magnitude of the effect on visits varied slightly. Even in the longer analysis, the binary fee model continues to show a slight advantage over the fee amount model.

To show the impact of the parking fee on disc golf course visits at Heckscher State Park, we present a residual plot in Figure 9. The plot shows the residuals, calculated as the difference between the actual number of visits and the predicted number of visits, with predictions controlling for the day of the week, weather conditions, and other relevant factors, but assuming no fee. The data covers the three weeks before and after the start of the seasonal parking fee for each year from 2020 to 2023. The colored dots represent the daily residuals for each year, and the bold black lines are trend lines before and after the fee is imposed. The plot shows a noticeable drop in residuals following the fee's introduction, indicating its negative effect on visits. The analysis isolates the true impact of the fee, independent of other control variables. For a more detailed examination, Appendix B includes a comprehensive set of plots explaining the timeline and construction of this residual graph.

Figure 9. Residuals of Visits Before and After Parking Fee Imposition at Heckscher State Park



The plot shows the residuals, calculated as the actual number of visits minus the predicted number of visits (assuming no fee and controlling for factors such as day of the week and weather), for three weeks before and after the start of the seasonal parking fee for each year from 2020 to 2023.

5.3 EXTRAPOLATION AND BROADER IMPLICATIONS

The main findings from our econometric analysis—particularly from Model 1—provide valuable insights into how parking fees affect disc golf course attendance. Our results show that the parking fee at Heckscher State Park causes a measurable decrease in daily visits, with an estimated reduction of about 4.756 visits on days when the fee is imposed.

With an average daily attendance of 19.077 across the 1,461 days in our sample, and an average of 23.028 visits on fee days, the reduction from the parking fee equates to a relative decrease of around 20% to 25%, calculated as 4.756/23.028 = 20.65% and 4.756/19.077 = 24.93%, depending on which average is the reference point. The decrease is substantial, indicating the fee's deterrent effect on potential park visitors.

Extrapolating from our results, one might expect that if another disc golf course implemented a $10 fee, its visits could drop by around 20% to 25%, meaning 75% to 80% of players would still be willing to pay the fee to play. However, directly applying the estimate to other courses could be misleading, as Heckscher's disc golf course has several unique characteristics that may not apply elsewhere.

One challenge with extrapolating results from a single course to the broader disc golf population is that each course has its own unique features, such as differences in hole layout, difficulty, available amenities, and the demographics of its visitors. As a result, while the parking fee effect observed at Heckscher provides useful insights, applying these findings broadly without accounting for local conditions and the specific context of other courses could lead to inaccurate predictions about the impact of similar fees.

Although our results show that parking fees significantly affect attendance at Heckscher State Park, these findings may not be directly applicable to courses with capacity constraints, where the response to fee changes might differ. Finally, Heckscher's situation is unique for two other reasons: (1) the availability of the Empire Pass, which alters the relationship between parking fees and visitation, and (2) its geographical location as the only disc golf course on a densely populated island.

Empire Pass Consideration: The deterrent effect of the parking fee is lessened by the widespread use of the Empire Pass among disc golfers at Heckscher. Our intercept survey revealed that a significant portion of players are not affected by the daily fee either because they hold an Empire Pass or are unaware of the fee, with some in the latter group being pass holders. As a result, the observed drop in visits primarily affects those without an Empire Pass, representing a smaller segment of the overall disc golfing population.

If all the disc golfers at Heckscher were subject to the parking fee—without the shield provided by the Empire Pass—the deterrent effect would likely be stronger. Access options, like annual passes, affect the real impact of parking fees on attendance. The consideration is especially relevant for policy-makers thinking about introducing a daily parking fee without offering an annual pass or alternative payment options, where the effect of the fee could be more substantial than what is observed here. However, the fact that many Empire Pass holders are willing to pay $75 for the annual pass may also suggest a higher tolerance for costs, potentially softening the fee's overall impact.

Geographical Location Consideration: Heckscher State Park's position as the only disc golf course on Long Island significantly affects how broadly our results can be applied. The exclusivity means that the decision to play at Heckscher often reflects a wider choice to engage in disc golf at all, given the lack of nearby alternatives. The presence of a parking fee at Heckscher does not only test demand for the course but also measures the fee's impact as a potential obstacle to participating in the sport within the region.

The availability of alternative courses in other regions could result in different behavioral responses, particularly through substitution effects that may amplify the impact of any fees. In areas where disc golfers have several nearby courses, imposing a parking fee might not only reduce visits by the observed 20-25% but could also divert some players to other free-access courses. Therefore, while our study finds a significant fee impact, the actual reduction in visits could be even greater in regions with more options for disc golf courses.

The distinction emphasizes the importance of placing our findings within Heckscher's unique market context. The observed 20-25% reduction in visits likely reflects a “best-case scenario” for imposing fees, where there are no nearby alternatives or where all competing courses in the area implement similar fees. Parks and disc golf courses considering similar fee strategies need to account for local competition and the potential for course substitution, which could significantly amplify the fee's deterrent effect beyond what we observed at Heckscher.

Extrapolation to Pay-to-Play Courses: When extending our findings, it is important to distinguish between Heckscher's parking fee and the broader concept of pay-to-play disc golf courses. Many premium pay-to-play courses justify their fees by offering superior amenities, course design, and maintenance—enhanced features that Heckscher, as a typical public park course, does not provide. Our study focuses specifically on the impact of a parking fee in a public park setting, without any corresponding improvements to course quality or amenities. As such, our results primarily capture the change in demand as it responds to an exogenous cost increase.

Extrapolating our findings to pay-to-play courses, where fees directly fund enhancements that may attract more players, requires careful consideration. Our study offers a baseline understanding of the effects of fees in a situation where the fees are sometimes charged when disc golfers enter the larger state park, but are somewhat separated from the disc golf course itself. The implications for courses implementing fees to improve quality could differ significantly. Park administrators and course designers must weigh the potential reduction in visits against the benefits of fee-supported improvements, balancing financial sustainability with maintaining and growing player engagement.

5.4 IMPLICATIONS FOR PARK MANAGEMENT AND POLICY

Our analysis emphasizes the significant effect that parking fees can have on recreational activity at public parks. Park administrators and policymakers should carefully weigh the benefits of implementing fees against the potential to deter visitors. Based on our study, we recommend that policymakers and park administrators consider the following:

- Offering mechanisms like the Empire Pass could help maintain or even boost park attendance, by mitigating the negative impact of parking fees and fostering visitor loyalty, which may encourage repeat visits.
- The potential for a substantial reduction in visits, absent such mitigating mechanisms, suggests that alternative revenue strategies or visitor management approaches might be necessary to balance financial objectives with the goal of promoting public engagement in recreational activities.
- The specific findings from Heckscher State Park provide a valuable case study on the effect of parking fees, but their broader applicability requires careful consideration of local market conditions, player preferences, and the unique value proposition of each disc golf course.

The parking fee at Heckscher State Park reduces visits by an estimated 20-25%, but the effect occurs in a context where there are no nearby courses and a significant portion of visitors can ignore the daily fee because they bought an annual pass. The broader implication for parks without similar mitigating mechanisms could be a much larger reduction in visitation. Since we cannot compare our setting, which includes the Empire Pass, to one where such a mechanism does not exist, we recommend exercising caution when applying these results to other courses. Park administrators should carefully consider the full range of potential impacts and explore strategies that balance financial sustainability with maintaining accessibility and visitor engagement.

As we build our understanding of the economics underlying disc golf, park visitation, and recreational participation, further research is needed to examine the elasticity of demand across different contexts. Studies that incorporate diverse park settings, fee structures, and visitor demographics will help broaden the discussion on sustainable park management and policy development. In addition, longitudinal analyses capturing the long-term effects of fee implementations and changes in visitor behavior can provide valuable insights into the evolving relationship between economic mechanisms and recreational engagement.

6. CONCLUSION

This study examined the impact of parking fees on disc golf course attendance at Heckscher State Park and found that the fees caused a reduction in daily visits by approximately 4.76, which translates to a 20-25% decrease in overall attendance. By controlling for factors such as weather, precipitation, and regional disc golf activity, the analysis isolates the specific effect of parking fees, ensuring that the results are not conflated with seasonal changes or broader market trends.

Our analysis at Heckscher State Park, where there are no nearby alternative courses, captures the demand for disc golf itself as a recreational activity rather than simply a preference for a particular course. The unique context mirrors a situation where all competing courses might impose similar fees, allowing us to more clearly isolate the impact of fees on overall participation in the sport. We also considered the implications of the Empire Pass—an annual pass offering unlimited access to most New York State parks, including Heckscher, for a flat fee of $75. Many frequent disc golfers purchase the pass, which partially offsets the deterrent effect of daily parking fees on visitation.

Heckscher's unique position as the only disc golf course on Long Island is a reminder of the importance of considering local market conditions when making fee-related decisions. The 20-25% decrease in visits observed with the fee imposition reflects a scenario where no close alternatives are available, potentially depicting the lower bound of a fee’s impact in more competitive areas. The findings emphasize the need for park administrators to factor in the broader recreational landscape and the availability of substitutes when developing pricing strategies to avoid significant unintended reductions in visits.

In summary, our research contributes to understanding the relationship between fee structures and park visitation. We find that fees are a deterrent to park visits, but the magnitude of the impact is dependent on local conditions and policy instruments. While not a definitive guide for park administration, this study provides valuable insights into how fee policies affect visitation and suggests that park management decisions should be adapted to specific local conditions.

The findings also point to opportunities for developing new disc golf courses in urban areas like New York City or additional locations on Long Island, where the sport's growing popularity could support such expansions. The insights from this study can help guide the evaluation of benefits and policy considerations for these developments, offering direction for future research. Moving forward, further studies should broaden the scope to include various parks and explore how different amenities and fee structures affect recreational participation. Longitudinal and comparative analyses could provide a deeper understanding of evolving recreational trends and the value that park amenities provide, supporting more informed, evidence-based policymaking in public recreation management.

ENDNOTES

1. We recognize that during the period of COVID-19, other options for leisure activities (gyms, bars, etc.) were limited, which could mute disc golfer response to entrance fees depending on the extent that these activities were substitutes. A more nuanced analysis of the effect of COVID-19 is beyond the scope of this paper and is left to future research.
2. We conducted the intercept survey at Heckscher State Park as part of a separate study focused on contingent valuation and travel cost models, which also includes surveys at other disc golf courses. However, for this paper, we use only responses from Heckscher, selectively drawing on specific questions relevant to modeling visitation dynamics. For a full account of the survey methodology, including the questionnaire, response rates, and findings, see Meyer et al. (2024).

REFERENCES

Banzhaf, H. S., & Smith, V. K. (2020). *Financing outdoor recreation* (No. w27541). National Bureau of Economic Research.

Bureau of Labor Statistics. (2024). *Consumer Price Index - New York-Newark-Jersey City, NY-NJ-PA*. U.S. Bureau of Labor Statistics. https://www.bls.gov/regions/northeast/data/xg-tables/ro2xgcpiny.htm

Brown, L. E., Basak-Smith, M., Bradley, K., Stearns, S. F., Morzillo, A. T., & Park, S. (2021). Exploring the implications of increased rural trail use during the COVID-19 pandemic on health, planning, equity, and inclusivity. *Choices, 36*(3), 1-7.

Buckley, R. (2003). Pay to play in parks: An Australian policy perspective on visitor fees in public protected areas. *Journal of Sustainable Tourism, 11*(1), 56-73.

Cohen, D. A., McKenzie, T. L., Sehgal, A., Williamson, S., Golinelli, D., & Lurie, N. (2007). Contribution of public parks to physical activity. *American Journal of Public Health, 97*(3), 509-514.

Ellis, D., & Schwartz, R. (2016). The roles of an urban parks system. *World Urban Parks, 10*.

González-Ramírez, J., Liao, K. A., Meyer, K., & Adams, V. (2023). The Effect of Seasonal Parking Fees on Disc Golfing. *Available at SSRN 4636565.*

Hanauer, M. M., & Reid, J. (2017). Valuing urban open space using the travel-cost method and the implications of measurement error. *Journal of Environmental Management, 198*, 50-65.

Hesseln, H., Loomis, J. B., González-Cabán, A., & Alexander, S. (2003). Wildfire effects on hiking and biking demand in New Mexico: a travel cost study. *Journal of Environmental Management, 69*(4), 359-368.

Larson, L. R., Jennings, V., & Cloutier, S. A. (2016). Public parks and wellbeing in urban areas of the United States. *PLoS one, 11*(4), e0153211.

Latinopoulos, D. (2014). The impact of economic recession on outdoor recreation demand: an application of the travel cost method in Greece. *Journal of Environmental Planning and Management, 57*(2), 254-272.

Lewison, M. (2017). Demand-based pricing in the US theme park industry. *International Journal of Leisure and Tourism Marketing, 5*(3), 271-293.

Mayer, M., Müller, M., Woltering, M., Arnegger, J., & Job, H. (2010). The economic impact of tourism in six German national parks. *Landscape and Urban Planning, 97*(2), 73-82.

Meyer, K., González-Ramírez, J., & Liao, K. A. (2024). Estimating the value of a disc golf course with revealed and stated preferences. *Working Paper.*

Middle, I., Hedgcock, D., Jones, R., & Tye, M. (2017). Understanding and planning for organized community sport in public parks: A case study of policy and practice in Perth. *Urban Policy and Research, 35*(4), 443-458.

Mukanjari, S., Muchapondwa, E., & Demeke, E. (2021). Recreation demand and pricing policy for international tourists in developing countries: evidence from South Africa. *Journal of Environmental Economics and Policy, 10*(3), 243-260.

Mulwa, R., Kabubo-Mariara, J., & Nyangena, W. (2018). Recreational value and optimal pricing of national parks: lessons from Maasai Mara in Kenya. *Journal of Environmental Economics and Policy, 7*(2), 204-222.

Orlowski, J., & Wicker, P. (2019). Monetary valuation of non-market goods and services: A review of conceptual approaches and empirical applications in sports. *European Sport Management Quarterly, 19*(4), 456-480.

Reynisdottir, M., Song, H., & Agrusa, J. (2008). Willingness to pay entrance fees to natural attractions: An Icelandic case study. *Tourism Management, 29*(6), 1076-1083.

Sibly, H. (2001). Pricing and management of recreational activities which use natural resources. *Environmental and Resource Economics, 18*, 339-354.

Smith, A. [Aaron] (2024). *Download US weather data: PRISM weather data by zipcode, county, or state.* https://asmith.ucdavis.edu/data/prism-weather

Smith, A. [Andrew]. (2018). Paying for parks. Ticketed events and the commercialisation of public space. *Leisure Studies, 37*(5), 533-546.

UDisc. (2023). *Disc golf growth report powered by UDisc*. https://udisc.com/disc-golf-growth-report

APPENDIX A: EXTENDED ANALYSIS USING FULL NINE-YEAR DATASET

This appendix extends the scope of our analysis by incorporating data from January 1, 2015, to December 31, 2023, encompassing a total of 3,287 observations. While the main paper focused on the most recent four years, the longer time frame offers a more comprehensive view of visitation patterns at Heckscher State Park. The nine-year period allows us to capture the early stages of disc golf's growing popularity and the increasing adoption of the UDisc app, providing valuable context for our findings.

