Mining Travel Logs to Understand the Benefits of Following the Rules

Eleuterio Ramirez, Jose Luis Haddad, Maximiliano Stuardo,
Juan Jose Martinez, Katherine Vergara, Benjamin Rivera, and Victor Galvez

Department of Computer Science
School of Engineering
Pontificia Universidad Catolica de Chile
{elramirez, jlhaddad, mestuardo,
jjmartinez, kyvergara, birivera, vagalvez}@uc.cl

Abstract. This report focuses on the divergences that exist between the process established by the university and the actual flow that users follow regarding the submission of permits for international trips with academic purposes. The process has three main stages: first, obtaining a travel permit, then the trip itself and lastly, the travel declaration for reimbursement of expenses.
The main focus of this report was on understanding the correlation between the throughput time of the process flow and the compliance of the established process steps, understanding the process of obtaining a travel permit itself and finally, analyzing the international trips that were carried out without the required permits to understand where deviations occurs in this part of the process related to the filling of a travel declaration form. The analyses showed that following the guidelines provided by the institution result in permits and declarations approved in shorter periods of time, however, there is a 23% of the cases where guidelines were not followed. Furthermore, there is a 7% of cases where reimbursement was made even though the travel permit was not approved.

Keywords: Process Mining · Workflow Management · Behaviour Analysis.

1 Introduction and Business Questions

In a world ruled by data, it is possible for every organization to retrieve information from their processes and activities, and also constantly analyzing such data with the goal of validating how the processes of the organization are conceived and updating them as needed, based on how they are being executed. Following this idea, the present work provides three interesting questions that guide an analysis that focuses on understanding the divergences between the process model established by Eindhoven University of Technology (TU/e) for their employees to declare international trips and the path that the employees took to achieve that exact same goal.

The process itself consists of three parts, the first one is a permit request where and employee is authorized to go on a trip, the second is the trip itself and the third is a declaration of expenses so that the employee can be reimbursed after his trip ends. The following questions will help us better understand how this process develops and will guide our analysis:

- **Question 1:** Knowing that not all of the employees follow the same steps when executing the process, *do employees that stick to the process model have any advantage in terms of throughput time and ease of flow?*
  The absence of any type of advantage for those employees who stick to the original process established by the institution and to the institution itself, may be an indicator for the need of changing the way in which the permit and travel declarations are managed.
- **Question 2:** *How different is the process flow projected by the university to obtain a travel permission, from the process flow obtained from the data log?*
  Following the previous question, it is interesting to analyze how different the process model is in some steps when comparing it to the reality. The process flow projected by the university has been contrasted with the process obtained from the datalog, to understand where there are deviations in the process, focusing only in the activities related to obtain the permit to travel.
- **Question 3:** *Are there international travels that do not have the corresponding permit and still received a refund?*
  Finally, another important step to analyze for possible deviations is the one that involves the payment itself. The original process stipulates that all international travels need a permit to be able to receive a refund. It would be interesting to know if this is being followed correctly or there are cases that received a refund without a permit and how that affects the flow.

These three questions will be addressed keeping in consideration the main goal of the organization that collected the data and this is to improve the throughput time and ease of flow through the process in the future. This approach does not imply that the authors seek to provide a solution for TU/e to make their travel process better, but only that keeping that consideration in mind would help following discovery paths that the organization may be interested in.

2 Question 1: Do employees that stick to the process model have any advantage in terms of throughput time and ease of flow throughout the process?

2.1 Analysis Method

To analyze this objective, the data was filtered to eliminate all the variants that were considered “not desirable”. We deleted all the cases that had any rejection

activity in any step and all the cases that did not comply with the described process flow. Specifically, the following filters in Celonis[1] were applied:

1. Only cases that did not have any rejection activity from any role during the process were considered
2. Only cases started after 01-01-2018 were considered, as that year the organization changed its process and the objective of this question is to understand the optimality of the current one.
3. Only cases that comply with the described process flow were considered. The seven criteria used to filter in this step are described in Table 1.

