Parametrization for the BAföG OCEL Simulation

Business Process Management (WS 25/26)
Group 10

Alina Thoden
Haya Edris
Mansour Dukhan
David Derr
Büsra Bugrahan
Jan Gödicke

January 2, 2026

1 Process Context

1.1 Scenario and Data Basis

The subject of the simulation is the influx of BAföG applications for Hamburg for the winter semester 2024/2025. The peak load is characterized by the start of the semester, with about 63 % of applications arriving around the start of the winter term (Studentenwerk SH, 2024). The data basis for Hamburg is 15,564 funded students in the year 2024 (Destatis, 2024).

1.2 Process Focus

The focus is on the processing of initial applications that are submitted exclusively digitally via BAföG-Digital. The scenario thus considers only applications via BAföG-Digital and excludes paper or mixed channels as well as follow-up applications, student and study-abroad BAföG. A separation between domestic and study-abroad BAföG is possible, as personnel capacity is divided between them (Bundesregierung, 2012).

The start point is "Application started", i.e., the time of data entry into the system. Preparatory work by the student is not considered, as it does not influence the office throughput time.

1.3 Volume Assumptions

We simulate a High-Load Scenario, where the entire application volume (9,800 at the start of the winter semester) is treated as initial applications to test the system's resilience under maximum complexity.