The extended dataset reinforces the core results presented in the main analysis, with most variables retaining their significance and directionality. However, as expected, the magnitude of some coefficients differs. For instance, the parking fee's impact on daily visits shows a reduction of approximately 1.93 visits per day, compared to the 4.76 visits seen in the four-year analysis. This smaller reduction reflects the lower average number of visits in the earlier years, where the nine-year average is 11.37 visits per day. In relative terms, the parking fee corresponds to a 17% decrease in visits over the full period, compared to a 20-25% decrease observed in recent years.

By examining the full dataset, we confirm the robustness of our main findings while also offering insights into visitation trends during the sport's early growth phase. Table A.1 contains summary statistics for the variables over the full nine years of data, offering a detailed view of the range and variability in daily visits, regional disc golf activity, parking fees, and weather conditions.

Table A.1 Summary Statistics – Full Nine-Year Dataset

Variable Description	Min	Median	Mean	SD	Max
Daily Visits to Heckscher	0	6	11.37	14.10	101
Daily Visits to All Other Courses (NY-NJ-CT)	0	231	593.21	720.40	3628
Parking Fee (1 = Yes, 0 = No)	0	0	0.34	0.48	1
Parking Fee Amount ($)	0	0	3.17	4.41	10
Parking Fee Amount in Dec 2023 $	0	0	3.64	5.08	12.43
Daily High Temperature (°C)	-10.18	16.58	16.25	9.25	35.99
Daily Precipitation (cm)	0	0	0.33	0.82	11.30

Summary statistics are based on *N = 3287* observation days from Jan 1, 2015, through Dec 31, 2023.

The summary statistics from Table A.1 provide a snapshot of the full nine-year dataset, capturing the evolving trends in disc golf participation at Heckscher State Park. With an average of 11.37 visits per day over this extended period, it is clear that overall visitation has increased in recent years, aligning with the sport's growing popularity and wider use of tracking apps like UDisc. Additionally, the dataset reveals broader fluctuations in regional visits and weather, both of which affect visitation trends at Heckscher.

To further investigate how these variables interact and to measure the effect of the parking fee on daily visits, we extend our econometric analysis to cover the full nine-year period. The expanded timeframe allows us to assess whether the relationships observed in the main analysis hold over a longer period with different visitation dynamics, and whether the impact of the parking fee remains consistent. In particular, we examine whether factors such as regional visits, weather conditions, and the parking fee itself have similar or varying effects compared to the more recent four-year period.

Table A.2 presents the results of our regression analysis using the full dataset. As with the primary analysis, we estimate two models: Model 1 assesses the binary impact of parking fee presence (whether a fee is imposed or not), while Model 2 examines the effect of the parking fee amount in real December 2023 dollars. Both models control for daily regional visits, weather variables, and fixed effects for day of the week, holidays, and year. The coefficients, standard errors, and statistical significance levels are provided for each variable.

Table A.2 Econometric Analysis Results – Full Nine-Year Dataset

Variable Description	Model 1 (Fee Presence)	Model 2 (Fee Amount)
Fee Presence (1=Yes, 0=No)	–1.928*** (0.501)	
Fee Amount (Dec '23 $)		–0.124*** (0.045)
Daily Regional Visits (NY-NJ-CT)	0.015*** (0.001)	0.015*** (0.001)
Daily High Temperature (°C)	0.145*** (0.052)	0.140*** (0.052)
Daily High Temperature Squared	–0.010*** (0.002)	–0.010*** (0.002)
Daily Precipitation (cm)	–0.956*** (0.190)	–0.966*** (0.191)
One-Day Lagged Precipitation	–0.180 (0.168)	–0.186 (0.169)
Daily Time Trend	0.001 (0.002)	0.001 (0.002)
Day of Week Fixed Effects	✓	✓
Holiday Fixed Effects	✓	✓
Year Fixed Effects	✓	✓
Observations	3286	3286
Adjusted R-squared	0.665	0.664
Akaike Information Criterion (AIC)	23153.0	23162.4
Root Mean Square Error (RMSE)	8.14	8.15

The dependent variable is Daily Visits to Heckscher. Standard errors are heteroskedasticity-robust and are reported in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The extended econometric analysis covering the full nine-year dataset reinforces the core findings from the main analysis. As shown in Table A.2, the parking fee continues to significantly reduce daily visits to Heckscher State Park. Specifically, the fee results in a reduction of approximately 1.93 visits per day, which, while smaller than the 4.76 reduction found in the four-year dataset, still reflects a substantial deterrent effect. This smaller reduction is consistent with the lower average number of visits in the earlier years, and it shows how the park's visitation has evolved over time as disc golf has gained popularity.

Similarly, in Model 2, each additional dollar of the parking fee leads to a reduction of 0.124 visits per day, which, although smaller than the 0.443 reduction in the recent four-year analysis, confirms that visitors remain sensitive to changes in the fee amount. The relationship between regional visits and Heckscher visits remains positive and significant, with every 100 additional visits to regional courses associated with a 1.5 visit increase at Heckscher.

The weather variables, including temperature and precipitation, retain their expected effects. The positive linear temperature coefficient and the negative quadratic term continue to validate the inverted U-shaped relationship between temperature and visits, while precipitation consistently deters attendance, as expected.

Overall, the extended analysis provides further evidence for the robustness of our findings. Despite some variations in magnitude, the significance and direction of the main variables remain consistent, underscoring the reliability of the results. The additional insights into long-term trends provide a broader understanding of how factors like parking fees, regional activity, and weather affect disc golf visits at Heckscher State Park.

APPENDIX B: ANALYSIS OF PARKING FEE IMPOSITION ON DAILY VISITS

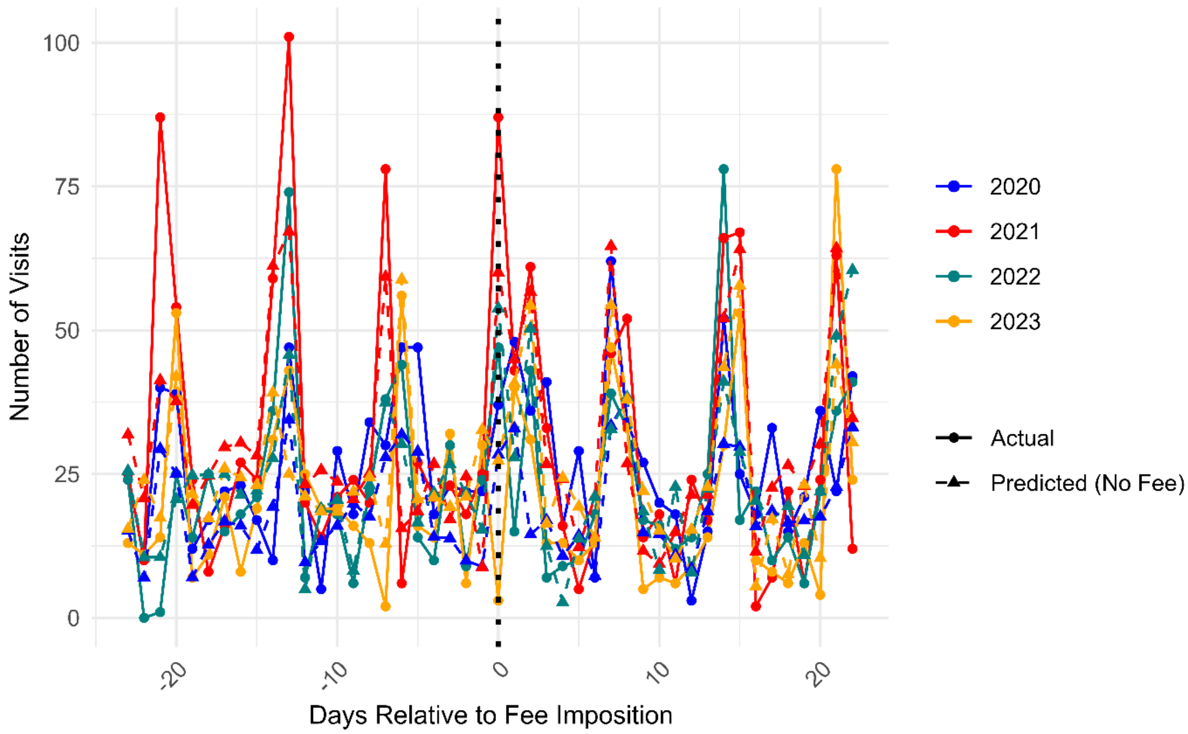
In this appendix, we present additional analysis examining the effect of the seasonal parking fee on daily visits to Heckscher State Park. While our main analysis includes data for the full calendar year, this appendix focuses specifically on the three weeks before and after the start of the seasonal parking fee for each year in our dataset (2020-2023). By narrowing the timeframe, we gain a clearer view of how the imposition of the parking fee directly affects daily attendance.

Figure B.1 shows the actual number of daily visits alongside the predicted number of visits, assuming no fee was applied. The predictions control for factors such as day of the week, weather conditions, and regional disc golf activity. The comparison allows us to see how the parking fee affects attendance while accounting for other variables that also impact visitation.

Although the graph shows the actual and predicted visits for the three weeks before and after the parking fee is introduced, the visual complexity—due to the multiple lines representing different years—makes it challenging to see clear differences in how the predictions align before versus after the fee imposition. While the predicted visits generally resemble the actual visits, particularly during high-demand periods like weekends, it is harder to pinpoint the fee's effect directly from this graph alone.

For this reason, the main text focuses on the residuals, which more clearly isolate the parking fee's impact by showing the difference between actual and predicted visits. The residual analysis, presented with a trend line in the main text, offers a cleaner visualization of how the fee reduces attendance, independent of other variables like weather or the day of the week. Together, the two visualizations complement each other, providing a comprehensive view of the fee's effect on visitation at Heckscher State Park.

Figure B.1 Actual and Predicted Visits Before and After Parking Fee Imposition at Heckscher State Park



The plot shows the actual number of visits and the predicted number of visits (assuming no fee, controlling for day of the week, weather, etc.) for three weeks before and after the fee imposition (*i.e.*, start of seasonal parking fee) for each year (2020-2023)

The Impact of CO_2 Emissions on Life Expectancy in Jamaica: A Cointegration and Quantile Regression Analysis

Adian McFarlane*, Leanora Brown†, Anupam Das‡

ABSTRACT

We examine the impact of carbon dioxide (CO_2) emissions on life expectancy in Jamaica from 1976 to 2019. We control for gross domestic product per capita, infant mortality and remittance inflows. Using the autoregressive bounds testing approach to cointegration complemented with quantile regression analysis, we find that a 1% increase in CO2 emissions is associated with a 0.07% decrease in life expectancy in the long run. The cointegrating relationship is stable across five quantiles, suggesting that the findings are reliable and statistically robust. We argue that policymakers should encourage sustainable energy options and discourage the consumption of fossil fuels.

INTRODUCTION

This study examines the relationship between carbon dioxide (CO_2) emissions and life expectancy in Jamaica. While there is some consensus on the deleterious impact of CO_2 emissions on life expectancy, the magnitude of this impact is very much country-specific, so we carry out a single-country study. In addition, with a single-country study, the findings provide a basis for providing insights into local policies and remedial environmental actions that could be applied given the peculiar national context of Jamaica. Such context relates to the country's life expectancy trajectory, level and sources of CO_2, emissions, consumption and production patterns, and industrial structure. Therefore, we empirically shed light on this impact for Jamaica and discuss the attendant policy implications of our findings.

Life expectancy, a frequently used human development indicator, summarizes mortality in the population and represents the average number of years a person is expected to live, typically at the time of birth. Although global life expectancy trends upward, there is a significant variation across regions and countries. Globally, life expectancy increased from 66.8 years in 2000 to 73.3 years in 2019. On a regional basis, life expectancy in Europe was 78.2 years in 2019, the highest across all regions; 77.7 years in the Western Pacific; 77.2 years in the Americas; and 64.5 years in Africa – the lowest life expectancy (Pan American Health Organization, n.d.). Life expectancy in Jamaica mirrors global trends, with an increase from 68.5 years in 1976 to 71.8 years in 2019 for both sexes but lags the regional average except for Africa.

The Ministry of Health and Wellness in Jamaica, in support of its global commitment to the sustainable development goals of the United Nations, launched its Vision for Health Plan to improve the health and

* King's University College at Western University Canada
† University of Tennessee at Chattanooga
‡ Mount Royal University

longevity and the social and environmental conditions of its population by 2030. To achieve this, the Vision for Health Plan aims to create conditions in which productive enterprises further greater wealth (The Ministry of Health and Wellness, n.d.). However, this objective is incongruent with the fact that Jamaica's productive enterprises are driven chiefly by fossil fuel imports, and many of these enterprises are generally net emitters of CO_2 emissions. It is widely acknowledged that burning fossil fuels releases large amounts of CO_2 in the air, thus making the country and, more broadly, the world susceptible to climate change hazards. According to the United Nations (2023), climate change has caused the premature deaths of 2 million people between 1970 and 2021 and thus reduces life expectancy.

Jamaica's annual CO_2 emissions have grown steadily since its independence in 1962. The country recorded CO_2 emissions of 2.12 million tonnes in 1962, which ballooned to 11.58 million tonnes in 2006 before falling back to 7.82 million tonnes in 2019 (Our World in Data, 2023). However, while the country's contribution to global CO_2 emissions is consistently low and was a mere 0.02% in 2019, the annual growth in emissions has triggered concerns because of its potential impact on public health outcomes, including life expectancy (Our World in Data, 2023). This concern is partly fuelled by the country's heavy dependence on fuel oil imports for energy generation. More than 85% of the country's CO_2 emissions are derived from fuel oil usage, which overexposes the country to climate risks and natural disasters (Our World in Data, 2023; Feruglio et al., 2023). Theoretically, it is argued that increasing amounts of CO_2 emissions can be harmful to the health and well-being of a population because it can cause air pollution that could subsequently give rise to illnesses such as respiratory diseases. Further, it is argued that rising levels of CO_2 emissions have led to rising average temperatures and, consequently, increased the frequency of prolonged high heat and droughts (Feruglio et al., 2023). As a result, rising levels of CO_2 emissions have been linked to respiratory diseases and food production and water quality disparities, especially among infant and elderly populations as well as vulnerable people from lower socioeconomic backgrounds (Rahman et al., 2022), all of which could potentially reduce life expectancy.

The contribution of our study to the literature is centred on filling the gap that is the absence of any study on Jamaica on the causal nexus between CO_2 emissions and life expectancy. With rising per capita CO_2 emissions, Jamaica is highly vulnerable to climate change hazards. As a result, there is some urgency for the country to move away from carbon-emitting production processes and towards renewable energy production. Concomitantly, economic growth in this country is primarily driven by non-renewable energy consumption. Thus, our study sheds light on the extent to which life expectancy considerations should form part of a sustainable economic development policy mix.

