

LEIBNIZ UNIVERSITÄT HANNOVER
WIRTSCHAFTSWISSENSCHAFTLICHE FAKULTÄT
INSTITUT FÜR WIRTSCHAFTSINFORMATIK


Untersuchung der preisbildenden Parameter von Dachflächen-Photovoltaiksystemen mit Data Analytics Methoden

MASTERARBEIT

zur Erlangung des akademischen Grades „Master of Science (M. Sc.)“ im Studiengang Wirtschaftsingenieur der
Fakultät für Elektrotechnik und Informatik, Fakultät für Maschinenbau und der Wirtschaftswissenschaftlichen
Fakultät der Leibniz Universität Hannover

vorgelegt von

Robert Mark Caspari



Prüfer:
Prof. Dr. Michael H. Breitner

Betreuer:
M. Sc. Tobias Kraschewski

Hannover, den 22. April 2022

Abstract

Rising energy demand in growing economies collides with the challenges of climate change. Therefore a cost-efficient and widespread adoption of renewable energy sources has been the goal of many policies in Germany in recent years. This research aims to explain cost differences in residential Photovoltaic-Systems to increase market transparency and reduce initial costs. The scope of this work is limited to the period from 2011 to 2021 and price analysis of offers, not operation and maintenance costs. Lastly, only residential PV-Systems with less than 30 kWp are considered. Relevant offer data from an online forum, including PV-System characteristics, approximate location, and price, is web-scraped and combined with time and location variant statistics. These statistics include among other factors, inflation, cost of living, and a count of nearby installers. Together this amounts to a dataset with 20.000 observations and 40 parameters. A purely descriptive analysis of cost trends indicates a decrease in total and component prices from 2011 until 2019. This trend comes to a halt in 2020 and 2021 when costs increase slightly. The data is also modeled and evaluated using OLS, Elastic Net, and General Additive Regression. OLS Regression provides the best results regarding interpretability and performance with a normalized RMSE of 14 %. Across most models, the following parameters exhibit a significant price increasing effect: Module-efficiency, battery-size, population, wage level, region size, energy prices, feed-in tariffs, the inclusion of frames, house prices, and the prices of individual components. However for the following parameters, a price decreasing effect is indicated: Scale, count of installed PV-Systems in the region, power per module, count of nearby installers, the share of equity and supporting the installer. Seasonality is also indicated in some models, although the resulting effects vary. Further recommendations include the design of a holistic online marketplace, which also serves as a future source of anonymized data available to the public. A focus on increased data quality compared to existing platforms is strongly advised.

Keywords: Photovoltaic System, residential PV, Cost, Analysis, Regression, General Additive Models, German Market, Consumer Cost, PV Adoption Challenges

Inhaltsverzeichnis

Abstract	I
Abbildungsverzeichnis	IV
Tabellenverzeichnis	VI
1 Einleitung	1
2 Status Quo und theoretische Grundlagen	2
2.1 Untersuchungsobjekt	2
2.2 Metriken	3
2.3 Photovoltaik Preisforschung Status Quo	7
2.3.1 Status Quo	8
2.3.2 Variablen	17
2.3.3 Hypothesenentwicklung	19
2.4 Statistische Grundlagen	23
2.4.1 Ausreißer und fehlende Werte	23
2.4.2 Korrelation und Modellierung	27
2.4.3 Modellevaluation	30
3 Entwicklung des Datenmodells zu den PV-Angebotspreisen	33
3.1 Datengrundlage	34
3.1.1 Datengewinnung	34
3.1.2 Datenverknüpfung und Feature Engineering	37
3.1.3 Datenaufbereitung	39
3.2 Finales Datenmodell	43
4 Statistische Analyse der PV-Angebotspreise	47
4.1 Datenauswahl und darauf aufbauende Modelle	48
4.1.1 Evaluation linearer Modelle	50
4.1.2 Evaluation additiver Modelle	55
4.2 Zusammenfassung Modellierung	57
4.3 Deskriptive Analyse	58
5 Diskussion der preisbildenden Parameter	62

5.1	Test der Hypothesen	62
5.2	Theoretischer Beitrag	67
5.3	Praktische Implikationen	69
5.4	Limitationen	71
6	Fazit und Ausblick	74
	Literatur	76
A	Anhang	84
A.1	Komplette Auflistung des prozentual fehlenden Anteils der Größen. . .	84
A.2	OLS, ElasticNet und GAM Modellierungsparameter des gesamten Mo- dells für die Median/Mode Imputation.	87
A.3	Python-Implementierung des Datenverarbeitungsprozesses	97
B	Research Summary	118
B.1	Einleitung	118
B.2	Untersuchungsobjekt	118
B.3	Sekundäranalyse	119
B.4	Daten	120
B.5	Statistische Analyse	121
B.6	Diskussion	123
B.7	Fazit und Ausblick	126
	Ehrenwörtliche Erklärung	128

1 Einleitung

Weltweit kann durch einen Wachstumstrend sowohl in der Bevölkerung als auch in der Wirtschaft ein steigender Energiebedarf beobachtet werden. Diesem Bedarf steht jedoch zugleich die Herausforderung des Klimawandels gegenüber. Um dieser Herausforderung zu begegnen ist es die Aufgabe der Gesellschaft, Politik und Wirtschaft geeignete Maßnahmen zu erarbeiten, die auch zukünftig den Wohlstand nachhaltig sicherstellen und vergrößern können. Die Bundesregierung hat sich aus diesem Anlass 2021 das Ziel gesetzt "[...] unser nationales Minderungsziel für 2030 um zehn Prozentpunkte auf 65 Prozent anheben und zugleich bereits 2045 Klimaneutralität anstreben." [1]. Im Fraunhofer-Institut wird zur Erreichung dieses Ziels die Erhöhung erneuerbarer Energien am Energiemix, und damit die Adaptionrate von PV-Anlagen, als fundamentaler Faktor identifiziert [2].