Criteria	**Explanation**
"permit approved by administration" is followed anytime by **"start trip"**	The trip only starts when permit was already approved.
"declaration approved by administration" followed anytime by **"payment handled"**	The payment is only made if the administration approved the declaration.
"start trip" is never followed by **"permit approved by supervisor"** or **"permit final_approved by supervisor"**	It should not happen that the supervisor approved the trip after the trip started.
"payment hadled" never followed by **"start trip"**	The payment is not made before making the trip, as the transaction is not yet executed.
"end trip" is followed anytime by **"declaration submitted by employee"**	The declaration is only made when the trip is finished, as if someone do it before the end of the trip, not all of the expenses of the entire trip will be included.
"end trip" and **"start trip"** is never followed by **"permit final_approved by director"**	It should not happen that the director approves the permit after the trip started or finished.
"request payment" followed anytime by **"payment handled"**	To make the payment, the request should have happened before.

Table 1. Filter criteria used to eliminate the variants considered as "not desirable".

4. Only cases without the activity SEND REMINDER were included, because sending a reminder to an employee who was supposed to already had done a mandatory step in the process, is considered as not desirable.
5. An additional filter was applied to divide the process into two parts, a first part considering the activities between the start of the process and the FINAL APPROVAL for a permit and a second part considering the activities between the DECLARATION SUBMITTED BY THE EMPLOYEE and the end of the process. This filter was applied because different people may choose to do their paperwork at different times, so that should not be considered in the overall time.
6. After applying all the filters, a Celonis OLAP table (a table with multiple dimensions) was created with the total case count for every organizational unit. Then, by activating or deactivating the set of filters described from points 1-5, comparable results were obtained.

[1] Process mining software developed by the company Celonis, a world leader in the field.

2.2 Question-Driven Technical Analysis Results

After applying all the filters described in subsection 2.1, the process model shown in Fig. 1 was obtained using Celonis's "Conformance Checking" option, a tool capable of deducting a model from a group of process traces.

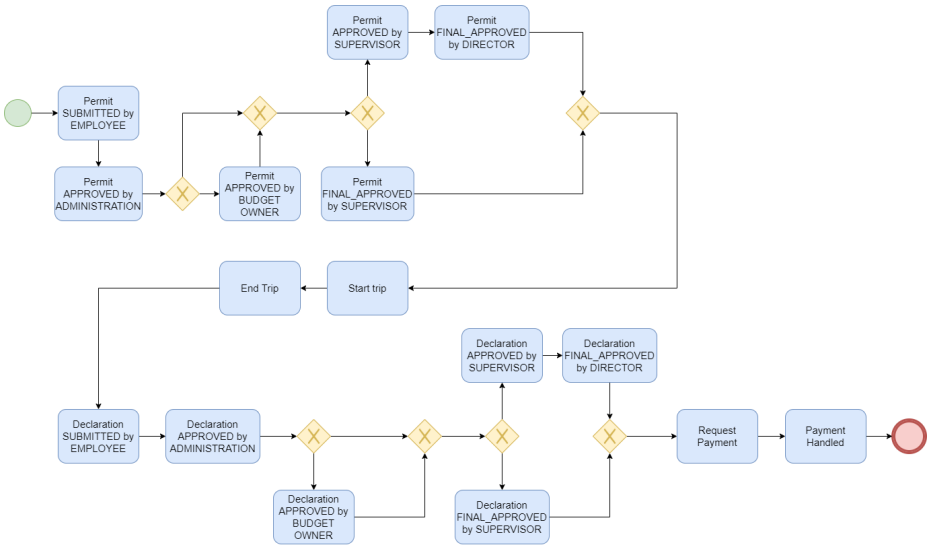

Fig. 1. International trips. Process model considered ideal.

Then, by observing the Process Explorer and the custom OLAP tables from Celonis, the following results were obtained:

- There are 2,269 cases that **comply with the ideal flow** from a total of 6,449, that corresponds to 35.18% of total cases. As only the cases generated since January first, 2018 were considered, our relevant total would be reduced to 4,897 cases. So the cases that comply with the ideal flow represent a total of 46.33% since 2018.
- There are **12 distinct variants** of the process (from the original 608 since 2018). They only differ in the steps of permit and declaration acceptance. The reason for this is that, as it was stated by the organization, in order for a permit or declaration to be approved, it needs to be reviewed by the budget owner and the supervisor or only one of the two if they are the same person. In addition, it is possible that in some cases, the director must also approve the request to be completely accepted. Therefore, there are several ways to do the process that are considered "correct" and that translates to several variants to be inside the "ideal path". This situation can be seen in the diagram of Fig. 2.