To elucidate the causal nexus between CO_2 emissions and life expectancy, we apply the autoregressive distributed lag (ARDL) bounds testing approach to cointegration to annual data over the 1976 to 2019 period. We use standard checks for model adequacy. Further, we verify the robustness of our model by estimating a quantile ARDL (QARDL) regression of the preferred ARDL model. The application of the QARDL regression is particularly noteworthy as it allows us to have a more nuanced understanding of how this causal relationship changes across the distribution of these variables. For our analysis, we control for national real value added, infant mortality, and remittance inflows (remittances). These inflows have been found by several scholars to significantly impact several aspects of the socioeconomic life of this small island developing state and so are of particular importance to be included as a control (e.g., Campell et al., 2022; Das, Brown, & McFarlane, 2023; McFarlane, Brown, & Das, 2023). Our analysis shows evidence of a statistically significant negative relationship between CO_2 emissions and life expectancy in the long run. We find that for every 1% increase in CO_2 emissions, there is a corresponding decrease of 0.07% in life expectancy, an inelastic response but a highly statistically significant one. One implication of our finding is that policymakers should formulate sustainable development policies accounting for the CO_2 emissions-life expectancy nexus as they strive to promote the use of renewable energy sources, such as wind and solar power, and reduce the use of fossil fuels. The rest of the study proceeds as follows. The next section provides a review of both theoretical and empirical literature. This is followed by an overview of CO_2 emissions in Jamaica and our contribution to the literature. The following section presents the data and methods. We discuss the findings in the next section. The last section concludes with a summary and the attendant policy recommendations.

LITERATURE REVIEW

The far-reaching effects of CO_2 pollution on human health and life expectancy are significant and complex. In its 2022 factsheet, the World Health Organization (2022) argued that in 2019, 4.2 million premature deaths were attributable to long-term exposure to particulate matter, which caused cardiovascular and respiratory disease and cancer. These conditions reduced life expectancy, particularly in urban areas, which tend to have high pollution levels. The theoretical reasons for the impact of air pollution on life expectancy include the formation of harmful secondary pollutants, extreme weather conditions that cause deaths, and health effects from pollution. The level of CO_2 that is found in the atmosphere is not directly harmful to human health. However, when CO_2 is released into the atmosphere from fossil fuel combustion, it can contribute to the formation of particulate matter, which increases the risk of death from cardiopulmonary causes (Jerett et al., 2009; Fiore et al., 2015). CO_2 emissions are responsible for extreme weather events and climate change. Higher levels of CO_2 in the atmosphere cause extreme heat waves, floods, and storms, which lead to injuries and deaths, particularly among vulnerable populations, and the spread of waterborne diseases (Haines et al., 2006). Further, Patz et al. (2005) found evidence of the impact of climate change on increased morbidity and mortality in environmentally vulnerable regions.

There are other indirect ways in which CO_2 emissions can impact life expectancy. Almost two billion people suffer from deficiencies of important minerals such as zinc and iron. These minerals in some crops are sensitive to the level of CO_2 in the atmosphere. Myers et al. (2014) showed that the elevated atmospheric CO_2 concentration is responsible for reducing zinc, iron and protein in some grains. As noted by the authors, dietary deficiencies of zinc and iron cause a loss of 63 million life-years annually. Therefore, elevated CO_2 concentrations in the atmosphere are partly responsible for the deterioration of global health.

Finally, the impact of climate change on life expectancy through mental health has not been given enough attention in the existing literature. Berry et al. (2010) explained three pathways through which climate change can affect mental health and life expectancy. First, frequent and intense natural disaster causes serious anxiety-related responses and chronic and severe mental health crises. Second, severe weather events can also cause long-term physical health problems that impact mental health. Third, mental health deterioration also occurs since climate change destroys natural and social environments that people rely on for their livelihoods and well-being. Williams et al. (2015) estimated the impact of climate change on the suicide rate. They concluded that in New Zealand, between 1998 and 2007, every 1-degree Celsius increase in temperature due to irregular variation in temperature is associated with 1.8% more suicides.

Several scholars have used different variables to investigate the dynamic link between CO_2 emissions and health outcomes with conflicting results. Some scholars have found positive and negative relationships between the variables, while others have found no association. Using the ARDL approach and time series data from 1995 to 2017 for Pakistan, Danish et al. (2019) found a positive relationship between CO_2 emissions and health expenditures. This result could suggest improved life expectancy, although this was not assessed directly. Similar results were found by scholars, including Beatty and Shimshack (2014), Yazdi, Tahmasedi, and Mastorakis (2014), Chaabouni, Zghidi, and Mbarek (2016), Yazdi and Khanalizadeh (2017), and Apergis et al. (2018).

Some scholars have also examined the relationship between energy consumption, the environment, and health outcomes and found this impact to be adverse. For instance, Wang et al. (2020), using time series data from 1972 to 2017, found that energy consumption reduces life expectancy in Pakistan. They noted that this finding may be indirect and operates through the impact of energy consumption on the environment. Others have established that higher energy consumption increases global energy-related CO_2 emissions that can lead to environmental degradation (Zakaria & Bibi, 2019; Danish et al., 2018; Bekhet et al., 2017), harming human health. Sarkodie et al. (2019) also examined the relationship between environmental degradation that results from energy consumption and European life expectancy. They showed that an increase in energy consumption by 1% increased per capita industrial particulate matter emissions by between 0.42% and 0.45%, lowering life expectancy and increasing mortality. In their research, Al Mulali and Ozturk (2015) and Zaidi and Saidi (2018) also found that emissions in the form of air pollution caused by energy consumption are harmful to human health.

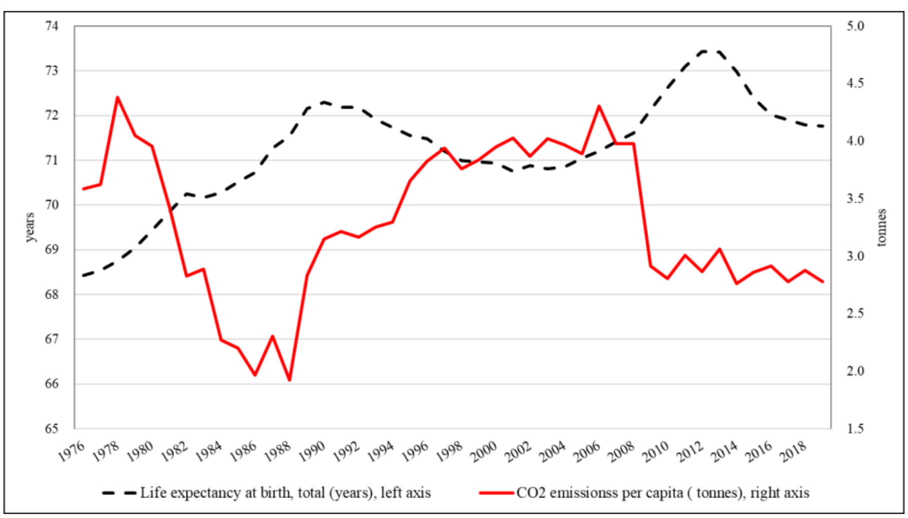
Other scholars have studied the direct relationship between CO_2 emissions per capita and life expectancy. For instance, Ali and Audi (2016) and Azam, Uddin, and Saqib (2023) conducted empirical investigations of this relationship with time series data. They utilized the ARDL approach for Pakistan and found that CO_2 emissions had a negative impact on life expectancy. In a related study by Azam and Adeleye (2023) for Asia and the Pacific region, they investigated the relationship between CO_2 emissions from liquid and solid fuels and environmental quality. They found that emissions from these sources create pollution, which lowers longevity and thus reduces life expectancy. Likewise, Nkalu and Edeme (2019) showed that environmental hazards regarding CO_2 emissions from solid fuel consumption reduced life expectancy by one month and three weeks in Nigeria. In a similar study on Nigeria, Matthew et al. (2018) utilized the ARDL econometric approach with data from 1985 to 2016 to show that a 1% increase in greenhouse gas emissions reduced life expectancy by 0.042%. Other studies that have also confirmed this negative relationship include Mahalik et al. (2022), Landrigan et al. (2018), Hill et al. (2019), Rahman, Rana, & Khanam (2022), Schwartz et al. (2018), and Demir et al. (2023).

CO_2 EMISSIONS IN JAMAICA

What do we know about CO_2 emissions in Jamaica?

From a global perspective, Jamaica's total CO_2 emissions are relatively small. The country's annual CO_2 emissions were only 7.7 million tonnes in 2021 compared to China (11.5 billion tonnes), the United States (6.3 billion tonnes), and India (2.7 billion tonnes) – major carbon-emitting countries – for the same period (Our World in Data, 2023). In tonnes per capita, CO_2 emissions were 2.3, 8.0, 14.9, and 1.89 for these countries in 2021, respectively (Our World in Data, 2023). Although Jamaica's CO_2 emissions fluctuated substantially in recent years, there has been steady growth in emissions between 1950 and 1978 and between 1988 and 2006. Much of this growth is caused by an over-reliance on fossil fuel imports for energy generation. With Jamaica's antiquated energy generation infrastructure, there is an overwhelming demand for fuel oil, a significant fossil fuel component. Likewise, the country's lack of a fully developed mass transportation infrastructure has led to high per capita fuel consumption due to the extensive use of cars and taxis (Feruglio et al., 2023). Fuel commodity imports accounted for 11% of the country's gross domestic product (GDP) for the fiscal year 2021/2022 (Our World in Data, 2023; World Bank, 2023).

Figure 1. Life Expectancy and CO_2 Emissions in Jamaica



Source: World Bank (2023)

With rising greenhouse gas emissions, this country has recently experienced sweltering temperatures and rising sea levels. For instance, the average annual temperature in Jamaica was projected to increase by 1.54 degrees Celsius by 2100 relative to the 1986-2005 baseline (Planning Institute of Jamaica, n.d.). According to the 2020 ND-GAIN Vulnerability Index (see Feruglio et al., 2023), the sea level is rising and threatening Jamaica's infrastructure and population, concentrated in coastal and low-lying areas. Thus, climate changes have predisposed the country to natural disasters, including more irregular rainfall that precipitates flooding or droughts and stronger tropical cyclones (Feruglio et al., 2023). Due to air pollution, climate change can also harm human health. Air pollution can induce respiratory illnesses, increase hospital and emergency room visits for the asthmatic, and increase premature deaths, ultimately reducing longevity and life expectancy. As Figure 1 suggests, longevity and life expectancy significantly declined between 1988 and 2001, which overlaps with the period of rising CO_2 emissions in Jamaica.

Situating our contribution to the literature

As noted above, several scholars have investigated the link between CO_2 emissions and health outcomes, but the results have been mixed. Some studies found a positive relationship between the variables, while others found a negative or no association. The studies that found a positive relationship between CO_2 emissions and health outcomes suggest that increased health expenditures may be associated with improved life expectancy. However, these studies did not directly use life expectancy. The studies that found a negative relationship between CO_2 emissions and health outcomes suggest that air pollution caused by energy consumption can lower life expectancy. The studies that found no association between CO_2 emissions and health outcomes suggest that other factors, such as socioeconomic status, could be more important determinants of health outcomes. Based on these mixed results, we focus on Jamaica to establish and quantify the impact of CO_2 emissions on health outcomes as reflected in life expectancy at birth. Thus, our contribution quantifies the causal nexus of the impact of CO_2 emissions on life expectancy in Jamaica. We do this in a particularly robust statistical framework by considering both the long-run and short-run dynamics in an ARDL bounds testing framework, which is widely used in the pollution literature (see, e.g., Alola & Adebayo, 2023; Shen et al., 2021; Adedoyin & Zakari, 2020) In addition, we apply the QARDL model to this framework to assess whether the causal nexus between both variables is stable or changes at different threshold levels as reflected in five quantile limits. Although the QARDL technique has gained more attention in recent years (Sharif et al., 2020; Wang et al., 2022; Kadir et al., 2023), to our knowledge, this is the first study that applies this method to examine the relationship between CO_2 emissions and health outcomes for Jamaica.

EMPIRICAL STRATEGY

Choice of control variables

A remark about the choice of control variables is in order: GDP, remittances, and infant mortality. There are many social and economic determinants of health and, by extension, life expectancy. These include income and social status, employment and working conditions, education and literacy, childhood experiences, physical environments, social supports and coping skills, healthy behaviours, access to health services, biology and genetic endowment, gender, culture, and race (Public Health Agency of Canada, 2023). In choosing variables for this study, we are guided by these determinations balanced against data availability for the period the analysis is conducted. We include infant mortality as it is known to provide a good indicator of a society's overall underlying health and social conditions. This measure partially reflects accessibility to quality health services, particularly good prenatal and postnatal childcare, and broader social and economic conditions (Centers for Disease Control and Prevention, 2022).

GDP is a standard measure included in the analysis of health outcomes, with higher levels generally associated with higher life expectancy (Deaton, 2015; Preston, 1975). Higher GDP gives a country better healthcare infrastructure and more effective public health interventions, improving general living conditions and health outcomes. Jamaica is a high-remittance-receiving country, and remittances have been found to impact GDP and other socioeconomic and environmental outcomes (Brown et al., 2020; Das et al., 2019; Das, McFarlane, & Jung, 2023). We include the remittance flow to Jamaica as a control variable. While remittances are not the main variable of interest in our analysis, it is important to understand how remittances impact life expectancy. On the one hand, remittances generally positively impact life expectancy. These flows increase recipient households' income, allowing them to increase expenditure on health-related nutrition and access to healthcare services (Amuedo-Dorantes & Pozo, 2009; Djeunankan & Tekam, 2022). In a 2009 study, Thomas-Hope et al. noted that medical expenses are the third highest component among the purposes for which Jamaican households used international remittances. Over 15% of total remittances were used to cover medical expenses. Beuermann et al. (2016) found that remittances offset the income loss caused by adverse health shocks in Jamaica. Thus, in the absence of health insurance, remittances are used as insurance against health shocks. All these have positive impacts on life expectancy. On the negative side, in a recent study, Campbell et al. (2024) found that there is a long-run association between remittances and homicides in Jamaica. This is mainly due to conflicts among criminals over funds entering the country as remittances from international lottery scams. Therefore, the net effect of remittances on life expectancy is not immediately apparent. Thus, this variable is a prime candidate for inclusion in our analysis. We would have liked to include variables such as some measure of education, health expenditure, access to safe drinking water, etcetera. Unfortunately, those variables were either not consistently available over the period or were found to be poor predictors (e.g., foreign direct investment).