1.4 Modeling Approach

The current BPMN Figure 2 is more data-driven than the initial version Figure 1. The current version concentrates on activities that would cause changes in the databases, such as CRUD operations.
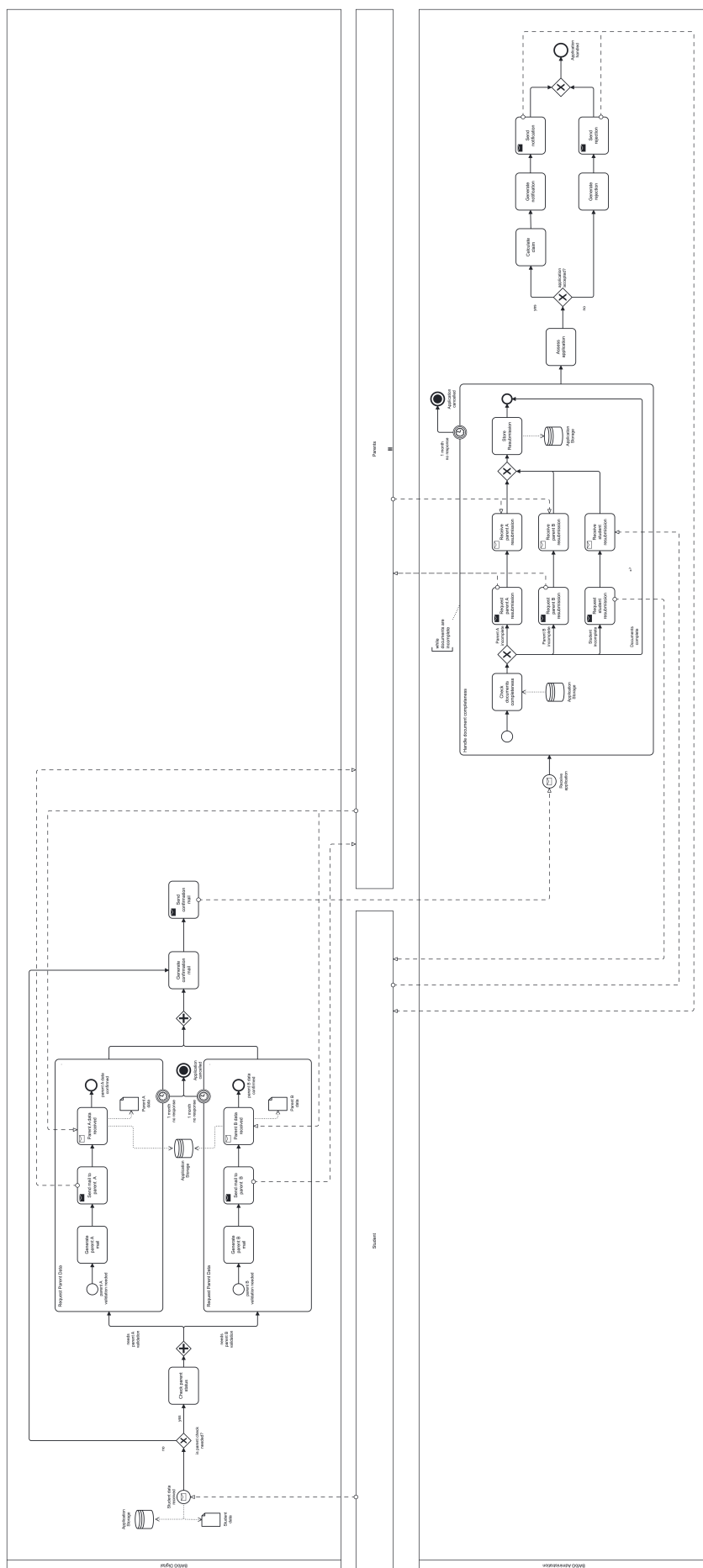
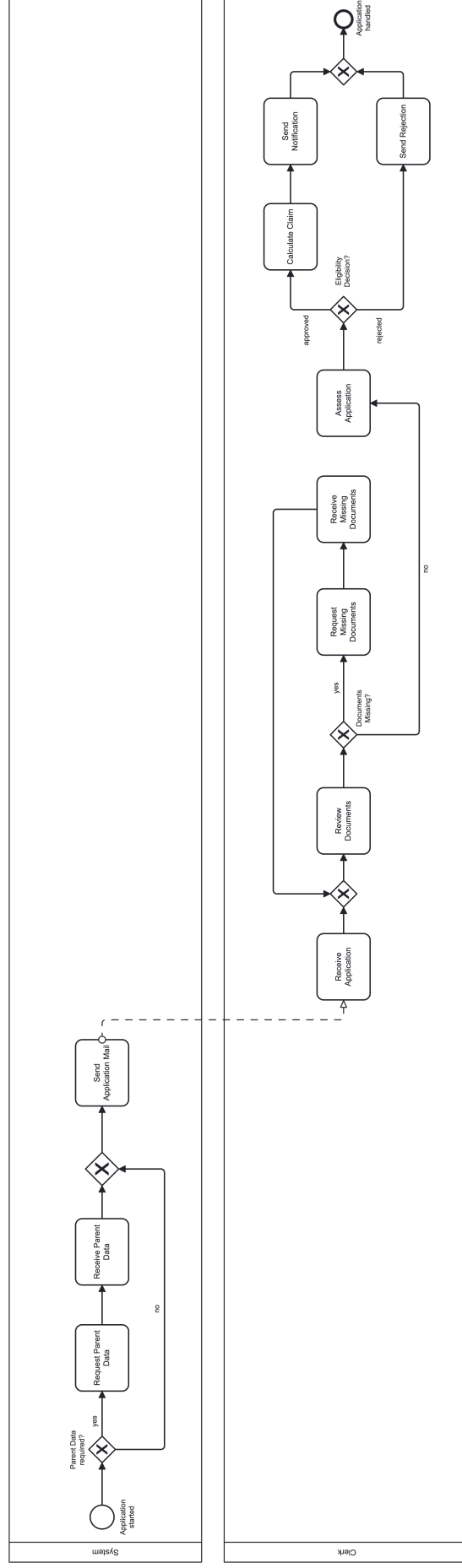Figure 1: Initial Version: BPMN of the BAföG Process

Figure 2: Current BPMN of the BAföG Process

3

1.5 Scenario Setup & Output

- **Timeframe:** The simulation starts on **15.09.2024** and runs for a period of **80 days** to cover the entire winter semester including follow-up time.