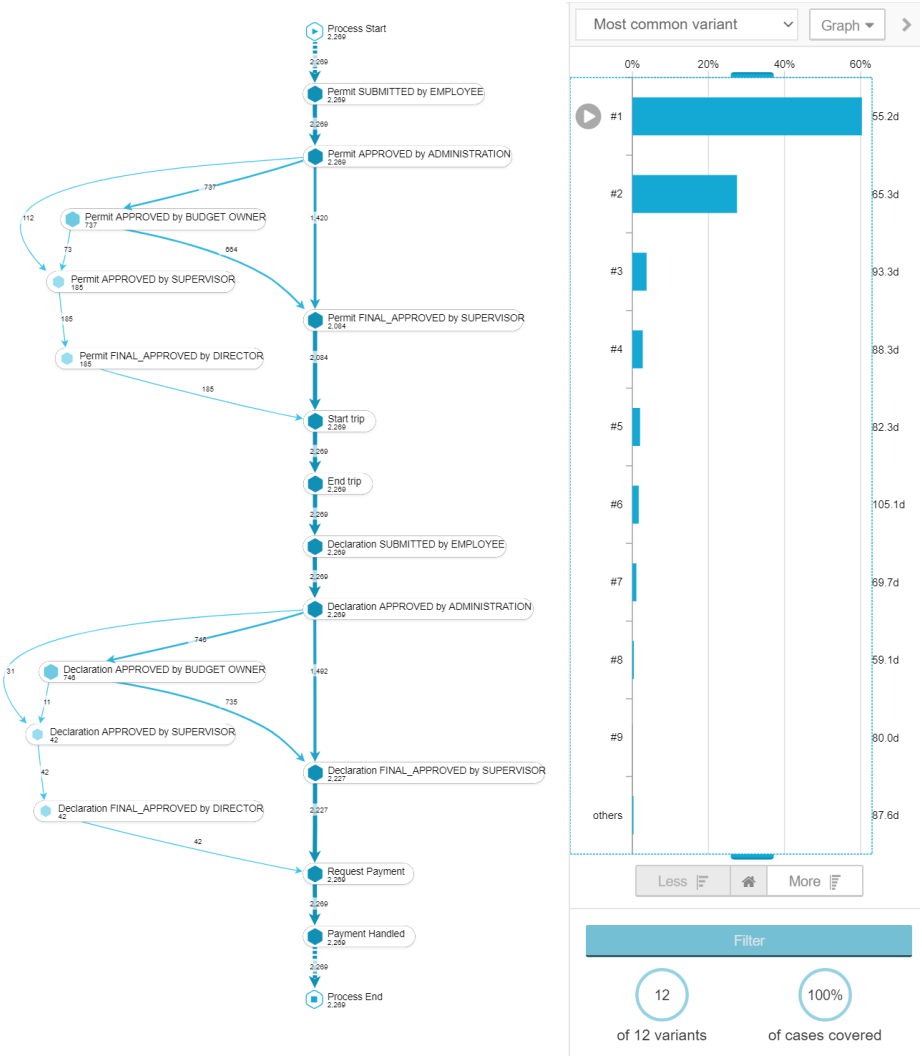

Fig. 2. Process considered ideal after applying all the filters mentioned in section 2.1. Visualized in Celonis's Variant Explorer.

Then, the percentage of cases that comply with the ideal flow for each organizational unit of the university was calculated. The detailed data used can be found in the Table A.1. in the appendix of this document. The results obtained are described bellow:

– Considering a percentage of compliance greater than 70% as "successfully complies", between 50% and 70% as "partially complies" and below 50% as "do not comply at all", organizational units 65475 and 65473 are the