Data sources

The data for Jamaica are annual and span the period from 1976 to 2019. CO_2 emissions, measured in tonnes, were sourced from Our World in Data (2023) and divided by the total population recorded by the World Bank (2023). The resultant CO_2 emissions per capita is denoted as $CO2PC$. Life expectancy, denoted $LIFE$, is the number of years a newborn is expected to live under current mortality conditions. GDP per capita is referred to as $GDPPC$. It is measured in constant 2015 prices expressed in US dollars per person. The infant mortality rate is denoted MOR. It shows the number of infants who do not survive past their first year per 1,000 live births annually. Personal remittances, which include individual transfers and wages earned by employees, encompass all current monetary or material transfers sent or received by local households from persons overseas. Remittances are measured in per capita constant 2015 prices expressed in US dollar terms. It is denoted as $REMPC$. The data on $LIFE$, $GDPPC$, MOR, and $REMPPC$ were all obtained from the World Bank (2023).

ARDL

We apply the natural logarithm to $CO2PC$, $LIFE$, $GDPPC$, MOR, and $REMPPC$. We denote them as $LNCO2PC$, $LNLIFE$, $LNGDPPC$, $LNMOR$, and $LNREMPPC$, respectively. Next, we employ the ARDL bounds testing approach to examine cointegration and analyze the corresponding short-run and long-run cointegrating forms. It is important to note that when using the ARDL bounds testing methodology, the order of integration of the variables cannot exceed one. Thus, we assess the orders of integration to confirm that none of our variables possess an integration order of two or higher. We use the Elliott-Rothenberg-Stock Unit Root Test, as proposed by Elliot, Rothenberg, and Stock (1996).

The ARDL testing methodology can be explained by referencing a generic example involving three variables: y, x, and z. In this example, we determine if a long-run relationship exists between y, x, and z where y is the dependent variable and x and z are the independent variables. We construct an unrestricted (unconstrained) error correction model to conduct this determination. This model is shown in Equation (1).

$$\Delta y_t = a_0 + \sum_1^q a_{1i}\Delta y_{t-i} + \sum_0^r a_{2i}\Delta x_{t-j} + \sum_0^s a_{3i}\Delta z_{t-k} + b_1 y_{t-1} + b_2 x_{t-1} + b_3 z_{t-1} + v_t \qquad (1)$$

We choose the lag lengths q, r, and s with reference to the Akaike information criterion so that the residuals are serially uncorrelated (Breusch–Godfrey test), normally distributed (Jarque–Bera test), and homoscedastic (Breusch–Pagan–Godfrey test). By utilizing the estimates derived from equation (1), we can perform the ARDL bounds test. This involves an F-test and a t-test. These tests allow us to assess the presence of a long-run relationship. To conduct the tests, we compare the lower and upper bound critical values, determined at selected levels of statistical significance, with the calculated F-test and t-test statistics obtained from equation (1). The null hypothesis for the F-test is $b_1 = b_2 = b_3 = 0$ and the one for the *t*-test is that $b_1 = 0$. Our decision criterion for both tests is to reject the null hypothesis at the chosen level of statistical significance if the calculated test statistic exceeds (in absolute terms) the upper-bound critical value. Therefore, we conclude that there exists a significant overall relationship at the selected level of statistical significance only when we reject the null hypotheses of both the F-test and t-test.

If there is a level relationship, we reformulate equation (1) in the error correction form to obtain equation (2).

$$\Delta y_t = a_0 + \sum_1^q a_{1i}\Delta y_{t-i} + \sum_0^r a_{2i}\Delta x_{t-j} + \sum_0^s a_{3i}\Delta z_{t-k} + \gamma e_{t-1} + v_t \qquad (2)$$

In equation (2), v_{t-1} represents the error correction term. This is defined as $e_{t-1} = y_{t-1} - \hat{b}_0 - \hat{b}_2 x_{t-1} + \hat{b}_3 z_{t-1}$. The coefficient γ on the error correction term is the speed of adjustment parameter. It is important to note that a cointegrating and error-correcting level relationship typically occurs when $-1 \leq \gamma < 0$. This relationship's stability is assessed using three different tests: the cumulative sum (CUSUM) of the recursive residuals, the CUSUM of the squares of the recursive residuals, and the regression specification error test formulated by Ramsey (1969).

QARDL

To further assess the adequacy of the ARDL model, we examine if the dynamics of the short and long run are invariant over the distribution of the dependent variable. To do this, we estimate a QARDL model, a quantile regression applied to the preferred ARDL model. This model will show whether independent variables have varied effects on the dependent variable across its distribution rather than just focusing on the mean effects. In effect, if there is consistency between the results of the ARDL and QARDL models, this adds robustness to the findings. The QARDL model allows us to examine the relationships between variables across different quantiles of the conditional distribution of the dependent variable. We examine five quantile limits: 0.2, 0.4, 0.5, 0.6, and 0.8. For a detailed explanation of the mechanics of the QARDL model, interested readers can refer to Cho, Kim and Shin (2015). Here we discuss, in generality, a generic form of the QARDL model as presented below in equation (3):

$$G_y(\tau|X_t) = \alpha(\tau) + \sum_{i=1}^{l} \beta_i(\tau) y_{t-i} + \sum_{j=0}^{m} \gamma_j(\tau) X_{t-j} + \epsilon_t(\tau) \qquad (3)$$

Here $G_y(\tau|X_t)$ is the τ-th quantile conditional distribution of y_t, and X_t the vector of independent variables. The specific quantile intercept and coefficients on the lagged dependent and independent variables are given by $\alpha(\tau)$, β_i and γ_j, respectively. The error term for a given quantile is denoted by $\epsilon_t(\tau)$. The parameters are estimated by minimizing the weighted sum of absolute residuals based on each specified quantile τ.

If the ARDL results are confirmed across multiple quantiles in the QARDL framework, the reliability of our findings is strengthened. We carry out two tests to determine if the ARDL results are confirmed. The first test is the quantile slope equality test, which is grounded in the work of Koenker and Bassett (1982). The null hypothesis of this test assumes that the coefficients of the variables across different quantiles are equal. The second test is the Newey and Powell (1987) test of conditional symmetry. This test's null hypothesis states that the variables' average coefficient values for symmetric quantiles around the median are statistically equal to the coefficient values at the median.

FINDINGS

Unit root test and ARDL bounds testing

Table 1 presents the results from the Elliot-Rothenberg-Stock unit root test. At levels, the variables are statistically significant. However, at their first differences, all variables are statistically significant at the 5% level or 1% level. Therefore, the null hypothesis that the series contains a unit root is rejected only at the first difference. These results allow us to apply the ARDL bounds test, which accommodates non-stationary variables.

Table 1: Elliott-Rothenberg-Stock unit root test

Variable	**t-statistic**	
$LNLIFE$	-2.672	
$\Delta LNLIFE$	-3.299	**
$LNREMPC$	-2.201	
$\Delta LNREMPC$	4.335	***
$LNGDPPC$	-2.310	
$\Delta LNGDPPC$	-3.947	***
$LNCO2PC$	-2.433	
$\Delta LNCO2PC$	-7.265	***
$LNMOR$	-2.545	
$\Delta LNMOR$	-2.224	**

Notes: (1) *** and ** represent statistical significance at the 1% and 5% levels, respectively. (2) The test's null hypothesis is that the series contains a unit root.

Table 2: ARDL bounds cointegration test

	F-test statistic		**t-test statistic**	
	5.204	***	-5.787	***
Statistical significance	**I(0)**	**I(1)**	**I(0)**	**I(1)**
10%	2.450	3.520	-2.570	3.660
5%	2.860	4.010	-2.860	3.990
1%	3.740	5.060	-3.430	-4.600

Notes: (1) The null hypothesis is no long-run level relationship. (2) Long run level equation is *LNLIFE =F(LNREMPC, LNGDPPC LNMOR, LNCO2PC)*. (3) *** indicates statistical significance at the 1% level. (2) F-test rejects the null hypothesis of no level relationship at 1% level as F-test statistic > F-test I(1) 1 % critical value (5.204 > 5.060). (3) t-test rejects the null hypothesis of no level relationship at 1% level as t-test statistic < t-test I(1) 1 % critical value (-5.787 < -4.600).

Table 2 reports both F-test and t-test statistics. The estimated F-test statistic is 5.204, greater than the upper bound value of 5.060 at the 1% level of statistical significance. The absolute value of the t-test statistic (5.787) is also greater than the absolute upper bound value (4.600) at the 1% level of statistical significance.

These results indicate a long-run cointegrating relationship among the level variables when *LNLIFE* is treated as the dependent variable. Therefore, we estimate the ARDL long and short-run equations.

ARDL results

We present the long-run coefficients and the short-run error correction term in Table 3. Not surprisingly, the coefficient of the error correction term is negative and statistically significant at the 1% level. The absolute size of this coefficient is 0.482, which suggests that approximately 48% of any movement from the long-run equation is corrected in the first year. In other words, the economy takes just over two years to converge to the long-run equilibrium.

Table 3: ARDL error correction term and long-run equation

	Coefficient		Standard error
Short run			
Error correction term	-0.482	***	0.088
Long Run			
$LNREMPC$	0.0003		0.006
$LNGDPPC$	0.173	***	0.027
$LNMOR$	0.016		0.013
$LNCO2PC$	-0.070	***	0.008
ARDL Model	ARDL (4, 1, 3, 0, 3)		
Observations	40		

Notes: (1) *** indicates a coefficient statistical significantly different from zero at the 1% level (2) Long run level equation is *LNLIFE =F(LNREMPC, LNGDPPC LNMOR, LNCO2PC)*.

In the long run equation, $LNGDPPC$ has a positive and statistically significant impact on $LNLIFE$ in Jamaica. The coefficient on $LNGDPC$ indicates that, on average, a 1% increase in GDP per capita increases life expectancy by 0.17%. Neither the infant mortality rate nor remittances per capita are statistically significant in the long-run life expectancy equation. It is rather surprising that we do not find any statistically significant impact of remittances on life expectancy. The existing literature rightfully argues that children in remittance-receiving households tend to have better health outcomes, including life expectancy (see Azizi, 2018). Nonetheless, we could not establish the same result using a macro-time series dataset for Jamaica. We contend that perhaps further analysis using household-level data is necessary to examine the association between remittances and life expectancy.

We are mainly interested in how CO_2 emissions dynamically impact Jamaica's life expectancy in the long run. The coefficient of $LNCO2PC$ is negative and statistically significant at the 1% level. We note that a 1% increase in per capita CO_2 reduces life expectancy by 0.07% in the long run. Thus, in line with Matthew et al. (2018), Azam, Uddin, and Saqib (2023), and Polcyn et al. (2023), we make the argument that air pollution could be responsible for reducing human life span.

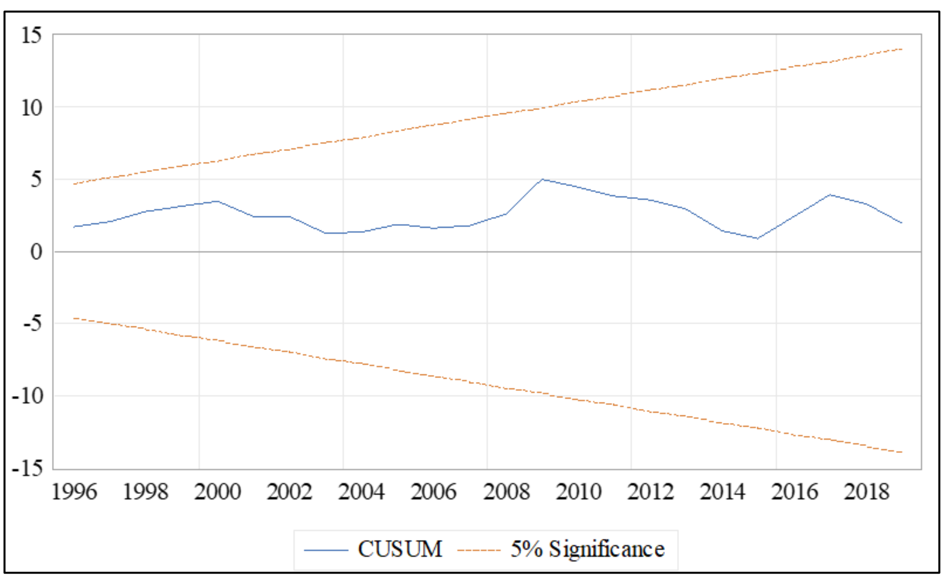
Next, we evaluate the statistical adequacy of the estimated model. Table 4 reports the results of the normality, serial correlation, heteroskedasticity and stability tests. The null hypotheses of no serial correlation, homoskedasticity, and normal distribution are not rejected even at the 10% level of statistical significance. The Ramsey RESET statistic is not statistically significant, which suggests the stability of the model's parameters. We present the CUSUM and CUSUM of squares diagrams in Figures 2 and 3. We note from these figures that recursive residual and recursive squares of residual lie within the 5% significance confidence bands. Therefore, the estimated model has no strong statistical evidence of variance or parameter instability.

Table 4: ARDL bounds test diagnostics

	Test statistic (Probability-value)	
Jarque-Bera normality, $\chi 2$ -test	1.951	(0.377)
Breusch-Godfrey serial correlation LM, F-test	0.916	(0.474)
Breusch-Pagan Godfrey heteroskedasticity, F-test	0.487	(0.924)
Ramsey RESET, F-test	0.007	(0.934)

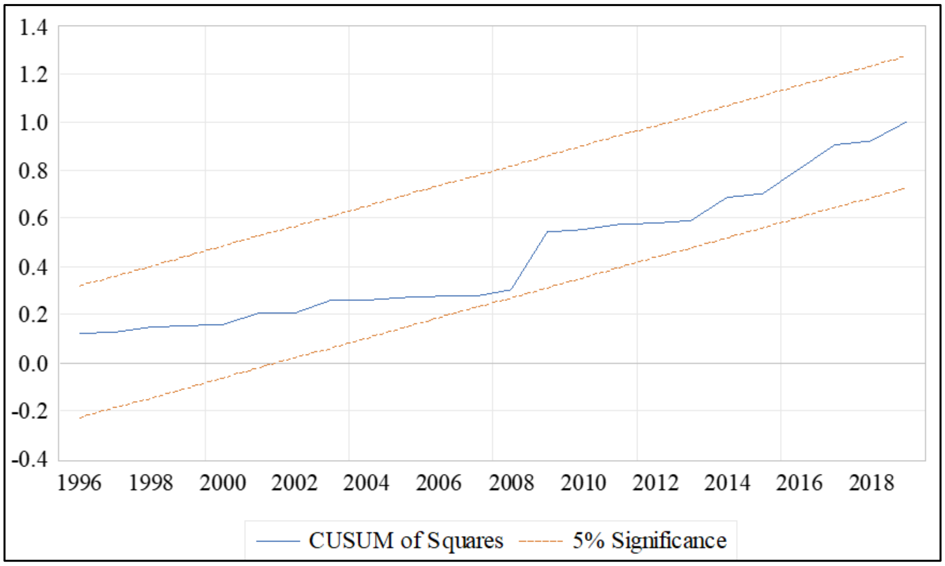
Notes: (1) The null hypotheses of none of the tests are rejected at the 10% level of statistical significance or less. (2) The null hypothesis of the Jarque-Berra test is that the data is normally distributed. (3) The null hypothesis of the Breusch-Godfrey Serial Correlation Lagrange Multiplier (LM) test is the absence of serial correlation. (4) The null hypothesis of the Breusch-Pagan-Godfrey test is homoskedasticity. (5) The null hypothesis of the Ramsey RESET test is that the ARDL model is not misspecified in the first order, wherein the powers of the fitted values of the dependent variable would have explanatory power.