- **Reproducibility:** To ensure reproducible results, a fixed **Random Seed (42)** is used.

- **Output:** The simulation generates relational **CSV tables** (Events, Applications, Documents, Object-Links) that correspond in their structure to an Object-Centric Event Log (OCEL). This allows for direct analysis in process mining tools that support object-centric data.

2 Interarrival

The load peaks are based on the start of the winter semester. According to (Studentenwerk SH, 2024), about 63 % of applications are received around the start of the winter term. Applied to Hamburg, this corresponds to about 9,800 of 15,564 funded students (Destatis, 2024). We assume the following interarrival times:

Time Window	Weekdays	Distribution	Parameter	Assumption
08:00–16:00	Mon–Fri	Exponential	Mean 1.5 min	Peak: Main business hours (High Load Scenario)
16:00–21:00	Mon–Fri	Exponential	Mean 2.7 min	Peak: After lectures (High Load Scenario)
21:00–24:00	Mon–Fri	Exponential	Mean 8.1 min	Late submission, moderate density
All day	Sat–Sun	Exponential	Mean 10.8 min	Weekend, reduced rate

Table 1: Interarrival times for the simulation

3 Gateways

3.1 Parent Data Required?

According to Statistics Bavaria 2022 (Statistik Bayern, 2022), about 18% of funded students are parent-independent (13,637 of 74,771). From this we derive that in 80% of cases parent documents are required, while 20% are funded independently of parents. This controls the attribute `Application.is_parent_independent` (FALSE = Parent documents required).

3.2 Documents Missing?

According to technical literature, only 1–2% of paper initial applications are complete (Bundesregierung, 2010), while for continued funding about 35% are submitted completely (Studentenwerk SH, 2024). Since the focus is exclusively on initial applications, the risk of incomplete documents is higher (approx. 98% for paper). Through the use of the digital assistant BAföG Digital, an improvement is assumed, which is why we set the risk "Documents Missing" to 40% and "Complete" to 60%. This gateway is controlled by the attribute `Document.status` ("Missing" vs. "Received").

3.3 Eligibility Decision?

According to (Bundesregierung, 2012), the share of unapproved applications is 16%. The path probabilities are thus 84% for approved and 16% for rejected. This gateway is controlled by the attribute `Application.status` ("Approved" vs. "Rejected").

4 Activity Durations

To derive the activity durations, the standard times from Table 22 ("Average standard times of application processing in the Student Union Hamburg for domestic funding") (Bundesregierung, 2012) were used. For the initial application scenario, this results in a total of 83 minutes. The shares of the individual activities were determined consistently with the standard times given in Table 22 and distributed to the activity groups listed below:

- **Receive Application:** corresponds to creation of the paper file $\approx 13\,\text{min}$

- **Review:** corresponds to completeness check $\approx 13\,\text{min}$.

- **Request Missing Data** corresponds to obtaining missing data or information $\approx 12\,\text{min}$.

- **Assess:** corresponds to half of the time of performing calculations/assessments $\approx 15\,\text{min}$.

- **Calculate:** corresponds to half of the time of performing calculations/assessments incl. checking and correcting results $\approx 20\,\text{min}$.

- **Notification/Rejection:** corresponds to preparation, sending $\approx 10\,\text{min}$.

The activity durations of the casework were transferred into stochastic distributions for parameterization. Manual tasks are modeled by normal distributions (with plausible standard deviations σ as well as specified minimum and maximum values), system-side activities by uniform distributions and waiting times by exponential distributions.