only ones that, having a relevant number of cases (considered to be equal or greater than 10 cases), have a percentage that "successfully complies". Specifically, they have both 75% of compliance.
- Organizational units 65456, 65458 and 65454 are the ones that have the most total cases (889, 842 and 957 after 01-01-2018 respectively). They show compliance percentages notably low, considered as "Do not comply at all". Specifically they have, in order, compliance percentages of 45.22%, 38.84% and 24.24%.
- The average throughput time achieved by those employees who stick to the original permit process model is 3 days for those who get their permit approved by the supervisor and 6 days for the ones that get their permit approved by a director. This part considers the activities between PROCESS START and PERMIT FINAL APPROVED BY SUPERVISOR or DIRECTOR.
- The average throughput time for the last part of the process, between the DECLARATION SUBMISSION by the employee and the end of the process takes 12 days on average.
- The average throughput time for the permit part of the process increases by one day when the cases that don't get rejected but do not stick to the original process model are considered. That would mean 4 days for the permit process approved by the supervisor, 7 for the permit approved by the director and 13 for the declaration process.
- When taking into account the processes that don't follow the ideal path, a small group of them take up to 100 days to be completed.
- When comparing the results with the throughput time obtained for an equivalent optimal path for the 2017 process, the total time for the permit part is 9 days for the permit approved by the supervisor and 12 for the permit approved by the director. The declaration part of the process, on the other hand, used to take 67 days to be completed on average, which is considerably longer than the time needed since 2018.
- The average throughput time of organizational units 65473 and 65475 (the ones with relevant number of cases and compliance level "successfully complies") in the year 2018 was 76.11 days. On the other hand, organizational units 65456, 65458 and 65454 (the 3 with the highest total case number and comply level "Do not comply at all") had an average throughput time of 98.71 days in the same year. That makes a difference of 22.6 days or 22.9% less total throughput time average for the first group.

2.3 Business Owner Conclusions

Taking in consideration the question presented, the short answer would be yes, the employees who stick to the process tend to get their submissions approved in a shorter period of time than those who do not get rejected but whose process don't satisfy the original model either. If we take a closer look at the differences that exist between both scenarios, we will notice that the average throughput time for both the permit and the declaration part of the process increases only

by one day when not following the guidelines. Nevertheless, it is also true that the average for the cases that do not follow the ideal path is affected by a few cases in which the time for the employee to get an approval raised up to 100 days. As an additional reason to support the hypothesis that sticking to the original process model implies advantages for the employees, we can add that the changes made to the process from 2017 to 2018 actually reduced the average throughput time for those employees by 66% for the permits approved by a supervisor and by 50% for the ones approved by a director. In the case of the declarations, the throughput time was reduced by a considerable 82%. It is also crucial to mention that the cases sticking to the model have a total of 12 variants, a more manageable number, instead of the 608 that the process originally had since 2018. Lastly, it is important to note that organizational units that had more percentage of compliance with the ideal model had, in average, 22.6 days less throughput time than the ones with low compliance percentage.

3 Question 2: How different is the process flow projected by the university to obtain a travel permission, from the process flow obtained from the data log?

3.1 Analysis Method

The objective is to understand where there are deviations in the process to obtain the TRAVEL PERMIT APPROVAL needed to travel internationally, to identify where there is room for improvement in that process in particular and how much is the throughput time. In order to do this, the process flow proposed by the university was compared with the log, so an optimal process flow or at least an ideal one could be identified. After this, the variants found in the Process Explorer of Celonis were identified with their respective case count.

1. First, the ideal process flow proposed by the university was identified and was used in the whole report as seen in section 2.
2. Second, the main activities from the process were identified.
3. Then, a verification was done in order to see if there's one that does not meet the ideal process flow from the university.
4. The process paths found in the Process Explorer were identified in the Variant Explorer of Celonis with their respective case count.
5. All cases where considered for this objective. We are aware that the organization changed its process and the objectives in January of 2018 but it was put anyway in the analysis in order to represent the total amount of cases.

3.2 Question-Driven Technical Analysis Results

From the data collected to get the TRAVEL PERMIT APPROVAL, four main paths were obtained from the Process Explorer, indicated in appendix Fig. A.1

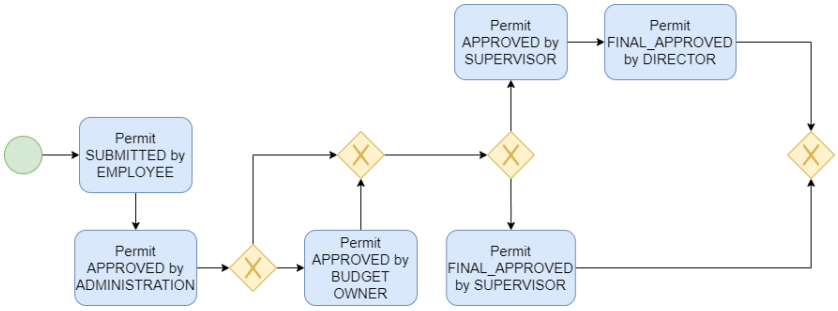

Fig. 3. Process flow proposed by the university for a travel permit approval, based on process model of Fig.2