Figure 2: CUSUM of recursive residuals test, forward recursion



Notes: (1) Test is based on Brown, Durbin, and Evans (1975). The CUSUM of recursive residuals is within the 5% significance bands, suggesting that the parameters of the ARDL model are stable.

Figure 3: CUSUM of squares of recursive residuals test, forward recursion



Notes: (1) Test is based on Brown, Durbin, and Evans (1975). (2) The CUSUM of squares of residuals is within the 5% significance bands, suggesting that the residual variance of the ARDL model is stable.

QARDL Results

The next step in our analysis is to conduct a robustness check of the ARDL model to reinforce further the statistical strength of the inference drawn from the short-run error correction term and long-run level coefficients. To do this, we apply the QARDL to the preferred ARDL model. After this, we conduct the Koenker and Bassett (1982) test, designed to check the equality of coefficients, and the Newey and Powell (1987) test, which examines conditional symmetry. These tests are performed at 0.2, 0.4, 0.5, 0.6, and 0.8 quantile limits. Table 5 provides these test results based on the Wald test statistics utilized for the equality and conditional symmetry of coefficients.

Table 5: Quantile slope equality and symmetric quantiles Wald tests

Coefficient(s) on	χ^{2}-test statistic	Degrees of freedom	Probability-value
		Quantile slope equality	
Error correction term	1.709	4	0.789
LNREMPC, *LNGDPPC*, *LNMOR*, and *LNCO2PC.*	14.651	16	0.550
		Symmetric quantiles	
Error correction term	1.064	2	0.587

LNREMPC, *LNGDPPC*, *LNMOR*, and *LNCO2PC.*	8.510	10	0.579

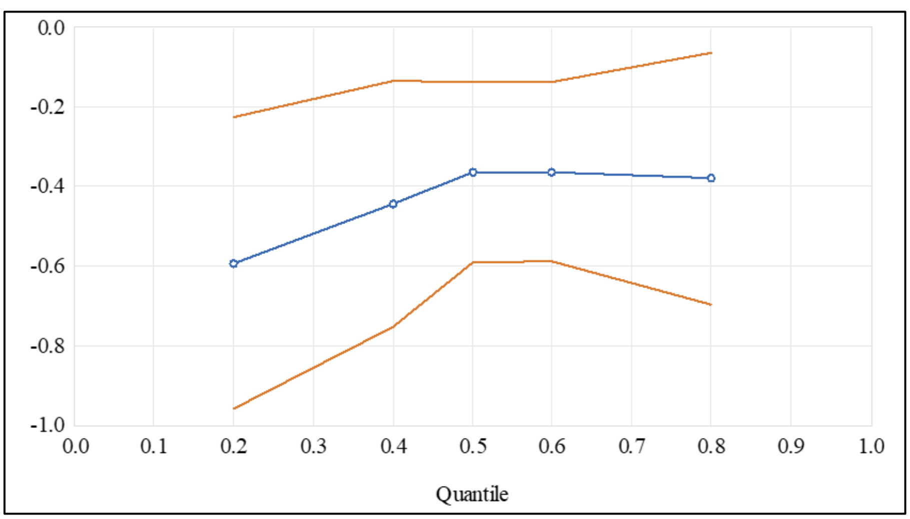
Notes: (1) Number of quantiles is 5. (2) Test statistic compares on indicated coefficient(s).

The results from this table indicate that there are no statistically significant differences across the different quantiles of the coefficients of the error correction term. The probability values associated with the test statistics exceed the 10% threshold. Similarly, we also find general stability across the quantiles for the dependent variables from the long-run level equation. The probability values associated with the test statistics for equality and conditional symmetry of the coefficients are also greater than the 10% threshold value.

Figure 4 and Figure 5 display visual plots to evaluate the stability of the coefficient values. The focus is also on the coefficient of the error correction term and those of the natural logarithm of the levels of independent variables from the long-run equation. This evaluation is performed at 0.2, 0.4, 0.5, 0.6, and 0.8 quantile limits. In addition to the quantiles, the plots depict each quantile's corresponding 95% confidence intervals. By incorporating confidence intervals, we allow for an assessment of the consistency and dependability of the coefficient values throughout the different quantiles. From these figures, we observe that the stability of the coefficients is maintained across the various quantiles. At the same time, it should be noted from Figure 5 that the $LNGDPPC$ and $LCO2PC$ have weaker significance at the 0.2 quantile limit as the 95% confidence band includes zero for that quantile.

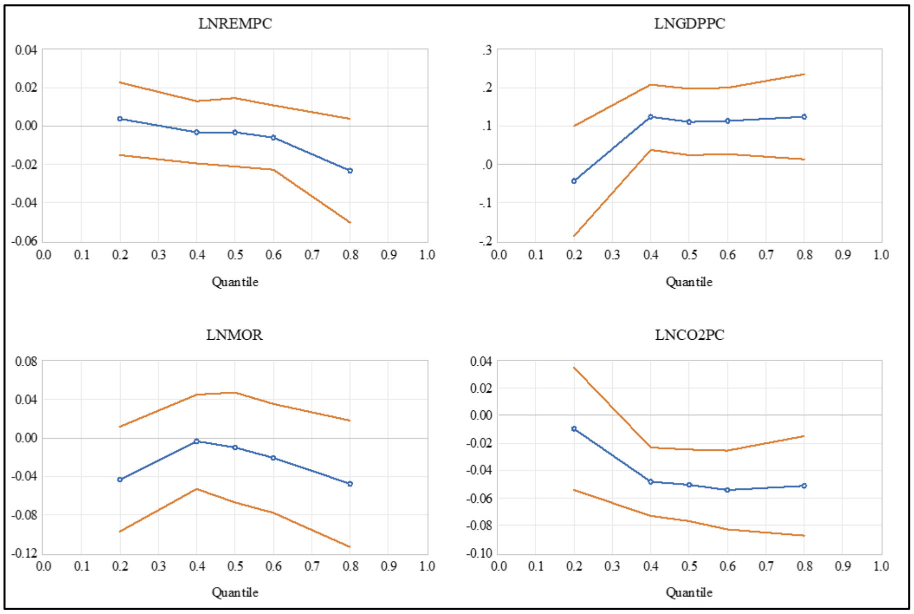
The general stability observed concerning the equality of coefficients and their conditional symmetry suggests that these coefficients tend to stay relatively constant despite varying quantile levels and the inherent variability within the data. This stability reinforces the soundness of the inference drawn from the ARDL model.

Figure 4: Quantile estimates of the speed of adjustment with its 95% confidence interval



Notes: (1) 95% confidence interval (CI) is the solid line without the marker. (2) Speed of adjustment refers to the coefficient on the error correction term.

Figure 5: Quantile estimates of the long run levels of the dependent variables with their 95% confidence interval



Notes: (1) 95% confidence interval (CI) is the solid line without the marker.

CONCLUSION AND DISCUSSION

We conducted a comprehensive and statistically robust study of the impact of CO_2 emissions on life expectancy in Jamaica using annual data from 1976 to 2019. Our study stands out from previous research as it is the first for this country that often grapples with the consequences of climate change. We accounted for changes over time in national real value added, infant mortality, and remittances to control for changes in socioeconomic conditions that could confound the impact of CO_2 emissions on life expectancy. An autoregressive bounds test modelling approach to cointegration was applied, with diagnostics tests and robustness checks by quantile regression confirming overall model adequacy across five quantile limits. Our central finding is that a long-run cointegrating relationship runs from CO_2 emissions to life expectancy, with a modest speed of adjustment. In this relationship, CO_2 emissions have a negative long-run elasticity with respect to life expectancy. We estimate this elasticity at approximately -0.07%, which is inelastic. Therefore, we contend that some air pollutants from CO_2 emissions impact life expectancy in the long run, even after controlling for changes in the socioeconomic environment. This study's results align with the World Health Organization's (2017) findings, which found that the mean annual temperature in Jamaica is expected to rise by almost 3.6 degrees by 2100 from its 1990 level and that air pollution poses a substantial health risk for Jamaica.

Our findings have important implications for policymakers. First, given the adverse effects of CO_2 emissions on life expectancy, it is essential to prioritize reducing CO_2 emissions to protect the health and well-being of the population. This calls for implementing sustainable practices and promoting renewable energy sources, which are relatively less polluting. In this regard, increased use of solar, wind and nuclear power as energy sources could mitigate CO_2 emissions. The government could consider subsidizing renewable energy production at the industrial and household levels. Second, the government could promote the development of green technologies to reduce CO_2 emissions. Historically, traditional and indigenous knowledge has been crucial in preserving the environment. In the *Climate Change Policy Framework* for Jamaica, the Government of Jamaica (2015) recognized the importance of traditional knowledge in combating climate change. The government must use such knowledge and scientific discovery to reduce CO_2 pollution.

Third, while the government has an important role, a holistic approach that incorporates the public, non-governmental organizations, and other stakeholders is essential to address the health impacts of CO_2 pollution. As such, anti-climate change training for the people could increase climate awareness. Fourth, in Jamaica, as the World Health Organization (2017) pointed out, air pollution is primarily responsible for mortality from respiratory infections, lung cancer and cardiovascular disease. Nonetheless, Jamaica's public healthcare system suffers from an acute shortage of healthcare providers, inadequate infrastructure, and a lack of public funding (Pan American Health Organization, n.d.). Given the impact of CO_2 emissions on life expectancy, it is crucial that the government mobilize sufficient funds to overhaul the healthcare system. Fifth, like other developing countries, Jamaica's women, children, and low-income people are the most vulnerable to climate change impacts (Government of Jamaica, 2015). The World Health Organization (2017) argued that compared to men, women and children are more likely to suffer from heart disease, stroke and lung cancer caused by household air pollution. A government initiative to replace solid fuels as energy sources with cleaner energy sources in the household will reduce health inequity in Jamaica.

Finally, Jamaica is highly vulnerable to external events (e.g., climatic ones) and relies heavily on imports and tourism (World Bank, 2024). The implication is that any adverse shocks to imports and tourism can act as stressors to the healthcare system, which, in turn, can have deleterious impacts on healthcare outcomes such as life expectancy. Thus, policymakers can prioritize interventions that address emissions reduction and resilience to external shocks. This dual approach can help mitigate the adverse effects on public health and enhance life expectancy.

REFERENCES

Adedoyin, F. F., & Zakari, A. (2020). Energy consumption, economic expansion, and CO_2 emission in the UK: the role of economic policy uncertainty. *Science of the Total Environment*, 738(October), 140014.

Ali, A., & Audi, M. (2016). The impact of income inequality, environmental degradation and globalization on life expectancy in Pakistan: An empirical analysis. MRPA Paper, No. 71112. Available at https://mpra.ub.uni-muenchen.de/71112/1/MPRA_paper_71112.pdf

Alola, A. A., & Adebayo, T. S. (2023). Analysing the waste management, industrial and agriculture greenhouse gas emissions of biomass, fossil fuel, and metallic ores utilization in Iceland. *Science of The Total Environment*, 887(August), 164115.

Amuedo-Dorantes, C., & Pozo, S. (2009). New evidence on the role of remittances on healthcare expenditures by Mexican households. *Review of Economics of the Household*, 9(1), 69–98.

Al-Mulali, U. & Ozturk, I. (2015). The effect of energy consumption, urbanization, trade openness, industrial output, and the political stability on the environmental degradation in the MENA (Middle East and North Africa) region. Energy, 84(May), 382–389.

Apergis, N., Gupta, R., Lau, C.K.M., & Mukherjee, Z. (2018). U.S. state-level carbon dioxide emissions: Does it affect healthcare expenditure? *Renewable and Sustainable Energy Reviews*. 91(March), 521–530.

Azam, M., & Adeleye, B.N. (2023). Impact of carbon emissions sources on life expectancy in Asia and the Pacific region. *Natural Resources Forum*. First published: 29 May 2023.

Azam, M., Uddin, I., & Saqib, N. (2023). The determinants of life expectancy and environmental degradation in Pakistan: evidence from ARDL bounds test approach. *Environmental Science and Pollution Research*, 30(1), 2233-2246.

Azizi, S. (2018). The impacts of workers' remittances on human capital and labor supply in developing countries. *Economic Modelling*, 75(November), 377-396.

Beatty, T.K.M and Shimshack, J.P. (2014). Air pollution and children's respiratory health: A cohort analysis. *Journal of Environmental Economics and Management*. 67(1), 39-57.

Bekhet, H.A., Matar, A., & Yasmin, T. (2017). CO_2 emissions, energy consumption, economic growth, and financial development in GCC countries: Dynamic simultaneous equation models. *Renewable and Sustainable Energy Reviews*, 70(April), 117–132.

Beuermann, D. W., Ruprah, I. J., & Sierra, R. E. (2016). Do remittances help smooth consumption during health shocks? Evidence from Jamaica. *The Journal of Developing Areas*, 50(3), 1-19.

Brown, L., McFarlane, A., Das, A., &. Campbell, K. (2020). Remittances and CO_2 Emissions in Jamaica: An Asymmetric Modified Environmental Kuznets Curve. *Journal of Economic Asymmetries*, 22, e00166.

Campbell, K., Das, A., Brown, L., & McFarlane, A. (2024). Remittances and homicides in Jamaica. *Journal of Money Laundering Control*, 27(4), 763-779.

Centers for Disease Control and Prevention. (2022). *Infant mortality*. Centers for Disease Control and Prevention. Accessed on August 1, 2023. Available:
https://www.cdc.gov/reproductivehealth/maternalinfanthealth/infantmortality.html

Chaabouni, S., Zghidi, N., & Mbarek, M.B. (2016). On the causal dynamics between CO_2 emissions, health expenditures and economic growth. *Sustainable Cities and Society*, 22(February), 184-191.

Cho, J. S., Kim, T., & Shin, Y. (2015). Quantile cointegration in the autoregressive distributed-lag modeling framework. *Journal of Econometrics*, 188(1), 281–300.

Danish, Wang, B., & Wang, Z. (2018). Imported technology and CO2 emission in China: Collecting evidence through bound testing and VECM approach. *Renewable and Sustainable Energy Reviews*, 82, 4204–4214.

Das, A., Brown, L., &., McFarlane, A. (2023). Economic misery and remittances in Jamaica. *Journal of Economic Development*, 48(2), 33-52.

Das, A., McFarlane, A., & Yung, J. (2019). Remittances and GDP in Jamaica: An ARDL bounds testing approach to cointegration. *International Economic Journal*, 33(2), 365-381.

Deaton, A. (2015). *The great escape: Health, wealth, and the origins of inequality*. Princeton University Press.

Demir, S., Demir, H., Karaduman, C., & Cetin, M. (2023). Environmental expenditure and health expenditure efficiency in Turkiye: The role of natural resources. *Environmental Science and Pollution Research*, 30, 15170-15185.