Activity	Resource	Distribution	Parameter	Justification
Application started	System	-	0 min	Start-Event (Data Entry)
Request Parent Data	System	Uniform	0.5–2 min	Automated mail to parents
Receive Parent Data	External Actor	Normal	$\mu = 10080, \sigma = 3600$	Waiting time (Assumption: Normal distribution)
Send Application Mail	System	Uniform	1–3 min	Automatic generation + sending
Receive Application	Clerk	Normal	$\mu = 13, \sigma = 4$	Creation of paper file
Review Documents	Clerk	Normal	$\mu = 5, \sigma = 2$	Time per document (Assumption sim_config)
Request Missing Documents	Clerk	Normal	$\mu = 12, \sigma = 3$	Obtain missing data
Receive Missing Documents	External Actor	Normal	$\mu = 10080, \sigma = 2880$	Waiting time (Assumption: Normal distribution)
Print Documents	Clerk	Uniform	2–5 min	Media disruption: Batch printing (Assumption: Deviation)
Assess Application	Clerk	Normal	$\mu = 15, \sigma = 4$	Assessment
Calculate Claim	Clerk	Normal	$\mu = 20, \sigma = 5$	Calculation
Send Notification	Clerk	Normal	$\mu = 10, \sigma = 3$	Sending notification
Send Rejection	Clerk	Normal	$\mu = 10, \sigma = 3$	Sending rejection
Archive Documents	Clerk	Uniform	2–5 min	Archiving (Assumption: Deviation)

Activity	Resource	Distribution	Parameter	Justification
Application handled	System	-	0 min	End-Event (Technical completion)

Table 2: Distribution of activity durations

4.1 Document Types and Complexity Factors

For a detailed simulation, the processing times are scaled based on the document type. The factors are based on assumptions.

Calculation of duration for "Review Documents":
The duration of this activity is calculated dynamically based on the sum of the complexity factors of all documents to be checked (batch processing):

$$\text{Duration} = \text{Base Duration}(\approx 5\,\text{min}) \times \sum(\text{Complexity Factors})$$

(See logic in `sim_ocel.ipynb`).

Document	Condition	Complexity Factor
Formblatt 1 (Application)	always	1.0
Enrollment Certificate	always	0.5
Formblatt 3 (Parent Income)	if `Application.is_parent_independent = FALSE`	1.5
Income Proof (Parents)	if Formblatt 3 present (usually 2 pieces)	1.3
Rent Certificate	if `Application.housing_type` $\neq$ `'Eltern'`	0.8

Table 3: Document types and complexity factors

5 Object-Centric Data & Logic

The simulation model is based on an object-centric approach (Object-Centric Process Mining), which goes beyond the classic Event Log standard.

5.1 Object Types

- **Application:** The central Case object that bundles the entire application process. Attributes: `application_id`, `student_id`, `is_parent_independent`, `housing_type`, `status`, `deviation_type`.

- **Document:** Secondary objects with their own lifecycle, which are linked to the application. Attributes: `document_id`, `doc_type`, `status` (Missing/Received), `submission_time`.

- **Event:** Links activities with one or more objects. Attributes: `activity`, `timestamp`, `org_resource`, `linked_objects`.

5.2 Document Generation and Lifecycle

Before the activity "Application started", all potential documents (e.g. Formblatt 1, Enrollment Certificate) are technically already instantiated. On Application started, all documents marked with `system_mandatory` in the config are set to `Received`. Furthermore, based on the gateway probabilities, the type of required documents is determined and additionally marked as `Missing` with a probability of `40%`.

- **Initial Status:** All documents start in the status `Missing`.

- **Status Change:** The status only changes to `Received` when the corresponding activity (e.g. "Application started", "Receive Application", "Receive Parent Data", "Review Documents") is actually executed.

This allows for the separate tracking of each individual document (incompleteness) parallel to the main application.

Furthermore, in Review Documents there is a parameter `p_invalid` with a probability of `10%` that a document is marked as `Missing` to simulate incomplete documents. As soon as these are successfully reviewed, these documents are no longer subjected to any check.

5.3 Object-to-Event Mapping

To depict the lifecycle of the application and the individual documents separately from each other, but synchronized, the following mapping scheme is applied:

- **Leading Object:** The object `Application` is linked with **all** activities to ensure the continuous process flow.