and specified in Table 2. This refers to four paths used to get the TRAVEL PERMIT APPROVAL, regardless of the person who approved this permit (it may be the supervisor, the director or the budget owner). Then the percentage of cases of the flows indicated according to the Variant Explorer is shown in Table 2, and indicates how many cases go through each path to get the permit approved. This allows us to see how many cases go through the ideal path that has been defined by the university, and how many cases use other paths that are less effective regarding timing to finish the process of obtaining the TRAVEL PERMIT APPROVAL as seen in Table 3. The average throughput time in cases where the ideal path was followed was 3 days when the approval of the travel permit came from the supervisor. However, those cases where the final approval came from the director took up to 3 days more.

	Path 1	Path 2	Path 3	Path 4
% of cases	7%	40%	22%	8%
Activities	-Permit submitted by EMPLOYEE -Permit approved by PRE_APPROVER -Permit approved by SUPERVISOR	-Permit submitted by EMPLOYEE -Permit approved by ADMINISTRATION -Permit final approved by SUPERVISOR	-Permit submitted by EMPLOYEE -Permit approved by ADMINISTRATION -Permit approved by BUDGET OWNER -Permit final approved by SUPERVISOR	-Permit submitted by EMPLOYEE -Permit approved by ADMINISTRATION -Permit approved by SUPERVISOR -Permit final approved by DIRECTOR

Table 2. Most followed paths to get the TRAVEL PERMIT APPROVAL

Also, other odd process flows were identified that should not take place. They compose 23% of the total process flows as seen in Table 3.

3.3 Business Owner Conclusions

Considering all the analysis we found the following conclusions.

	Path 5	Path 6	Path 7	Path 8
% of cases	8%	3%	9%	3%
Activities	-Start trip -End trip -Permit submitted by EMPLOYEE	-Start trip -Permit submitted by EMPLOYEE -Permit approved by ADMINISTRATION -Permit final approved by SUPERVISOR -End trip	-Permit submitted by EMPLOYEE -Permit approved by SUPERVISOR	-Permit submitted by EMPLOYEE -Permit approved by SUPERVISOR -Permit final approved by DIRECTOR

Table 3. Less efficient paths to get the TRAVEL PERMIT APPROVAL

- First, the TRAVEL PERMIT APPROVAL process found in the Process Path 1 in Table 2 is no longer relevant since it is has ceased to happen after the beginning of 2018.
- The process paths 2, 3 and 4 in Table 2 are the ideal paths this process should follow according to what the university wants. The 23% remaining cases in Table 3 should be completely avoided in the future travel permit requests.
- Also, it has been found out that most of the TRAVEL PERMIT APPROVAL by the director have a higher chance to get a REJECTED DECLARATION or a declaration process that can take up to 12 days. We believe that may be possible that those travels as do not go through the supervisor review are not properly drafted and this generates problems at the moment of filling up the declaration form. Overall, every activity that goes through the director's approval takes up to 3 days more than the other reviews made by the supervisor, the pre-approver and the budget owner.
- Furthermore, it was discovered that a small number of cases get a reimbursement without the mandatory travel permit. This issue needs further analysis and it will be analyzed in section 4.

4 Question 3: Are there international travels that do not have the corresponding permit and still received a refund?

4.1 Analysis Method

The objective is to characterize and quantify international travels where refund was made without the approved travel permission, in order to understand when, where and why there are cases that do not follow all the sub-processes described in Process Flow of section 2, but still receive refund.

This analysis is focused on the sub-process related to the "Payment Handled" activity, which occurs mostly after the travel declaration is approved (Fig.4), and in particular, the cases in which there is no PERMIT APPROVED activity, which occurs mostly after the travel permit is requested. It must be noted that

what was stated above is independent of the role that approved the permit, since in the process, the SUPERVISOR, BUDGET OWNER, ADMINISTRATION and DIRECTOR have the faculty to do so.