Djeunankan, R., & Tekam, H. (2022). Do remittances matter for health outcomes in developing countries? Fresh evidence from a panel vector autoregressive (PVAR) model. *International Journal of Development Issues*, *21*(3), 458–482.

Elliott, G., Rothenberg, T. J., & Stock, J. H. (1996). Efficient Tests for an Autoregressive Unit Root. *Econometrica: Journal of the Econometric Society*, 64(4), 813-836.

Feruglio, N., Saxena, S., & von Thadden-Kostopoulos, S. (2023). Jamaica: Technical Assistance Report – Climate Public Investment Management Assessment (C-PIMA), IMF Country Report, No. 23/236, Washington, D.C.: International Monetary Fund. Available at:
https://www.elibrary.imf.org/downloadpdf/journals/002/2023/236/article-A000-en.xml

Fiore, A. M., Naik, V., & Leibensperger, E. M. (2015). Air quality and climate connections. *Journal of the Air & Waste Management Association*, *65*(6), 645-685.

Government of Jamaica (2015). Climate Change Policy Framework for Jamaica. Available at: https://www.lse.ac.uk/GranthamInstitute/wp-content/uploads/2016/05/Jamaica-Climate-Change-Policy-fwL-2015.pdf

Haines, A., Kovats, R. S., Campbell-Lendrum, D., & Corvalán, C. (2006). Climate change and human health: Impacts, vulnerability and public health. *Public Health*, *120*(7), 585-596.

Hill, T.D., Jorgenson, A.K., Ore, P., Balistreri, K.S., Clark, B. (2019). Air quality and life expectancy in the United States: An analysis of the moderating effect of income inequality. SSM – Population Health. 7(April), 100346.

Jerrett, M., Burnett, R. T., Pope III, C. A., Ito, K., Thurston, G., Krewski, D., Shi, Y., Calle, E. & Thun, M. (2009). Long-term ozone exposure and mortality. *New England Journal of Medicine*, *360*(11), 1085-1095.

Kadir, M. O., Deka, A., Ozdeser, H., Seraj, M., & Turuc, F. (2023). The impact of energy efficiency and renewable energy on GDP growth: new evidence from RALS-EG cointegration test and QARDL technique. *Energy Efficiency*, 16(5), 46.

Koenker, R., & Bassett, G. (1982). Robust tests for heteroscedasticity based on regression quantiles. *Econometrica*, 50(1), 43–61.

Landrigan, P.J., Fuller, R., Hu, H., Caravanos, J., Cropper, M.L., Hanrahan, D., Sandilya, K., Chiles, T.C., Kumar, P., & Suk, W.A. (2018). Pollution and Global Health – An Agenda for Prevention. *Environmental Health Perspectives*, 126(8), 084501-084501-6.

Mahalik, M.K., Le, T-H., Le, H-C., & Mallick, H. (2022). How do sources of carbon dioxide emissions affect life expectancy? Insights from 68 developing and emerging economies. *World Development Sustainability*, 1, 100003.

Matthew, O., Osabohien, R., Fasina, F., & Fasina, A. (2018). Greenhouse gas emissions and health outcomes in Nigeria: Empirical insight from ARDL technique. *International Journal of Energy Economics and Policy*, 8(3), 43-50.

McFarlane, A., Brown, L., & Das, A. (2023). Remittances and fertility in Jamaica. *Applied Economics Letters*, Published online: 5 March 2023.

Ministry of Health and Wellness (n.d.) Vision for Health 2030. Available at: https://www.moh.gov.jm/wp-content/uploads/2019/05/MOHW-Vision-for-Health-2030-Final.pdf

Myers, S. S., Zanobetti, A., Kloog, I., Huybers, P., Leakey, A. D., Bloom, A. J., Carlisle, E., Dietterich. L. H., Fitzgerald, G., Hasegawa, T., Holbrook, N. M., Nelson, R. L., Ottman, M. J., Raboy, V., Sakai, H., Sartor, K. A., Schwartz, J., Seneweera, S., Tausz, M., & Usui, Y. (2014). Increasing CO_2 threatens human nutrition. *Nature*, *510*(7503), 139-142.

Newey, W. K., & Powell, J. L. (1987). Asymmetric least squares estimation and testing. *Econometrica*, 55(4), 819–847.

Nkalu, C.N., & Edeme, R.K. (2019). Environmental hazards and life expectancy in Africa: Evidence from GARCH model, *Sage Open*. 9(1), 2158244019830500.

Our World in Data. (2023). *CO_2 Emissions*. Retrieved June 1, 2023. https://ourworldindata.org/co2-emissions

Pan American Health Organization (n.d.). Strengthening Health Systems & Services in Jamaica. Available at: https://www.paho.org/en/stories/strengthening-health-systems-services-jamaica

Patz, J. A., Campbell-Lendrum, D., Holloway, T., & Foley, J. A. (2005). Impact of regional climate change on human health. *Nature*, *438*(7066), 310-317.

Planning Institute of Jamaica (n.d.). Terms of Reference for the Project Final Evaluation Improving Climate Data & Information Management Pilot Program for Climate Resilience in Jamaica. Available at: https://www.pioj.gov.jm/wp-content/uploads/2022/02/TOR-Project-Final-Evaluation-final.pdf

Polcyn, J., Voumik, L. C., Ridwan, M., Ray, S., & Vovk, V. (2023). Evaluating the influences of health expenditure, energy consumption, and environmental pollution on life expectancy in Asia. *International Journal of Environmental Research and Public Health*, 20(5), 4000.

Preston, S. H. (1975). The changing relation between mortality and level of economic development. *Population Studies*, 29(2), 231–248.

Public Health Agency of Canada (2023). Social determinants of health and health inequalities - Canada.ca. Available at: https://www.canada.ca/en/public-health/services/health-promotion/population-health/what-determines-health.html

Rahman, M. M., Rana R., & Khanam, R. (2022). Determinants of life expectancy in most polluted countries: Exploring the effect of environmental degradation. PLoS ONE, 17(1), e0262802.

Ramsey, J. B. (1969). Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society: Series B*, 31(2), 350–371.

Sarkodie, S.A., Strevoz, V., Jiang, Y., & Evans, T. (2019). Proximate determinants of particulate matter ($PM_{2.5}$) emission, mortality and life expectancy in Europe, Central Asia, Australia, Canada, and the US. *Science of the Total Environment*, 683(May), 489-497.

Schwartz, J.D., Wang, Y., Kloog, I., Yitshak-Sade, M., Dominici, F., & Zanobetti, A. (2018). Estimating the effects of PM2.5 on life expectancy using causal modeling methods. *Environmental Health Perspectives*, 126(12), 127002-1 – 127002-9.

Sharif, A., Baris-Tuzemen, O., Uzuner, G., Ozturk, I., & Sinha, A. (2020). Revisiting the role of renewable and non-renewable energy consumption on Turkey's ecological footprint: Evidence from Quantile ARDL approach. *Sustainable Cities and Society*, 57(June), 102138.

Shen, Y., Su, Z. W., Malik, M. Y., Umar, M., Khan, Z., & Khan, M. (2021). Does green investment, financial development and natural resources rent limit carbon emissions? A provincial panel analysis of China. *Science of the Total Environment*, 755(February), 142538.

United Nations (2023). Extreme weather caused two million deaths, cost $4 trillion over last 50 years. Available at: https://turkiye.un.org/en/232828-extreme-weather-caused-two-million-deaths-cost-4-trillion-over-last-50-years

Wang, Z., Asghar, M.M., Zaidi, S.A.H., & Wang, B. (2019). Dynamic linkages among CO_2 emissions, health expenditures and economic growth: empirical evidence from Pakistan. *Environmental Science and Pollution Research*, 26(March), 15285-15299.

Wang, Z., Asghar, M.M., Zaidi, S.A.H., Nawaz, K., Wang, B., Zhao, W., & Xu, F. (2020). The dynamic relationship between economic growth and life expectancy: Contradictory role of energy consumption and financial development in Pakistan. *Structural Change and Economic Dynamic*, 53(March). 257-266.

Wang, Q., Wang, X., & Li, R. (2022). Does urbanization redefine the environmental Kuznets curve? An empirical analysis of 134 Countries. *Sustainable Cities and Society*, 76(January), 103382.

World Bank. (2023). *World Development Indicators*. World Bank. Retrieved June 1, 2023. https://databank.worldbank.org/source/world-development-indicators.

World Bank. (2024). The World Bank In Jamaica Overview. Retrieved June 9, 2024. https://www.worldbank.org/en/country/jamaica/overview

World Health Organization (2022). Ambient (outdoor) air pollution. Available at: https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health

World Health Organization (2017). Climate and Health Country Profile- 2017: Jamaica. Available at: https://apps.who.int/iris/rest/bitstreams/1097478/retrieve

Yazdi, S., Zahra, T., & Nikos, M. (2014). Public healthcare expenditure and environmental quality in Iran. *Recent Advances in Applied Economics*, *1*, 126-134.

Yazdi, S. K., & Khanalizadeh, B. (2017). Air pollution, economic growth, and health care expenditure. *Economic Research-Ekonomska Istraˇzivanja*, 30(1), 1181–1190.

Zaidi, S. & Saidi, K. (2018). Environmental pollution, health expenditure and economic growth in the sub-Saharan Africa countries: Panel ARDL approach. *Sustainable Cities and Society*, 41(August), 833-840.

Zakaria, M. & Bibi, S. (2019). Financial development and environment in South Asia: The role of institutional quality. *Environmental Science and Pollution Research*, 26 (8), 7926–7937.

The Motherhood Penalty

Mohammad (Mitu) Ashraf*

ABSTRACT

This paper presents a formal model that shows the size of the motherhood penalty—an earnings penalty due to bearing and rearing children. Lower earnings and a lower financial status often translate into lower socio-political status of women as compared with their male counterparts. An implication of the model presented in this paper is that legislating a paid maternity leave may not be enough to ameliorate the motherhood penalty and bring the financial status of female workers equal to their male counterparts. The paper also suggests a solution to remedy the motherhood penalty.

JEL Code: J24, J31

Key Words: Gender Wage Gap, Motherhood Penalty, Externalities

INTRODUCTION

It is a biological imperative that women bear children. Children borne by the current generation of women become labor force in the next generation. Women also tend to be the primary caretakers of children and other family members (Cravey and Mitra 2011, 306-311; Pelkowski 2005, 17-39; Wolf and Soldo 1994, 1259-1276). To bear and rear children, however, women must either take time off from labor market activities or significantly decrease labor market participation. This results in a decreased human capital accumulation of female workers which, in turn, leads to lower labor market earnings (Blau and Kahn 2000, 75-99; Blau and Kahn 2013, 251-256; Goldin 2014, 1091-1119; Browning 1992, 1434-1475; Lundborg, Plug, and Rasmussen 2017, 1611-1637; Waldfogel 1998, 137-156; Ferber and Lowry 1976, 377-387; Fernandez 2014, 37-80; Blau and Kahn 2017, 789-865; Hotchkiss and Pitts 2007, 417-421; Hotchkiss, Pitts, and Walker 2017, 3509-3522; Barth, Kerr, and Olivetti 2017; Ferber and Spaeth 1984, 260-264; Adda, Dustmann, and Stevens 2017, 293-337). A lower financial status of women often translates into their lower socio-political standing.

A rich body of literature has provided empirical evidence of gender-wage gap in the labor market. Studies have identified numerous causes. These range from gender discrimination to job characteristics to work intermittency due to childbearing and rearing. Most of these studies use reduced-form models and regression analyses to tease out factors that show the male-female wage gap.[1] The present study fills this

*Department of Economics, Marketing, Entrepreneurship and Analytics, Thomas College of Business and Economics, The University of North Carolina at Pembroke, One University Drive, P.O. Box 1510, Pembroke, NC 28372-1510, Phone: (910) 521-6464, Email: ashraf@uncp.edu
I am grateful to the anonymous referees for their invaluable comments. All errors belong to the author.

gap in the literature. I present a simple theoretical model that shows how labor-market intermittency due to childbearing and rearing may generate gender wage and income gap.

Along with capital, the quality and quantity of the labor force is a main source of economic growth for an economy (Solow 1956, 65-94; Mankiw, Romer, and Weil 1992, 407-437; Barro 1991, 407-443; Aghion and Howitt 1998; Romer 1990, S71-S102). The model presented in this paper shows that even if there did not exist any other forms of gender discrimination or constraints in the labor market, there remains gender-wage gap due to the biological imperative that women bear children. I also argue that because the society benefits from women having children (*i.e.*, future labor force), without being compensated for producing future labor force, there exist positive externalities. An implication of the model presented in this paper is that legislating a paid maternity leave may not be enough to bring the financial status of female workers equal to their male counterparts. I suggest a way to compensate women for producing future labor force.

Note that the scope of the model presented in this paper is rather narrow. I do not question the motivation for having children, which may range from insurance against the uncertain future (Pörtner 2001, 119-136; Cain 1981, 435-474), to a host of personal preferences for procreation (Benatar 2017; Lundborg, Plug, and Rasmussen 2017, 1611-1637). I take the choices of women having children and being in the labor market as given. This means that the model presented here is only applicable to a society where women have such choices; they are neither forced to have children nor they are forced to enter labor market or barred from the labor market. As a result, the setup the model in this paper may not apply to societies where women do not have such choices either by fiat or custom.

The rest of the paper is arranged as follows. The next section presents a review of the relevant literature, followed by the section that presents a theoretical model that incorporates empirical findings. This is followed by a section that presents a compensation model to correct for externalities and discusses its implications. Finally, the last section gives some concluding remarks.

A REVIEW OF THE RELEVANT LITERATURE

I briefly review economic literature related to the effects of childbearing and rearing on gender-wage gap. The literature review presented here is not meant to be exhaustive. An excellent and exhaustive review of studies highlighting the various aspects of male-female wage differentials is presented in Francine Blau and Lawrence Kahn (Blau and Kahn 2017, 789-865). See also Vennessa Burbano, Nicolas Padilla, and Stephen Meier (Burbano, Padilla, and Meier 2024, 61-94) for a more recent exploration in this topic. Burbano et al. look at gender differences in the “meaningfulness” of various occupations, and how this perceived meaningfulness may affect the labor market outcome. The authors define meaningfulness at work as “an individual’s sense of impact as a result of their work: their understanding of the purpose and what they believe is achieved as a result of their work.” (p. 62). Their results indicate that male and female attitudes towards the impact of a given occupation differ significantly, and that these differences help explain (at least) some of the gender differences as they relate to labor market outcomes.

Studies show that having children impacts a female's labor force participation and earnings. This is the so-called "motherhood penalty" (Blau and Kahn 2017, 789-865; Lundborg, Plug, and Rasmussen 2017, 1611-1637). Economic theory and empirical evidence suggest that amount of penalty may vary by occupation. Occupations that require particular and long hours will have larger penalty as compared with occupations that have more flexible schedule, all else constant. This is because occupations in the latter category will allow a female employee to set work schedule around her motherly duties. Occupations that allow for schedule flexibility will also have higher female labor force participation rates and lower gender-wage gaps.

