Neural networks have seen a spike in popularity in natural language processing in recent years. They consistently outperform the traditional methods and require less human labor to perfect as they are trained unsupervised on large text corpora. However, these corpora may contain unwanted elements such as biases. We inspect multiple language models, primarily focusing on a Czech monolingual model – RobeCzech. In the first part of this work, we present a dynamic benchmarking tool for identifying gender stereotypes in a language model. We present the tool to a group of annotators to create a dataset of biased sentences. In the second part, we introduce a method of measuring the model's perceived political values of men and women and compare them to real-world data. We argue that our proposed method provides significant advantages over other methods in our knowledge. We find no strong systematic beliefs or gender biases in the measured political values. We include all the code and created datasets in the attachment.