In order to guide and focus the analysis of the research, the following questions were proposed to be resolved:

- "What is the percentage of cases where the situation described above occurs?" It is important to know if there is a significant percentage regarding to the main process, that have this behavior.
- "Which is the role that approves most declarations without a permit?" It is interesting to know whether this practice is specific to a particular person, or if her role entails the authorization to do so.
- "Are the refunds made after or before the trip starts?"
- "What is the average throughput time for this type of flow?" With this time in our hands we can find if there is a relationship between those who take longer and refunds without permits, since a longer time could mean more pressure on the worker, leading to a higher percentage of error.
- "Is this behavior typical for any organizational unit?" If this is the case for any unit, it would be interesting to guide the research in this direction to discover why the ideal process is not being followed correctly in this units.

For this research, Celonis was used to model and study the process. In particular, the Process Explorer tool was used because it allows to visualize the process flows in an orderly manner and provides relevant numbers of the same, whether it is the number of times an activity occurs, average time, different connections, etc. The Variant Explorer tool was used to analyze different variants of the process, grouping them according to the percentage of occurrence; and finally, the Process Overview tool, was used to obtain different general metrics of the process under study, either the frequency of daily cases, average throughput time, among others.

4.2 Question-Driven Technical Analysis Results

As mentioned above in methods of analysis, it was decided to eliminate from the original flow all activities related to an approved permit. Then, with the Process Explorer tool, a new process map was discovered which is composed of 448 cases, equivalent to 7% of all international trips. Of the latter, there were 448 travel declarations of which only 426 were approved, with the detail that 82 cases were rejected in the first instance and they had to declare it again where the permit was finally approved.

Along these lines, it was noted that of the 426 approved declarations, only 422 were refund, i.e. 6.5% of international travels. With this we can note that, firstly, not all declarations without permission are approved in the first or second instance, and secondly, having been approved a declaration, there is 1 case in which no refund was made.

On the other hand, it was observed that not all refunds from this path followed the steps of the process proposed in section 2, since 381 were made before

the person travelled, whereas this activity should (ideally) be made after the trip. Going deeper into the analysis, it was possible to realize that exists a happy-path under the conditions mentioned before. This path can be seen in Fig. 4.

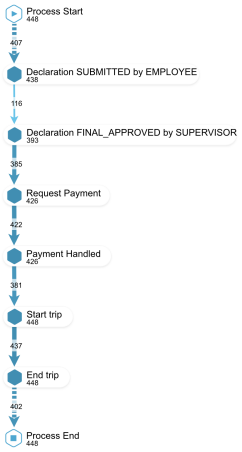

Fig. 4. Flow without approved travel permits.

This occurs 111 times, which is equivalent to 24.78% of the case studies. At first glance it seems to be the correct path, however, it could be noticed that the average throughput time is much higher than the average of the original process, 194 versus 387 days respectively. In Fig. 5 can be seen how the cases are distributed according to the number of days in which it develops:

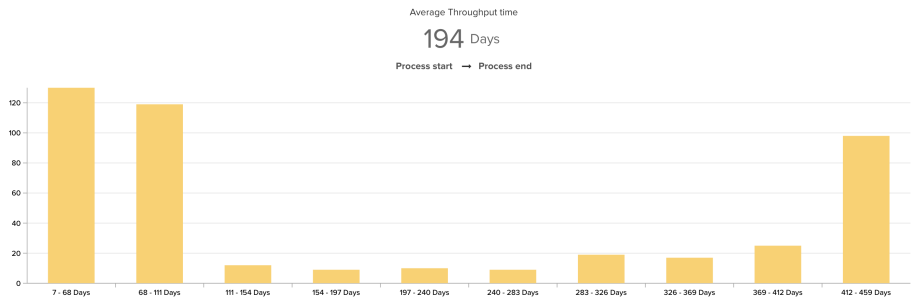

Fig. 5. Flow without approved travel permits.

It can be noted that there is no clear distribution and a large part of the cases are resolved quickly (less than 111 days), while others are resolved very slowly (more than 412 days). From the above, it could also be noted that it takes more than twice (on average) as long as the original process, with the latter taking an average of 75 days per case.

The latter suggests that there is probably a phase in the process where some bottleneck is slowing down the flow of the process. For this reason it was decided to analyze the times between process activities with the Process Explorer tool. The 448 case studies were taken and the frequency visualization was used to create the model below in Fig. 6.