Claudia Goldin and Lawrence Katz (Goldin and Katz 2008, 363-369) find that the financial penalty for taking time off from labor market differs by occupation and education level. They use data for students graduating from some of the most elite universities. According to their results female earnings are negatively related with the number of children. However, "the negative impact of women's earnings is entirely accounted for by hours worked." (p.367).

Goldin and Katz (2008) also find that the time taken off labor market activities also differs by occupation and degree. "Physicians, for example, took the briefest nonemployment spells after having a child. PhDs were next in terms of length of spell, followed by lawyers, then MBAs and those with other types of master's or no further degrees, who took the greatest amount of time off for family reasons." (p.368). Furthermore, earnings penalty differed by the occupation. According to their findings, "the smallest earnings penalty was for physicians and other medical professionals (0.16 log points). Next in penalty size was those with a JD or a PhD (0.34 log points). The largest earnings penalty for time off was for those with an MBA (0.53 log points)." (p.368)

In her 2014 study Claudia Goldin (Goldin 2014, 1091-1119) looks at various occupations with respect to gender-wage gap. Goldin (2014) shows that "given age and time worked residual differences for Business occupations are large and residual differences in Technology and Science are small. In fact, for the 'young' group (less than 45 years old) some Technology and Science occupations have positive coefficients." (p.8). The reason for positive coefficients on "Technology and Science" occupations is the schedule flexibility afforded to employees in such occupations.

In a 2016 study Claudia Goldin and Lawrence Katz (Goldin and Katz 2016, 705-746) use data from the US Census, the American Community Surveys, the Current Population Surveys, and surveys of pharmacists to explore the development of the pharmacy profession over time. The authors compare the pharmacy profession with other professions in the US. Their findings lead them to conclude that in the US the pharmacists' profession is the most "egalitarian" as compared with other occupations that require similar level of training. The pharmacy profession has the lowest gender wage and income gaps. The authors argue that "Pharmacy is an egalitarian occupation not only in terms of gender. It also has smaller racial and ethnic wage differentials than are typical of college graduate occupations." (p.732). Goldin and Katz (2016) show that the changes that have taken place in the pharmacy profession are "less consistent with those stemming solely from an increase in the demand for family-friendly workplace amenities. The changes, moreover, do not appear to have resulted from legislation, antidiscrimination policies, licensing requirements, or regulations specific to the pharmacy profession. Rather, a host of structural changes outside the realm of the labor market increased the demand for pharmacists and reorganized work in ways that have made pharmacy a more family-friendly and female-friendly profession." (p.739)

Julie Hotchkiss, Melinda Pitts, and Mary Beth Walker (Hotchkiss, Pitts, and Walker 2017, 3509-3522) use a dataset that combines information about females' childbearing information (collected by the state of Georgia between 1994 and 2002) with the employer-employee data through 2008. The uniqueness of the dataset allows them to tease out the effects of intermittency due to the birth of the first child on a female's earnings. Their results suggest that "women with an average amount of intermittency during five years after the birth of her first child experience earnings that are roughly 18 percent lower than a woman with no intermittency." (p.16). This dataset also allows the authors to separate the effects of intermittency on a female's career path by education and the industry structure. The authors find that in their dataset the motherhood penalty for women with a college degree is twice the motherhood penalty of women with only a high school diploma. Their results also indicate that higher-earning women tend to self-select into more competitive industries and firms with greater number of establishments.

Nickolay Angelov, Per Johansson, and Erica Lindahl (Angelov, Johansson, and Lindahl 2016, 545-579) use Swedish data of couples to see if there is a motherhood penalty. They find that "15 years after the first child has been born, the male-female gender gaps in income and wages have increased by 32 and 10 percentage points, respectively." (p.545). The results of their study further indicate that women never catch up to their male counterparts. The reason, they argue, is that while there is not an immediate decline in human capital due to a maternity leave, male-female human capital differences however do accumulate as women have to allocate more time to childrearing as compared with their male counterparts. They argue that so long as women have the main responsibility to rear children, it is unlikely that the male-female wage gap will disappear.

Jerome Add, Christian Dustmann, and Katrien Stevens (Adda, Dustmann, and Stevens 2017, 293-337) use a dynamic life-cycle model to estimate the cost of having children for a woman along various dimensions. The authors use a unique dataset for Germany that allows them to measure the impact of having children on her career choices, her labor supply, her saving behavior, and her wealth accumulation over her labor market working life. The dataset also allows the authors to decompose lost earnings due to career intermittency versus skill loss. Their estimates indicate that "about three-quarters of the career costs of children stem from lost earnings due to intermittency or reduced labor supply, while the remainder is due to wage responses, as a result of lost investments in skills and depreciation." (p.295). Comparing two scenarios—women with and without children—their estimates show that "the overall costs of children are close to 35 percent of the net present value of income at age 15." (p.320). To estimate the impact of children on gender-wage gap, the authors construct a "counterfactual wage profile (conditional on working) of a woman who remains childless and conditions on that knowledge from the start of her career." According to their estimates, “The gender gap closes by about 0.2 log points when women are in their 30s, which corresponds to about a third of the overall gap.” (pp.322-323)

It is possible that women who have children would have had children regardless of their employment status. In the literature this is referred to as the adverse selection explanation. If that is the case, then ascribing career choices of mothers to motherhood will be incorrect. Another explanation of the effects of children on the career paths of mothers is called the causation explanation. That is, women make certain career choices because of having had children and the demands on their time that children place.

Peter Lundborg, Eric Plug, and Astrid Rasmussen (Lundborg, Plug, and Rasmussen 2017, 1611-1637), using a unique dataset of Danish women who had in vitro fertilization (IVF) treatment try to find which explanation—adverse selection or causation—is empirically supported. The IVF treatment for some women was successful (*i.e.,* resulted in childbirth) and for other it was not successful. Using the success of an IVF treatment as an instrument the authors find support for the causal explanation. They find that having children does have negative consequences for women’s career. Motherhood leads to a decline in incomes of females that persists throughout their careers. Their results also show that “women with children reveal a stronger preference for working close to home than women without children, and as a consequence end up working in less-well-paid jobs.” (p.1629)

From this brief literature review presented, and studies cited above, a few points become clear. One, there remains male-female wage gap even after controlling for education levels, industry structure, occupational preferences, and other labor market characteristics. Two, gender-wage gap is present in all countries that have been studies so far. Three, career intermittency has a negative impact on earnings and that this negative impact is larger for female workers than it is for their male counterparts. Four, career intermittency due to childbearing has a lasting negative impact on the earning profiles of female workers. Five, having policies that allow for paid maternity leave are no guarantee for equal financial status of males and females; countries where there are relatively generous maternity leave policies, mothers’ earnings profiles still stay below those of fathers after the birth of the child. And finally, the higher the education level the steeper the earning profile and the lager the motherhood penalty, all else constant. The theoretical model presented below incorporates these features of the labor market.

MODEL

This section presents a simple theoretical model of motherhood penalty in the labor market. It highlights the main contours of the male-female wage gap that arises due to childbearing and rearing that empirical research has found over the past half century. It is not meant to capture all the various nuances of gender discrimination in the labor market.

To keep the model simple, I assume that labor (L) is the only factor of production. At time t, the aggregate production function takes the form

(1) $$Y_t = F(L_t)$$

Workers are paid a wage (w) according to their marginal product (MP_L)

(2) $$w_t = MP_{L,t}$$

Total labor force at any time t is L_t. It is composed of male workers ($L_{m,t}$) and female workers ($L_{f,t}$).

$$L_t = L_{m,t} + L_{f,t}$$

At any given time, t, a fraction, θ, of female workers leaves the labor force to give birth. Females stay out of the labor force for a period, s, and return to labor force at time $t + s$. At time t total labor engaged in production is

(3) $$L_t = L_{m,t} + (1 - \theta)L_{f,t}$$

And the proportion of female labor force that is on maternity leave is $\theta L_{f,t}$. Equation (1) may be written as

(4) $$Y_t = F(L_{m,t} + [1 - \theta]L_{f,t})$$

This setup also implies that, with zero net immigration and the birth rate equal to death rate, the change in labor force is also equal to $\theta L_{f,t}$. For the ease of exposition, I ignore the subscript t for now.

Assume that both male and female workers have a total number of hours, H_m and H_f, respectively, and that $H_m = H_f$. Also assume that both male and female workers start in the labor market with equal level of human capital and workers learn on the job equally well which leads to changes in human capital.

Out of the total number of hours, H_m and H_f, both male and female workers spend certain number of hours on labor market activities, $h_{m,L}$ and $h_{f,L}$, and the remaining number of hours bearing (female workers) and rearing (both male and female workers) children $h_{m,C}$ and $h_{f,C}$.

$$H_m = h_{m,L} + h_{m,C}$$

$$H_f = h_{f,L} + h_{f,C}$$

Given the biological imperative that only females bear children and the empirical evidence that females are the main caregivers (Cravey and Mitra 2011, 306-311; Pelkowski 2005, 17-39; Wolf and Soldo 1994, 1259-1276), we have $h_{f,C} > h_{m,C}$ and that $h_{f,L} < h_{m,L}$ for at least part of the working life.

Total wage income of male and female workers, W_m and W_f, respectively, is a function of marginal product of labor *and* the number of hours worked.

(5) $$W_m = \omega_m(\beta[h_{m,L}, h_{m,C}], h_{m,L})$$

(6) $$W_f = \omega_f(\alpha[h_{f,L}, h_{f,C}], h_{f,L})$$

where $\beta(.)$ and $\alpha(.)$ are the male and female marginal productivities, respectively, and are functions of time spent on labor market activities and time spent on bearing and rearing children. Since both male workers and female workers learn-by-doing, we have

$$\frac{\partial\beta(.)}{\partial h_{m,L}} > 0; \frac{\partial\beta(.)}{\partial h_{m,C}} = 0$$

$$\frac{\partial\alpha(.)}{\partial h_{f,L}} > 0; \frac{\partial\alpha(.)}{\partial h_{f,C}} = 0$$

The equalities, $\frac{\partial\beta(.)}{\partial h_{m,C}} = 0$, and $\frac{\partial\alpha(.)}{\partial h_{f,C}} = 0$, follow from the assumption that learning-by-doing takes place only when workers are involved in labor market activities.

Furthermore, since male and female workers learn-by-doing equally well, we have

$$\frac{\partial\beta(.)}{\partial h_{m,L}} = \frac{\partial\alpha(.)}{\partial h_{f,L}}$$

From (5) and (6) it follows that the *change* in male wage income is

(7) $$\frac{dW_m}{dh_{m,L}} = \frac{\partial\omega_m}{\partial\beta(h_{m,L},h_{m,C})} \times \frac{\partial\beta(h_{m,L},h_{m,C})}{\partial h_{m,L}} + \frac{\partial\omega_m}{\partial h_{m,L}}$$

Analogously, the *change* in female wage income is

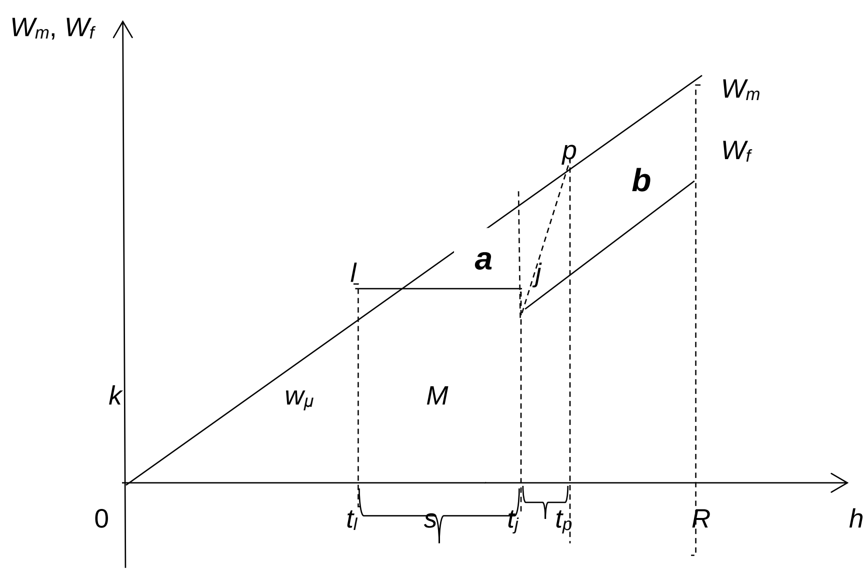
(8) $$\frac{dW_f}{dh_{f,L}} = \frac{\partial\omega_f}{\partial\alpha(h_{f,L},h_{f,C})} \times \frac{\partial\alpha(h_{f,L},h_{f,C})}{\partial h_{f,L}} + \frac{\partial\omega_f}{\partial h_{f,L}}$$

Equations (7) and (8) say that as the number of hours worked changes, it affects the change in a worker's wage income through two sources. One, via affecting the marginal productivities, β(.) and α(.)—learning-by-doing—and the other directly by the number of hours worked. Because during the paid maternity leave, $\frac{\partial\alpha(h_{f,L},h_{f,C})}{\partial h_{f,L}} = 0$, and $\frac{\partial\omega_f}{\partial h_{f,L}} = 0$, the *change* in the female wage is $\frac{dW_f}{dh_{f,L}} = 0$. The first term is equal to zero because her productivity remains unchanged, and the second term is equal to zero because she is not working in the labor market; she is being paid the wage rate when she started the maternity leave.

Put another way, all else constant, while earning curves of male and female workers without intermittency have positive and equal slopes throughout their careers, the earnings curve of a female worker who bears (and rears) children during the paid maternity leave has a slope equal to zero.

Figure 1 presents this argument graphically. For the sake of simplicity, I assume that the learning-by-doing is a linear function of number of hours worked during a given period.

Figure 1: Earning Curves of Male and Female Workers



I plot male and female earnings, W_m and W_f, respectively, along the vertical axis and the number of hours worked in the labor market, h, along the horizontal axis. (Note that with appropriate normalization, time and hours can be used interchangeably.) Both male and female workers start at the same starting wage, k. Learning-on-the-job takes place that leads to an increase in human capital and hence the wage rate. It is represented by the positive slope of the wage curve. Both male and female workers learn equally well and, hence, have the same slope of the earnings curves. A male employee's earning profile is represented by the curve kW_m. A female employee takes a paid maternity leave at time t_l. She rejoins the work force at time t_j. Her earning profile is represented by the curve $kljW_f$. Both male and female workers retire at time R.

In this diagram s is the period for which a female takes maternity leave. During the maternity leave her earning profile is flat, $\frac{dW_f}{dh_{f,L}} = 0$, and she is earning a wage rate equal to when she started the leave. This is represented by w_μ. Her income during that period is equal to the area M (= $w_\mu \times s$). Once she returns to labor market activities her earning profile has the same slope as that of her male counterpart, but this portion of the curve has shifted downward. As a result of bearing a child she has lost ground equal to the area ***a*** + ***b***. This area represents the motherhood penalty.