- **Secondary Object:** The object `Document` is linked **exclusively** with activities that represent a physical processing or state change of a document (e.g. *Request, Receive, Review*).

This allows the analysis of 1:n relationships in Process Mining (one application has n documents that behave differently).

Activity	Linked Objects	Justification
Application started	Application	Initialization of the Case
Request Parent Data	Application, **Document**	Generation/Request of Parent Objects
Receive Parent Data	Application, **Document**	Receipt of Parent Documents
Send Application Mail	Application	System notification to the BAföG office Hamburg about application receipt
Receive Application	Application, **Document**	Receipt of Basic Documents
Review Documents	Application, **Document**	Content check per single document
Request Missing Documents	Application, **Document**	Requesting missing documents
Receive Missing Documents	Application, **Document**	Receipt of submitted documents
Assess Application	Application	Check of the entire application
Calculate Claim	Application	Calculation of the BAföG rate
Send Notification	Application	Positive notification (Approval)
Send Rejection	Application	Negative notification (Rejection)

Activity	Linked Objects	Justification
Archive Documents	Application, **Document**	Archiving of all documents
Print Documents	Application, **Document**	Printing of all documents
Application handled	Application	Technical completion of the Case

Table 4: Object-to-Event Mapping

5.4 Synchronization Logic and Document Categories

To reduce model complexity, parent documents are not split into individual roles (mother/father), but treated as an aggregated proof object. The OCPM logic applies here through the synchronization of different document categories:

1. **Differentiated waiting times:** The document of the category "Parent" (e.g. Income Proof) receives a significantly longer waiting time in the simulation (exponential distribution, $\mu \approx 7$ days) than documents of the category "Student" ($\mu \approx 2$ days).

2. **Synchronization at the Gateway:** The process step "Assess Application" acts as a synchronization point. It may only start when **all** linked documents (both Student and Parent) have reached the status "Received".

This demonstrates the Core Benefit of Object-Centric Process Mining: The analysis shows that the overall process is often blocked by a single, complex object (here: Parent Proof), while other objects (e.g. Enrollment Certificate) are already present (Waiting Time Paradox).

6 Resources

6.1 System

- Available 24/7, high capacity (9999)

- Availability: 00:00–23:59, all days.

6.2 Clerk (Caseworker)

- Number: Domestic funding with 10 clerks, in reality 32(Bundesregierung, 2012), however 10 to simulate queues.

- Availability: Mon–Fri 07:30–16:00.

- **Queuing:** Tasks that arrive outside service hours (e.g. through online application in the evening) are collected in a queue (backlog) and only processed at the beginning of the next shift (07:30 am). This explains waiting times despite calculably sufficient capacity.

7 Planned Deviations

To demonstrate a Conformance Check, deviations are deliberately generated in the simulation data that deviate from the standard process (Happy Path).

7.1 Switched Activities

Scenario: In **1 %** of cases, the step "Assess Application" is swapped with "Calculate Claim" in the sequence (Simulation Parameter `activity_switch`).

Meaning: This simulates a deviation in the process sequence. In the Conformance Check, this must be detected as a **Sequence Violation**.

7.2 Direct Rejection (Shortened Path)

Scenario: In **3 %** of cases, the process aborts immediately with a rejection after "Review Documents" (Simulation Parameter `direct_rejection`).

Meaning: Simulates immediate rejection (e.g. in case of obvious lack of jurisdiction). Recognizable as **Skipped Activities**.

7.3 Blind Approval (Compliance Violation)

Scenario: In **2 %**, the application is approved directly without check ("Assess") (Simulation Parameter `blind_approval`).

Meaning: A serious compliance violation (approval without check).

7.4 document_invalid (Rework)

Scenario: In **20 %** of cases, a document is rejected as invalid or incorrect during the review (Simulation Parameter `document_invalid`).

Meaning: This leads to a further review round and increases the cycle time massively.

7.5 Rejection due to Timeout

Scenario: In **4 %**, the process jumps directly to rejection after a waiting time of >30 days.