Analyzing this model, you can see that all activities take about 4 days (on average) to move to the next one, except for Payment Handled, which takes 194 days to move to Start Trip. Consequently, it is possible to say that there is a "bottleneck" since the employee sends his travel statement at least 190 days in advance, which tends to increase the average time of the process in question too much.

With regard to the statements approved according to the role, it was determined that the majority are approved by the supervisor with 393 cases, and then followed by the director with 34 cases, for a total of 427 cases as shown in the Table 4. In addition, when the travel permit is not approved, the director approves 7.9% of the statements; however, when the process follows the happy-path proposed in section 2, the director approves only 3.9% of the statements. Therefore, for this process you can see that the director's activity increases approximately twice as much.

Role	No. of approved cases (without permission)	No. of approved cases (with permission)
Supervisor	393	5723
Director	34	235

Table 4. Less efficient paths to get the TRAVEL PERMIT APPROVAL

Approaching from the perspective of the organizational units of the university, we wanted to know which ones were involved in this study process. In order to analyze the distribution of cases by unit, the data were grouped by the attribute CASE-PERMIT ORGANIZATIONAL ENTITY. It could be seen that this behavior is typical of a particular unit, since of the 27 organizational units that exist, this variant of the original process only occurs in the organizational unit 65454. So it was decided to analyze this particular unit.

For this, from the original process, only the events with a CASE-PERMIT ORGANIZATIONAL ENTITY attribute equal to 65454 were left. This resulted in 957 cases, equivalent to 15% of the total. Therefore, taking into account the results above, 448 out of 957 cases in this unit are carried out without obtaining a permit while the other 509 cases proceed on a regular basis as indicated in the original process.

Therefore, in unit 65454, in 53.2% of the cases it is operated according to the original process while in 46.8% it is done without requiring the approved permit. Since the two processes have a very similar frequency within the unit, it follows that these cases are not exceptions. Finally, it can be stated that this unit is in charge of these special cases, while at the same time they carry out procedures according to the original model.

4.3 Business Owner Conclusions

In conclusion, it can be stated that there are international travels that receive refund, even if they never obtained the due permission. However, in relation to the total of travels, the resulting amount is not as relevant as originally thought. It should be mentioned that there is an academic unit in the university that carries out two different models to make the refund with similar frequency. Due to the existence of the latter, it can be inferred that there is a certain bias in the unit to determine the refunds, since approximately half of the cases do not have the travel permit that is required for international trips and yet are refunded. On the other hand, in these cases the money is returned, on average, more than half a year before the start of the trip, which could be interesting to analyze in more detail. Finally, it is recommended to investigate this unit in order to establish a regular channel, as other units do, since in 425 cases the money is refunded without being in line with the normal process; and also to rule out the possibility of some fraud.

5 Conclusions

In short terms, the answer to the first question is yes, there is a notable advantage in throughput time when comparing employees that stick to the process model to the one that do not. Specifically, they take 22.6 less days (22.9%) total time average to complete the process and they have only 12 variants, a manageable number, instead of the astonishing high original number of 608 different variants since 2018. Regarding the question number two the answer is yes, there is a difference from what is expected as an ideal flow and what happens in the process obtained from the cases, regarding the obtained TRAVEL PERMIT APPROVAL. This delay in the total throughput time to obtain the permit is up to 23% of the cases and it is especially significant when the flow goes through the approval of the director, as it get up to 3 days more to be approved. This is it also connected with a higher percent of DECLARATION REJECTED and declarations resubmitted, adding even more time, up to 12 days. Is strongly suggested that the travel permits that go through the director are always previously approved by the supervisor as is expected this way to reduce the time it takes to claim the declaration and subsequent reimbursement. Finally, the third question can be easily answered with a yes. There is an important quantity of cases that are refunded even though they do not have an approved permit to travel, around a 7% of the total. Those cases came from only one organizational unit, the unit 65454. The cases of this unit that do not have an approved permit represents 46.8% of their total cases so it can be inferred that these are not exceptions. Also, it can be noted that the other 53.2% flows by the regular path. The final conclusion is that following "the rules", the designed ideal path, **does** have a significant benefit in the execution of the process for the university.