Lifetime earnings of both male and female workers are the areas below their respective curves. More formally, a male worker's lifetime earnings (LTE_m) are equal to the integral

(9) $$LTE_m = \int_{h_{m,L}=0}^{h_{m,L}=R} \frac{dW_m}{dh_{m,L}}\, dh_{m,L}$$

And a female worker's lifetime earnings (LTE_f) are equal to the integral

(10) $$LTE_f = \int_{h_{f,L}=0}^{h_{f,L}=R} \frac{dW_f}{dh_{f,w}}\, dh_{f,w} = \int_{h_{m,L}=0}^{h_{m,L}=R} \frac{dW_m}{dh_{m,L}}\, dh_{m,L} - (a+b)$$

The equality on the right-hand side follows from the assumption that both male and female workers are identical except that a female worker takes a paid maternity leave and loses income equal to the area ***a*** + ***b*** in Figure 1.

It is easy to see in Figure 1, and in Equations 9 and 10, that the steeper the slope of the W_m curve, the more costly is the maternity leave of a given length, and the larger the male-female lifetime earning difference—the larger the area ***a*** + ***b*** (Barth, Kerr, and Olivetti 2017; Hotchkiss, Pitts, and Walker 2017, 3509-3522; Lundborg, Plug, and Rasmussen 2017, 1611-1637).

Another point that Figure 1, and Equations 9 and 10, make is that a paid maternity leave is not enough to equate male and female wages (*i.e.*, make the area ***a*** + ***b*** = 0), all else constant. For a female worker who has taken a maternity leave to catch up to their male and female counterparts without intermittencies it must be the case that the slope of her earnings curve be greater as compared with her counterparts after she returns from the maternity leave. (It is the productivity part, $\frac{\partial \alpha(h_{f,L}, h_{f,C})}{\partial h_{f,L}}$, in Equation 8). This is during the period t_p-t_j and it is represented by the curve segment *jp* in Figure 1. A female who catches up to her (male and female) counterparts without intermittency has the earnings curve $kljpW_m$ in Figure 1. Furthermore, in terms of Figure 1, the shorter the period t_p-t_j, the steeper the slope of the curve segment *jp*, and the smaller the area ***b***, all else constant. Studies cited above show, however, that women who bear and rear children pay a financial penalty; they usually do not catch up to their male counterparts and female counterparts without children.

As I state above, the society benefits from women bearing and rearing children; they provide labor force. Women, however, are not compensated for providing the labor input; indeed, they are penalized. In terms of Figure 1 the penalty is equal to the area ***a*** + ***b***. This area also represents the size of the positive externality.

A COMPENSATION MODEL

In this section I present a compensation model that accounts for the positive externalities and discuss its implications. I will also bring back the time subscript while presenting the compensation model.

As before, assume that a certain fraction of females, θ, of the current generation, g_t, bear and rear children. These children become labor force in the next generation, g_{t+1}. To this end females in the current generation take time off from the labor market which leads to a lifetime lost income that is equal to the area ***a*** + ***b*** in Figure 1. It follows then that the compensation should (at least) be equal to the area ***a*** + ***b***. One may write the compensation function of a given female *j* as follows.

$$(\boldsymbol{a} + \boldsymbol{b})_j = \zeta PV(LTE_{m,j,g_{t+1}} + LTE_{f,j,g_{t+1}}) \tag{11}$$

Where 0 ≤ ζ ≤ 1 is the fraction that equates the area ***a*** + ***b*** to the present discounted value (*PV*) of the next generation's lifetime earnings, $LTE_{i,j,g_{t+1}}$, for *i* = male, female, who is (are) the offspring(s) of female *j*. The larger is the area ***a*** + ***b***, the more sacrifice the female *j* has made, and the higher will be the value of *ζ*, all else constant. And the larger the present value of the next generation's earnings, the lower the ζ for a given area ***a*** + ***b***.

Note also that the present discounted value of life-time earnings of an individual of a given generation are a function of the human capital acquired at the start of her/his career (*k* in Figure 1) and the career choice which determines the slope of the earnings curve (Equations 7 and 8), all else constant. The value of *k* of the next generation depends upon the current generation because the current generation provides schooling to the next generation. The value of $PV\left(LTE_{m,j,g_{t+1}} + LTE_{f,j,g_{t+1}}\right)$ also depends, at least partly, upon the current generation. This is because of the opportunities that the current generation has provided to the next generation when the next generation enters the labor force. Examples of these opportunities may include the institutional infrastructure and the physical capital. This fact ties the fates of any two generations together, and the link is the females who bear and rear children.

On first blush, the assumptions made in this paper may seem overly restrictive and the conclusions too strong. Note, however, that the purpose of this model is to provide a formal framework that shows what has been found in empirical research regarding motherhood penalty. That is, it focuses only on the impact of childbearing and rearing on the female's earnings. It shows that even if the labor market rids itself of any other form of gender discrimination, the motherhood penalty remains. It also suggests a way to remedy the motherhood penalty.

CONCLUDING REMARKS

Studies show that there is a financial cost for women who take time off from the labor market for childbearing and rearing. This is called motherhood penalty in the literature. I present a theoretical model in this paper that quantifies the motherhood penalty. Given that women, by bearing and rearing children, provide labor force from which the society benefits, without being compensated, there exist positive externalities. I also show that a paid maternity leave may not be enough to compensate mothers and bring their financial status equal to that of their male counterparts and female counterparts who do not have children. The model presented in this paper also suggests a solution.

The importance of well-trained labor force is essential to any economy's health. Data show that birthrates in developed countries have been declining for decades. For instance, in the US, the birthrate has declined from 23.7 per 1,000 people in 1960, to 11.0 per 1,000 people in 2022.[2] Combine this with an increasing life expectancy due to medical innovations, and you see that there will be fewer workers to support an aging population.

One solution for a declining labor force maybe to increase immigration to the US from-less developed countries that tend to have higher birthrates. This, however, seems to be a politically fraught proposition under the current political environment. Not only that, even in less-developed countries the birthrate has been declining; the birthrate in the "least" developed countries was 47.85 per 1,000 people in 1960, and in 2022 the birthrate was 30.96 per 1,000 people.[3] A more politically palatable solution maybe to increase domestic birth rates. As the literature indicates, however, there is a penalty for females for childbearing and rearing.

That incentives matter is a basic tenet of economics. To increase domestic birthrates, a country must provide incentives (or at the very least, not penalize) for childbearing and rearing. What practical steps maybe taken to remedy the motherhood penalty? A fiscal policy that provides incentives to females for childbearing and rearing, instead of penalizing them, can go a long way.

One avenue maybe to increase the amount of child-tax credit, that already exist in the US tax code, to match the motherhood penalty—area ***a + b*** in Figure 1. Simply tying the child-tax credit to the motherhood penalty—area ***a + b*** in Figure 1—however, may have a negative side effect. As the model shows, well-educated women will have steeper earning profiles and women with less education will have flatter earning profiles, all else constant. This means that women in the former category will receive a larger child-tax credit (the area ***a + b*** is larger), than those in the latter category (the area ***a + b*** is smaller). Since the education of mothers' positively affects the education and wellbeing of their offsprings (Currie and Moretti 2003, 1495-1532), this is may further widen the income gap between the haves and have-nots in each successive generation. Designing a fiscal policy that takes such effects into consideration may take some effort. Future research may focus on these issues.

Another point that the future research may focus is the impact of increased labor market hours on the part of fathers to compensate for the decreased labor market hours on the part of mothers, such that the family income stays the same. As the model shows that this may further increase the gap between mothers' and fathers' lifetime earnings. Will this decreased financial status of mothers further lead to their declining socio-political influence? Empirical evidence is required to answer this question more fully.[4]

ENDNOTES

[1] To my knowledge, an exception is the Edward Lazear and Sherwin Rosen study (Lazear and Rosen 1990, S106-S123). Lazear and Rosen (1990), however, focus on worker and job characteristic matching process.

[2] World Bank, Crude Birth Rate for the United States [SPDYNCBRTINUSA], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/SPDYNCBRTINUSA, August 1, 2024. Note: According to the definition, "Crude birth rate indicates the number of live births occurring during the year, per 1,000 population estimated at midyear. Subtracting the crude death rate from the crude birth rate provides the rate of natural increase, which is equal to the rate of population change in the absence of migration."

[3] World Bank, Crude Birth Rate for Least Developed Countries [SPDYNCBRTINLDC], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/SPDYNCBRTINLDC, August 1, 2024.

[4] I am grateful to the anonymous referee for bringing this point to my attention.

REFERENCES

Adda, Jerome, Christian Dustmann, and Katrien Stevens. 2017. "The Career Costs of Children." *Journal of Political Economy* 125 (2): 293-337.

Aghion, Philippe and Peter Howitt. 1998. *Endogenous Growth Theory*. Cambridge, MA: The MIT Press.

Angelov, Nikolay, Per Johansson, and Erica Lindahl. 2016. "Parenthood and the Gender Gap in Pay." *Journal of Labor Economics* 34 (3): 545-579.

Barro, Robert J. 1991. "Economic Growth in a Cross Section of Countries." *The Quarterly Journal of Economics* 106 (2): 407-443.

Barth, Erling, Sari Pekkala Kerr, and Claudia Olivetti. 2017. "The Dynamics of Gender Earnings Differentials: Evidence from Establishment Data." *NBER Working Paper Series* Working Paper No. 23381.

Benatar, David. 2017. "The Human Predicament : A Candid Guide to Life's Biggest Questions." . doi:10.1093/acprof:oso/9780190633813.001.0001.

Blau, Francine D. and Lawrence M. Kahn. 2013. "Female Labor Supply: Why is the United States Falling Behind?" *American Economic Review* 103 (3): 251-256.

———. 2000. "Gender Differences in Pay." *Journal of Economic Perspectives* 14 (4): 75-99.

———. 2017. "The Gender Wage Gap: Extent, Trends, and Explanations." *Journal of Economic Literature* 55 (3): 789-865.

Browning, Martin. 1992. "Children and Household Economic Behavior." *Journal of Economic Literature* 30 (3): 1434-1475.

Burbano, Vanessa, Nicolas Padilla, and Stephan Meier. 2024. "Gender Differences in Preferences for Meaning at Work." *American Economic Journal. Economic Policy* 16 (3): 61-94. doi:10.1257/pol.20220121.

Cain, Mead. 1981. *Risk and Insurance: Perspectives on Fertility and Agrarian Change in India and Bangladesh*. Vol. 7.

Cravey, Tiffany and Aparna Mitra. 2011. "Demographics of the Sandwich Generation by Race and Ethnicity in the United States." *Journal of Socio-Economics* 40 (3): 306-311.

Currie, Janet and Enrico Moretti. 2003. "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings." *The Quarterly Journal of Economics* 118 (4): 1495-1532. doi:10.1162/003355303322552856.

Ferber, Marianne A. and Helen M. Lowry. 1976. "The Sex Differential in Earnings: A Reappraisal." *Industrial and Labor Relations Review* 29 (3): 377-387.

Ferber, Marianne A. and Joe L. Spaeth. 1984. "Work Characteristics and the Male-Female Earnings Gap." *American Economic Review* 74 (2): 260-264.

Fernandez, Raquel. 2014. "Women's Rights and Development." *Journal of Economic Growth* 19 (1): 37-80.

Goldin, Claudia. 2014. "A Grand Gender Convergence: Its Last Chapter." *The American Economic Review* 104 (4): 1091-1119.

Goldin, Claudia and Lawrence F. Katz. 2016. "A most Egalitarian Profession: Pharmacy and the Evolution of a Family-Friendly Occupation." *Journal of Labor Economics* 34 (3): 705-746.

———. 2008. "Transitions: Career and Family Life Cycles of the Educational Elite." *American Economic Review: Papers and Proceedings* 98 (2): 363-369.

Hotchkiss, Julie L. and M. Melinda Pitts. 2007. "The Role of Labor Market Intermittency in Explaining Gender Wage Differentials." *The American Economic Review Papers and Proceedings* 97 (2): 417-421.

Hotchkiss, Julie L., M. Melinda Pitts, and Mary Beth Walker. 2017. "Impact of First Birth Career Interruption on Earnings: Evidence from Administrative Data." *Applied Economics* 49 (35): 3509-3522.

Lazear, Edward P. and Sherwin Rosen. 1990. "Male-Female Wage Differentials in Job Ladders." *Journal of Labor Economics* 8 (1): S106-S123.

Lundborg, Peter, Erik Plug, and Astrid W. Rasmussen. 2017. "Can Women have Children and a Career? IV Evidence from IVF Treatments." *American Economic Review* 107 (6): 1611-1637.

Mankiw, N. Gregory, David Romer, and David N. Weil. 1992. "A Contribution to the Empirics of Economic Growth." *The Quarterly Journal of Economics* 107 (2): 407-437.

Pelkowski, Jodi Messer. 2005. "Adult Children's Propensity to Care for an Elderly Parent: Does the Marital Status of the Parent Matter?" *Journal of Economics (MVEA)* 31 (1): 17-39.

Pörtner, Claus Chr. 2001. "Children as Insurance." *Journal of Population Economics* 14 (1): 119-136. doi:10.1007/s001480050162.

Romer, Paul M. 1990. "Endogenous Technological Change." *The Journal of Political Economy* 98 (5, Part 2: The Problem of Development: A Conference of the Institute for the Study of Free Enterprise Systems): S71-S102.

Solow, Robert M. 1956. "A Contribution to the Theory of Economic Growth." *The Quarterly Journal of Economics* 70 (1): 65-94.

Waldfogel, Jane. 1998. "Understanding the "Family Gap" in Pay for Women with Children." *The Journal of Economic Perspectives* 12 (1): 137-156.

Wolf, Douglas A. and Beth J. Soldo. 1994. "Married Women's Allocation of Time to Employment and Care of Elderly Parents." *Journal of Human Resources* 29 (4): 1259-1276.

REFEREE ACKNOWLEDGMENT

The New York State Economics Association is grateful for the time spent by the following referees in recent years. Without your efforts, we would not be able to produce this journal.

1. Faisal Abbas
2. Bebonchu Atems
3. James T. Bang
4. Supravat Bagli
5. Prasun Bhattacharjee
6. Joseph Cheng
7. Brian Chezum
8. Christian Conrad
9. Jeffrey Cross
10. Michael Davis
11. Anupam Das
12. Prithviraj Guha
13. Joanne Guo
14. Rupayan Gupta
15. L. Chukwudi Ikwueze
16. Michael Jerison
17. Gisella Kagy
18. Joshua Mask
19. Michael McAvoy
20. Michael O'Hara
21. Kasia Platt
22. Mitali Pradhan
23. Jungjun Park
24. Jonathan Stone
25. Xuguang Sheng
26. Elias Shukralla
27. Della Lee Sue
28. Masashi Tsutsumi
29. Hang Yi
30. Ling Yan
31. Jenny Zhao
32. Xu Zhang