Um diesen Ansprüchen gerecht zu werden, sind die ökonomische Effizienz und Effektivität von PV-Anlagen ein wesentlicher Treiber. Hierbei sind insbesondere die Anschaffungskosten elementar. Allerdings besteht, insbesondere bei Privathaushalten und Dachflächen PV-Anlagen, eine signifikante Intransparenz in der Preisbildung einer PV-Anlage. Diese Intransparenz hemmt die PV-Adaption durch eine Erhöhung der Suchkosten bei Angebotsinteressierten und möglicherweise auch die Anschaffungskosten selbst. Damit ist eine Erklärung und Untersuchung der preisbildenden Parameter bedeutsam, um die PV-Adaptionrate zur Erreichung der Klimaziele zu erhöhen. Um diese Thematik aufzugreifen, muss die Forschungsfrage somit lauten:

Welche Parameter erklären die Varianz der Anschaffungskosten von Dachflächen Photovoltaik Anlagen für Privathaushalte?

Mit dem Ziel, diese Frage fundiert zu beantworten, wurde in dieser Arbeit eine Sekundäranalyse bestehender Forschung in Kap. 2 ausgeführt. Im Anschluss daran werden Hypothesen, über die preisbildenden Parameter, formuliert. Diese Hypothesen werden zusammen mit einem dafür in Kap. 3 entwickelten Datensatz und auf diesem in Kap. 4 aufbauender statistischer Modellierungen geprüft. Abschließend werden die gewonnenen Ergebnisse mit ihren theoretischen und praktischen Limitationen als Basis für weitere Forschung in Kap. 5 reflektierend diskutiert.

6 Fazit und Ausblick

Die Motivation dieser Arbeit war es den Herausforderungen des Klimawandels und der Energieunabhängigkeit zu begegnen. Dazu stellt die Installation von Dachflächen PV-Anlagen in Privathaushalten einen Teil eines möglichen Lösungsansatzes dar. Dabei sind in der ökonomischen Perspektive die Anschaffungskosten für den Haushalt elementar. Gleichwohl wurde für Dachflächen PV-Anlagen bei Privathaushalten eine Forschungslücke identifiziert. Daher wurde das Schließen der Lücke mit der Forschungsfrage:

Welche Parameter erklären die Varianz der Anschaffungskosten von Dachflächen Photovoltaik Anlagen für Privathaushalte?

als Ziel dieser Arbeit definiert. Zur Erklärung der Varianz des Anschaffungspreises folgte der Ermittlung, in Kap. 2 und 3, die Untersuchung der preisbildenden Parameter in Kap. 4 und 5. Für die Untersuchung wurde ein neuer Datensatz mit ca. 20.000 Beobachtungen und über 40 Parametern für den Zeitraum von 2011 bis 2021 entwickelt. Auf dieser Basis wurden sowohl preisbildende Parameter identifiziert, als auch eine Modellierung zur Vorhersage formuliert. Für wenige Parameter konnte jedoch kein eindeutiges Urteil gebildet werden. Die Modellierung schätzt den Preis ebenfalls lediglich näherungsweise. So liegt die Abweichung der besten Modellierung, anhand des mit 5-facher Kreuz-Validierung bestimmten normierten RMSE, bei 14 %. Als mögliche Ursache, wurde die teilweise mangelnde Datenqualität eines ausgewerteten Forums, sowie das Fehlen präziserer statistischer Daten, identifiziert. Aufgrund dieser Limitation konnte die Forschungslücke nur zum Teil geschlossen werden. Dennoch konnten mit der durchgeführten Sekundäranalyse, in Verbindung mit den eigenen Untersuchungen, praktische Maßnahmen in Kap. 5 entwickelt werden, welche zukünftig die Markttransparenz weiter erhöhen und damit die Varianz der Anschaffungskosten noch erklärbarer gestalten würden.

Die Erkenntnisse dieser Arbeit können für die einzelnen Parameter festgehalten werden. Die Größen Moduleffizienz, Batteriekapazität, Einwohnerzahl, Lohnniveau, Fläche der Region, Netzpreis, Einspeisevergütung, Gerüst, Hauspreisindex, sowie die Kosten der individuellen Komponenten, zeigen in der Mehrheit der Modell und Modellierungs Kombinationen einen preissteigernden Effekt. Zugleich deuten die Größen PV-Anlagen Dichte, Anlagengröße, Modulleistung, Solarteurdichte, Eigenkapitalquote und Mithilfe

einen preissenkenden Effekt an. Für die weiteren untersuchten Parameter war es entweder nicht möglich eine eindeutige Wirkung zu bestimmen oder diese war insignifikant.

Somit wird abschließend für das weitere Vorgehen empfohlen, die gewonnenen Erkenntnisse den Marktakteuren noch zugänglicher bereitzustellen. Dies könnte durch das Design eines neuen Marktplatzes gewährleistet werden. Dieser Marktplatz kann gleichzeitig als Quelle anonymisierter Daten die zukünftige Forschung unterstützen. Damit würden ebenfalls die Limitationen, durch die eingeschränkte Verfügbarkeit von Daten, aufgegriffen werden. Zukünftig könnten die Erkenntnisse dieser Arbeit hiermit noch präziser gestaltet werden. Zusammen mit diesem Baustein, wird der Markt für PV-Anlagen effizienter, senkt finanzielle Hürden und begegnet den gesellschaftlichen Herausforderungen des Klimawandels in der ökonomischen Dimension.