Acknowledgments

The authors of this report want to thank the collaboration of the Celonis Academic Alliance, in particular to Jerome Geyer-Klingeberg, who provided us access to Celonis software and for all his support with technical issues.

A Appendix

Organizational Unit	Ideal count [num]	Total count [num]	Percentage ideal
organizational unit 65488	1	1	100,00%
organizational unit 65478	1	1	100,00%
organizational unit 65462	1	1	100,00%
organizational unit 65475	12	16	75,00%
organizational unit 65473	9	12	75,00%
organizational unit 65482	5	7	71,43%
organizational unit 65484	2	3	66,67%
organizational unit 65469	30	50	60,00%
organizational unit 65471	4	7	57,14%
organizational unit 65467	10	18	55,56%
organizational unit 65460	243	480	50,63%
organizational unit 65486	1	2	50,00%
organizational unit 65465	4	8	50,00%
organizational unit 65461	37	77	48,05%
organizational unit 65468	6	13	46,15%
organizational unit 65457	138	304	45,39%
organizational unit 65456	402	889	45,22%
organizational unit 65470	18	40	45,00%
organizational unit 65464	137	305	44,92%
organizational unit 65459	212	493	43,00%
organizational unit 65466	78	220	35,45%
organizational unit 65455	352	1081	32,56%
organizational unit 65477	1	4	25,00%
organizational unit 65454	232	957	24,24%
organizational unit 65458	327	1432	22,84%
organizational unit 65480	2	9	22,22%
organizational unit 65472	4	19	21,05%

Table A.1. Ideal cases registered for each Organizational Unit. Considering a percentage of compliance greater than 70% as “successfully complies” (green), between 50% and 70% as “partially complies” (yellow) and below 50% as “do not comply at all” (red)

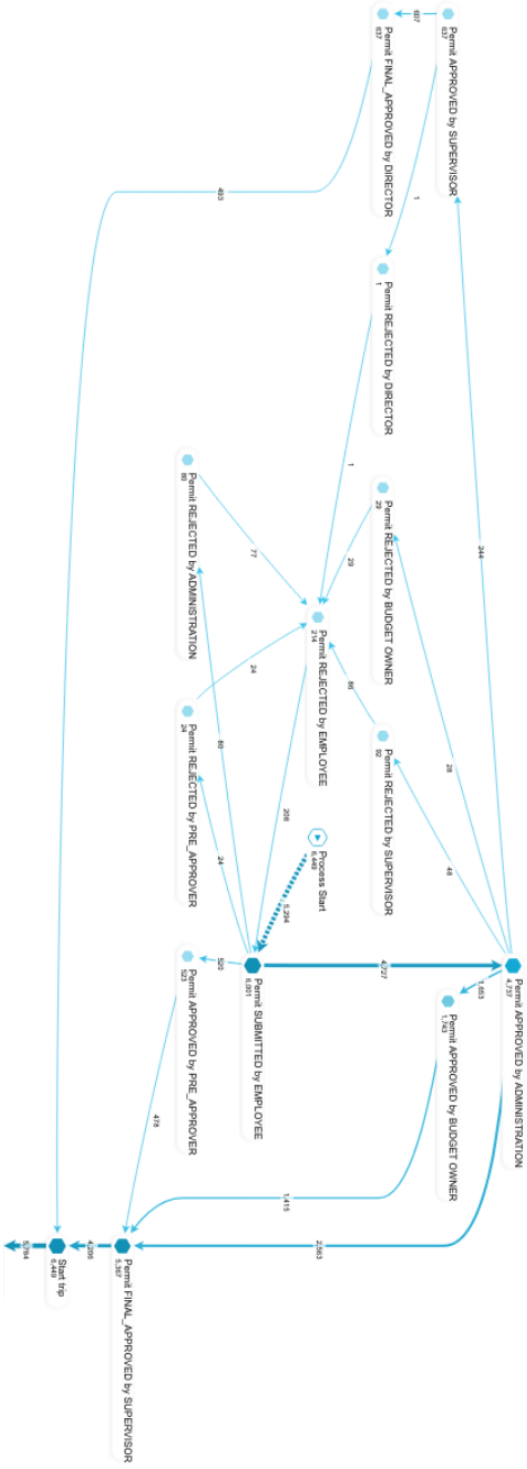

Fig. A.1. Process map with the paths that allow the TRAVEL PERMIT APPROVAL.