Meaning: Automatic rejection in case of inactivity.

7.6 Unmapped Activities (Model Moves)

Scenario: The activities "Print Documents" and "Archive Documents" are performed by clerks, but are **not included in the target process (BPMN)**.

Meaning: These steps appear as **Log Moves** (in the log, but not in the model) in Conformance Checking and must be filtered or accepted as a deviation (Assumption: Media disruptions in reality).

8 Simulation Recommendations

- KPIs: Cycle Time per application, utilization of clerks, waiting times per queue, approval rate.

- Experiments: Sensitivity to document completeness (Gateway), variation Receive-Document mean, personnel capacity (e.g. 10–16 FTE), seasonal peaks vs. normal semester.

9   Limitations

The simulation is subject to the following limitations:

- **Focus on initial applications:** The simulation depicts exclusively initial applications. Follow-up applications, student BAföG and study-abroad BAföG are not considered.

- **Exclusively digital applications:** Only applications via BAföG-Digital are simulated. Paper applications and hybrid submission channels are excluded.

- **Aggregated parent documents:** Parent documents (Income Mother/Father) are not distinguished individually, but treated as an aggregated proof object.

- **Simplified document logic:** The probability that a document is initially missing is fixed at 40 %. In reality, this varies depending on document type and applicant profile.

- **No resource outages:** Sickness, leave or training of clerks are not modeled. Capacity remains constant at 10 FTE.

- **Fixed review rounds:** The maximum number of review rounds is limited to 10 (`max_review_rounds`). Endless request loops are prevented.

- **No seasonal variations within the semester:** The interarrival rates are exponentially distributed over the entire simulation period and do not depict intra-seasonal fluctuations (e.g. exam phases).

- **No application time windows 00:00–08:00 weekdays:** In the current configuration, an explicit interarrival time window for the night (00:00–08:00 am) on weekdays is missing; these periods are treated with a default value.

- **Static rejection rate:** The approval rate of 84 % is constant and does not take into account variation by applicant profile or document quality.

References

Bundesrechnungshof (2024): *Bericht zum BAföG – Volltext.* Textziffern (Tz.) 2.2.1, 3.2, 9.1, 9.2 u. a. URL: `https://www.bundesrechnungshof.de/SharedDocs/Downloads/DE/Berichte/2024/bafoeg-volltext.pdf?__blob=publicationFile&v=2`

Bundesregierung; Nationale Normenkontrollrat (2010): *Einfacher zum Studierenden-BAföG – Abschlussbericht.* URL: `https://www.bundesregierung.de/resource/blob/2065474/396444/d606b826f7415afb0c091c30e968e383/2010-03-17-abschlussbericht-einfacher-zum-bafoeg-data.pdf?download=1`

Bundesregierung (2012): *Projektbericht zur BAföG-Verwaltung – Bewilligungsquoten Inland.* Tabellen 22 und 23. URL: `https://www.bundesregierung.de/resource/blob/974430/455514/00c68b52fe5801d0f59643345aa942a2/2012-06-22-projektbericht-7-data.pdf?download=1`

Studentenwerk Schleswig-Holstein (2024): *Warum dauert die BAföG-Bearbeitung länger? – FAQ.* URL: `https://studentenwerk.sh/de/bafoeg-warum-die-antragstellung-laenger-dauert`

Statistisches Bundesamt (Destatis) (2024): *BAföG-Geförderte nach Ländern.* GENESIS-Tabelle 21411-0020. URL: `https://www-genesis.destatis.de/datenbank/online/statistic/21411/table/21411-0020`

Bayerisches Landesamt für Statistik (2022): *BAföG-Geförderte – Anteil elternunabhängiger Förderung.* Statistische Berichte, Kennziffer K9100C. URL: `https://www.statistik.bayern.de/mam/produkte/veroffentlichungen/statistische_berichte/k9100c_202200.pdf`